ICAROB 2014

THE PROCEEDINGS OF INTERNATIONAL CONFERENCE ON ARTIFICIAL LIFE AND ROBOTICS

January 11-13, 2014
Compal Hall, Oita, JAPAN
International Meeting Series

Editor-in-Chief
Masanori Sugisaka
Editors: Takao Ito, Ju-Jang Lee, Yingmin Jia
ISBN 978-4-9902880-8-2

The Proceedings of the International Conference on

ARTIFICIAL LIFE AND ROBOTICS

(ICAROB 2014)

January 11-13, 2014
Compal Hall Oita, Japan

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

Contents

ORGANIZED

International Steering Committee of International Conference on Artificial Life and Robotics (ICAROB) and ALife Robotics Corporation Ltd.

CO-ORGANIZED BY

IEEE Fukuoka Section (IEEE Robotics and Automation Society, Fukuoka Section, Japan)
Chinese Association for Artificial Intelligence (CAAI, P. R. China)
University of Sultan Zainal Abidin (UniSZA)

ADVISORY COMMITTEE CHAIR

Moshe Kam (Drexel University, Former IEEE President, USA)

ADVISORY COMMITTEE

F. Harashima (The University of Tokyo, Japan)
B. Eisenstein (Drexel University, Former IEEE President, USA)Eisenst
H. Kimura (RIKEN, Japan)
M. Tomizuka (University of California Berkeley, USA)
M. Kam (Drexel University, Former IEEE President, USA)
A. Grzech (Wroclaw University of Technology, Poland)
K. Aihara (The University of Tokyo, Japan)
D. J.G. James (Coventry University, United Kingdom)
J. Johnson (The Open University, UK)
K. Kyuma (Mitsubishi Electric Corporation, Japan)
S. Rasmussen (University of Southern Denmark, Denmark)
J. M. Epstein (The Johns Hopkins University, USA)
J. Świątek (Wroclaw University of Technology, Poland)
P. Kalata (Drexel University, USA)
P. Oh (Drexel University, USA)
R. Fischl (Drexel University, USA)
T. Fukuda (Nagoya University, Japan)

GENERAL CHAIRS

Yingmin Jia (Beihang Univesty)
Takao Ito (Ube National College of Technology)
Ju-Jang Lee (KAIST)

VICE GENERAL CHAIRS

C. Zhang (Tinghua Univesity, P. R. China)

H. H. Lund (Technical University of Denmark, Denmark)
J. L. Casti (International Institute for Applied Systems Analysis, Austria)
J.M. Lee (Pusan National University, Korea)
Y. G. Zhang (Academia Sinica, P. R. China)

PROGRAM CHAIRMAN

Makoto Sakamoto (University of Miyazaki, Japan)

SUB PROGRAM CHAIR

M. Oswald (The Vienna University of Technology, Austria)

HONORARY GENERAL CHAIR

M. Sugisaka (ALife Robotics Co., Ltd., Japan,
University of Sultan Zainal Abidin (UniSZA),
University of Malaysia-Peris, Malaysia)

INTERNATIONAL ORGANIZING COMMITTEE

E. Hayashi (Kyushu Institute of Technology, Japan)
H. Desa (University of Malaysia, Perlis, Malaysia)
H. Suzuki (The University of Tokyo, Japan)
H. Furutani (The University of Miyazaki)
H. Matsuno (Yamaguchi University, Japan)
J. Wang (Beijing Jiaotong University, P. R. China)
K. Shimohara (Doshisha University, Japan)
K. Ohtsu (Tokyo University of Marine Science and Technology)
K-L. Su (National Yunlin University of Science and Technology, Taiwan)
M. Kubo (National Defense Academy of Japan, Japan)
M. Obayashi (Yamaguchi University, Japan)
M. Yamashita (Kyushu University, Japan)
M. Rajiv (New Jersey Institute of Technology)
M. Rizon (University of Sultan Zainal Abidin, Malaysia)
P. Sapaty (Ukrainian Academy of Science, Ukraine)
Q. Yanbin (Harbin Institute of Technology, P. R. China)
S. Ishikawa (Kyushu Institute of Technology, Japan)
T. Kohno (The University of Tokyo, Japan)
T. Hattori (Kagawa University, Japan)
T. S. Ray (University of Oklahoma, USA)
V. Berdonosov (Komsomolsk-on-Amur State University of Technology, Russia)
Y. Yoshitomi (Kyoto Prefectural University, Japan)

INTERNATIONAL PROGRAM COMMITTEE

A. Nakamura (AIST, Japan)
A. Selamat(University of Technology of Malaysia(UTM), Malaysia)
B. Fu(Shanghai Jiatong University, P. R. China)
D. Ai (University of Science & Technology Beijing, China)
E. Joelianto (Bandung Institute of Technology, Indonesia)
F. Dai (Tianjin University of Science & Technology, P. R. China)
H. Ogai (Waseda University, Japan)
H. Yanagimoto (Osaka Prefecture University, Japan)
H. Umeo (Osaka Electro-Communication University, Japan)
H. Iizuka (Osaka University, Japan)
H. Zhao (Shanghai Institute of Technology, P. R. China)
H. Abbass (University of New South Wales and ADFA, Australia)
I. Tanev (Doshisha University, Japan)
J. Zhao (Beijing Jiaotong University, P. R. China)
K. E. Merrick (University of New South Wales and ADFA, Australia)
K. Kobayashi (Aichi Prefectural University)
K. Kurashige (Muroran Institute Technology, Japan)
K. Sugawara (Tohoku Gakuin University, Japan)
L. Pagliarini (Technical University of Denmark, Denmark)
M. Eaton (University of Limerick, Ireland)
M. Yokomichi (The University of Miyazaki, Japan)
M. Tabuse (Kyoto Prefectural University, Japan)
M. Hatakeyama (university of Zurich, Switzerland)
M. Svinin (Kyushyu University, Japan)
M. Watanabe (Kagoshima University, Japan)
N. Mokhtar (University of Malaya, Malaysia)
P. S. Sapaty (National Academy of Sciences of Ukraine, Ukraine)
S. Ikeda (The University of Miyazaki, Japan)
S. S. Joshi (College of Engineering, UCLA, USA)
S. Omatu (Osaka Institute of Technology, Japan)
S-M Chen (National Taichung University of Education, Taiwan)
T. Zhang (Tsinghua University, P. R. China)
T. Iwamoto (Mitsubishi Electric Corporation, Advanced Technology R&D Center, Japan)
T. Katayama (the University of Miyazaki, Japan)
T. Yoshinaga (Tokuyama National College of Technology)
Y. Fan (University of California-Davis, USA)
Y. Liu (University of Aizu, Japan)
Y. I. Cho (The University of Suwon, Korea)

LOCAL ARRANGEMENT COMMITTEE

M. Sakamoto (University of Miyazaki, Japan)
M. Sugisaka (ALife Robotics Co., Ltd., Japan,
University of Sultan Zainal Abidin (UniSZA),
University of Malaysia-Peris, Malaysia)
Takao Ito (Ube National College of Technology)

HISTORY

This symposium was founded in 1996 by the support of Science and International Affairs Bureau, Ministry of Education, Culture, Sports, Science and Technology, Japanese Government. Since then, 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). We changed this symposium name as The International Conference on Artificial Life and Robotics newly. This conference invites you all.

OBJECTIVE

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 will discuss new results in the field of artificial life and robotics.

GENERAL SESSION TOPICS

GS1 Artificial intelligence & Complexity
GS2 Reinforcement Learning System and genetic programming
GS3 Human-machine cooperative systems & Human welfare robotics
GS4 Neural Networks
GS5 Pattern recognition
GS6 Robotics I
GS7 Robotics II

ORGANIZED SESSION TOPICS

OS1 Intelligent Control
OS2 Environment Navigation and Localization
OS3 Facial Expression Analysis, Music Recommendation and Augmented Reality
OS4 Applications in Kansei Engineering
OS5 Analysis of Randomized Algorithms Intelligence in Biological Systems
OS6 Empirical research on Network and MOT
OS7 Intelligence in Biological Systems
OS8 Software Development Support Method
OS9 Foundation of computation and its Application

COPYRIGHTS

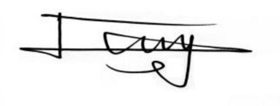
Accepted papers will be published in the proceeding of ICAROB and some of high quality papers in the proceeding will be requested to re-submit their papers for the consideration of publication in an international journal ARTIFICIAL LIFE, COMPLEXITY AND ROBOTICS under negotiation. All correspondence related to the conference should be addressed to ICAROB Office.

ICAROB Office

ALife Robotics Corporation Ltd.
3661-8 Oaza Shimohanda, Oita 870-1112, JAPAN
TEL/FAX：+81-97-597-7760
E-MAIL： **icalr@alife-robotics.co.jp**
Home Page：http://alife-robotics.co.jp/

MESSAGES

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



Yingmin Jia

General Chair of ICAROB

It is my great pleasure to invite you to the International Conference on Artificial Life and Robotics (ICAROB), in Oita City, Oita, Japan from Jan. 11th to 13th, 2014.

ICAROB 2014 develops from the AROB that was created in 1996 by Prof. Masanori Sugisaka and celebrated her birthday of eighteen years old in 2013. Doubtless, new mission and big challenges in the field of artificial life and robotics will promote ICAROB to start a new stage 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 Oita City and wishing you enjoy your stay in Japan.

Takao Ito
General Chair
(Professor of Ube National College of Technology, Japan)

Takao Ito

General Chair of ICAROB

It is my great honor to invite you all to The International Conference on Artificial Life and Robotics (ICAROB 2014).This Conference is changed as the old symposium from the first (1996) to the Eighteenth I am pleased to welcome you to the 2013 International Conference on Artificial Life and Robotics (ICAROB) in the wonderful city of Oita, Japan

The ICAROB has long history. The former organization of the ICAROB was developed under the strong leadership of the President, Professor. Masanori Sugisaka, the father of AROB. We gathered many researchers, faculty members, graduate students from all over the world, and published numerous high-quality proceedings and journals every year.

Over the years, dramatic improvements have been made in the field of artificial life and its applications. The ICAROB has becoming the unifying the exchange of scientific information on the study of man-made systems that exhibit the behavioral characteristic of natural living systems including software, hardware and/or 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, good ideas, and valuable information about artificial intelligence, complex systems theories, robotics, management of technology, etc.

In order to provide an outstanding technical level for the presentations at the conference, we have invited more than 60 distinguished experts in the field of artificial life in the organizing committee and program committee. We will have 23 sessions during 3 days of conference, including 3 invited sessions.

The conference site is the Compal Hall, one of the finest congress centers in Oita. It is situated near the center of the city. You can find many fantastic scenic spots and splendid hot-springs. Enjoy your stay and take your time to visit the city of Oita.

I am looking forward to meeting you in Oita during 2013 ICAROB and to sharing a most pleasant, interesting and fruitful conference

Ju-Jang Lee

General Chair of ICAROB

Ju-Jang Lee

General Chair

(Professor, KAIST)

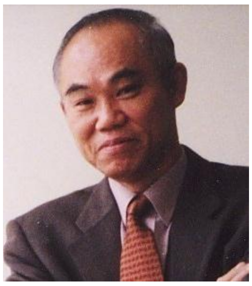
The First International Conference on Artificial Life and Robotics (ICAROB) will be held in Oita City, Oita, Japan from Jan. 11th to 13th, 2014. This year's Conference will be held amidst the high expectation of the increasingly important role of the new interdisciplinary paradigm of science and engineering 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 for me as a General Chair of the 1st ICAROB 2014 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 1st ICAROB.

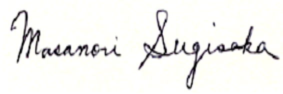
I'm looking forward to meeting you at the 1st ICAROB in Oita City and wishing you all the best.



Masanori Sugisaka

Honorary General Char

**(Professors,
University of Sultan Zainal Abidin(UniSZA), Malaysia,
University of Malaysia -Perlis and President,
ALife Robotics Co., Ltd. , Japan)**



Masanori Sugisaka

Honorary General Chair of ICAROB

It is my great honor to invite you all to The International Conference on Artificial Life and Robotics (ICAROB 2014).

This Conference is changed as the old symposium from the first (1996) to the Eigteenth(2013) anually 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 are inherited also from now. From now on, ALife Robotics Corporation Ltd. is in charge of management. The future of The ICAROB 2014 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 Coference (ICAROB 2014) 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.

TIME TABLE (1/11)

1/11	Room301	Room304	Room305
9:30~	Registration		
10:00-11:20		GS3 (4) Chair M. Rizon	GS7 (4) Chair J. Wang
11:20-11:40	Coffee break		
11:40-12:00	Opening Ceremony (room300)		
12:00-13:00	Lunch		
13:00-14:00	IS (room300) Chair J. M. Lee		
14:00-14:20	Coffee break		
14:20-16:00		GS1(5) Chair M. Kubo	GS4(4) Chair A. Selamata (will be end at 15:40)
16:00-16:20	Coffee break		
16:20-18:00		GS5(5) Chair S. Ishikawa	GS 2(5) Chair K. Kobayashi

GS1 Artificial intelligence & Complexity
GS2 Reinforcement learning system & Genetic programing
GS3 Human-machine cooperative systems & Human welfare robotic
GS4 Neural Networks
GS5 Pattern recognition Ⅰ
GS6 Robotics Ⅰ
GS7 Robotics Ⅱ

OS1 Intelligent Control
OS2 Environment Navigation and Localization
OS3 Facial Expression Analysis, Music Recommendation and Augmented Reality
OS4 Applications in Kansei Engineering
OS5 Analysis of Randomized Algorithms
OS6 Empirical research on Network and MOT
OS7 Intelligence in Biological Systems
OS8 Software Development Support Method
OS9 Foundation of computation and its Application

TIME TEBLE (1/12)

1/12	Room301	Room304	Room305
9:30～	Registration		
10:00-11:20	PS1(6) Chair D. Hazry	OS5(3) Chair H. Furutani (will be end at 11:00)	OS8(4) Chair T. Katayama
11:20-13:00	Lunch		
13:00-14:40	PS1(6)	OS1(5) Chair Y. Jia	OS3(5) Chair Y. Yoshitomi
14:40-15:00	Coffee break		
15:00-17:00	PS1(6)	OS2(6) Chair J. M. Lee	OS4(5) Chair: T. Hattori (will be end at 16:40)
18:00-20:00	Banquet: HOTEL HOKKE CLUB OITA		

TIME TABLE (1/13)

1/13	Room301	Room304	Room305
9:30～	Registration		
10:00-12:00		GS6(5) Chair F. Dai (will be end at 11:40)	OS7(6) Chair J. Nishii
12:00-13:00	Lunch		
13:00-14:40		OS6(5) Chair: T. Ito	OS9(4) Chair: M. Sakamoto (will be end at14:20)

The International Conference on

ARTIFICIAL LIFE AND ROBOTICS 2014

(ICAROB 2014)

January 11 (Saturday)

Room 300: 11:40-12:00

Opening Ceremony

Chair: M. Sakamoto (University of Miyazaki, Japan)

Welcome Addresses

1. General Chairman of ICAROB	Y. M. Jia (Beihang University, China)
2. General Chairman of ICAROB	T. Ito (Ube National College of Technology, Japan)
3. General Chairman of ICAROB	J. J. Lee (KAIST, Korea)
4. Honorary General Chairman of ICAROB	M. Sugisaka(ALife Robotics Co., Ltd. Japan)

January 12 (Sunday)

HOTEL HOKKE CLUB OITA

18:00-20:00

Banquet

Chair: M. Oswald (The Vienna University of Technology, Austria)

Welcome Addresses

Y. M. Jia (Beihang University, China)

H.H.Lund(Denmark Technical University of Denmark,Denmark)

L.Pagliarini(Denmark Technical University of Denmark,Denmark)

TECHNICAL PAPER INDEX

January 11 (Saturday)

Room 300

13:00-14:00 Invited Session
Chair: J. M. Lee(Pusan National University, Korea)

IS1-1 *Playware Research – Methodological Considerations*
Henrik Hautop Lund (Centre for Playware, Technical University of Denmark, Denmark)

IS1-2 *ALife for Real and Virtual Audio-Video Performances*
Luigi Pagliarini (Centre for Playware, Technical University of Denmark, Denmark
Academy of Fine Arts of Macerata,Italy)
Henrik Hautop Lund (Centre for Playware, Technical University of Denmark, Denmark)

IS1-3 *Heart-pulse Biofeedback in Playful Exercise using a Wearable device and Modular Interactive Tiles*
Tomoya Shimokakimoto (University of Tsukuba, Japan)
Henrik H. Lund (Technical University of Denmark, Denmark),
Kenji Suzuki (University of Tsukuba/ JST, Japan)

January 11 (Saturday)

9:30- Registration

Room 304

10:00-11:20 GS3 Human-machine cooperative system & Human welfare robotics
Chair: Mohamed Rizon (Universiti Sultan Zainal Abidin (UniSZA), Malaysia)

GS3-1 *Development of Measurement System for Quantitative Evaluation of Skillfulness of Lower Extremities*
Kazunori Yamazaki, Donggun Kim, Yoshifumi Morita, Hiroyuki Ukai
(Nagoya Institute of Technology, Japan),
Kenji Kozakai, Satoru Shibata, Shigenori Onishi, Akihiro Ito, Daisuke Mizuno
(Sanyo Machine Works, Ltd., Japan),
Tatsuya Hirai, Haruna Takeda, Yuka Sugiura (Tanakakai Nishio Hospital, Japan)

GS3-2 *Glove-Based Virtual Interaction for the Rehabilitation of Hemiparesis Stroke Patient*
Khairunizam WAN, Aswad A.R, Rashidah Suhaimi, Nazrul H. ADNAN, D. Hazry,

Zuradzman M. Razlan, Syed Faiz Ahmed (Universiti Malaysia Perlis, Malaysia)

GS3-3 *Measuring system of therapist's guiding motion for standing up training of patient with hemiplegia after stroke and analysis results of patient's motion*
JungTang Wang, Kazunori Yamazaki, Yoshifumi Morita, Noritaka Sato
(Nagoya Institute of Technology, Japan)
Hirofumi Tanabe (Self-Defense Force Central Hospital, Japan)

GS3-4 *Image Segmentation of Coronary Artery Plaque Using Intuitionistic Fuzzy C-Means Algorithm*
Zahra Rezaeia, Ali Selamata, Mohd Shafry Mohd Rahima,
Mohammed Rafiq Abdul Kadirb (Universiti Teknologi Malaysia, Malaysia)

14:20-16:00 GS1 Artificial intelligence & Complexity

Chair: M. Kubo (National Defense Academy of Japan)

GS1-1 *Cognitive approach to Computer Go programming*
N. Tenys, I. Tanev, K. Shimohara (Doshisha University, Japan)

GS1-2 *Circulative Narrative Generation Based on the Mutual Transformation between Narrative Conceptual Structures and Music in the Integrated Narrative Generation System*
Taisuke Akimoto, Takashi Ogata (Iwate Prefectural University, Japan)

GS1-3 *Dynamic Analysis of Tensegrity Systems Subject to Arbitrary Joint Constraints*
Youngsu Cho, Joono Cheong (Korea University, Korea)

GS1-4 *Road map generation from Smartphone data*
Masao KUBO, Chau viet DAN, Hiroshi SATO, Akira NAMATAME
(National Defense Academy of Japan)

GS1-5 *Zigzag-Perceptually Important Points for Financial Time Series Indexing*
Chawalsak Phetchanchai, Ali Selamat , Md Hafiz Selamat
(University Teknologi Malaysia, Malaysia)

16:20-18:00 GS5 Pattern recognition

Chair: S. Ishikawa (Kyushu Institute of Technology, Japan)

GS5-1 *Recognizing a Road Environment Using Multiple-window Bag of Features*
Shou Morita, Joo Kooi Tan, Hyoungseop Kim, Seiji Ishikawa
(Kyushu Institute of Technology, Japan)

GS5-2 *Recognized Face Tracking for CONBE Robot*
Sakmongkon Chumkamon, Eiji Hayashi (Kyushu Institute of Technology)

GS5-3 *Eye detection Using Composite Cross-Correlation form Face Images in Varied Illumination*
Kutiba Nanaa, Mohamed Rizon, Mohd Nordin Abd Rahman
(Universiti Sultan Zainal Abidin (UniSZA), Terengganu, Malaysia)

GS5-4 *Development of an autonomous-drive personal robot "An object recognition system using BoF and SOM"*
Keisuke Ito, Eiji Hayashi (Kyusyu Institute of Technology, Japan)

GS5-5 *Comparison of Feature Detectors for Obstacles Detection*
Shaohua Qian, Joo Kooi Tan, Hyoungseop Kim, Seiji Ishikawa
Takashi Morie, (Kyushu Institute of Technology, Japan)
Takashi Shinomiya(Japan University of Economics, Japan)

Room 305

10:00-11:20 GS7 Robotics Ⅱ

Chair: Jiwu Wang (Beijing Jiaotong University, China)

GS7-1 *Decision Making System of Robots introducing a Re-construction of Emotions Based on Their Own Experiences*
Shogo Watada, Masanao Obayashi, Takashi Kuremoto, Shingo Mabu
(Yamaguchi University, Japan)
Kunikazu Kobayashi (Aichi Prefectural University, Japan)

GS7-2 *Visual Servoing and Sound Localization in a Surveillance Robot*
So-Yeon Park, Yeoun-Jae Kim, Ju-Jang Lee (KAIST, Korea)

GS7-3 *Study on the improvement of flexibility for an industrial robot based on machine vision*
Jiwu Wang, Xianwen Zhang, Weining Zhang (Beijing Jiaotong University, China)
Sugisaka Masanori(Nippon Bunri University, Japan)

GS7-4 *Improved Map Generation by Addition of Gaussian Noise for Indoor SLAM using ROS*
Khairul Salleh Mohamed Sahari, Barry Loh Tze Yuen
(Universiti Tenaga Nasional,Malaysia)

14:20-15:40 GS4 Neural Networks

Chair: Ali Selamat (University Teknologi Malaysia, Malaysia)

GS4-1 *The recollection characteristics of Generalized MCNNusing different control methods*
Shun Watanabe, Takashi Kuremoto, Shingo Mabu, Masanao Obayashi
(Yamaguchi University, Japan),
Kunikazu Kobayashi (Aichi Prefectural University, Japan)

GS4-2 *Revealing Terrorism Contents form Web Page Using Frequency Weighting Techniques*
Thabit Sabbah, Ali Selamat, Md Hafiz Selamat(Universiti Teknologi Malaysia, Malaysia)

GS4-3 *Application of Self Organizing Map for analyzing of robotic arm's action with Consciousness-Based Architecture module*
Wisanu Jitviriya, Eiji Hayashi (Kyushu Institute of Technology, Japan)

GS4-4 *Study on Time Synchronization Algorithm of Wireless Sensor Networks Based on Maximum Likelihood Estimation*
Ai Dongmei, Zeng Guangping, Tu Xuyan, He Di
(University of Science and Technology Beijing, China)

16:20-18:00 GS2 Reinforcement Learning System & Genetic programming
Chair: K. Kobayashi (Aichi Prefectural University, Japan)

GS2-1 *Development of phrase and music search engine by humming*
Kiminori Satou, Eiji Hayashi (Kyushu Institute of Technology, Japan)

GS2-2 *Human Recognition based on Gait feature and Genetic Programming*
Dipak Gaire Sharma, Ivan Tanev, Katsunori Shimohara (Doshisha University, Japan)

GS2-3 *An Action Selection Method Using Degree of Cooperation in a Multi-agent System*
Masanori Kawamura, Kunikazu Kobayashi (Aichi Prefectural University, Japan)

GS2-4 *Cooperative Action Acquisition Based on Intention Estimation Method in a Multi-agent Reinforcement Learning System*
Tatsuya Tsubakimoto, Kunikazu Kobayashi (Aichi Prefectural University, Japan)

GS2-5 *Integrating the Event Generation Mechanism in the Propp-based Story Generation Mechanism into the Integrated Narrative Generation System*
Shohei Imabuchi, and Takashi Ogata (Iwate Prefectural University, Japan)

January 12 (Sunday)

9:30- Registration

Room 301

10:00-17:00 PS1 Poster Session
Chair: D. Harzy (Universiti Malaysia Perlis (UniMAP), Malaysia)

PS1-1 *Relationship input object position and optimal error diffusion coefficients for Kinoform using error diffusion method*
Daisuke Kashima, Ken-ichi Tanaka (Meiji University, Japan)

PS1-*2* Optimization of dither matrix by hybrid of Genetic Algorithm and Simulated Annealing

Kohei Kato, Ken-ichi Tanaka (Meiji University, Japan)

PS1-3 *Inverse Halftoning using Multi-Layer Feed-Forward Neural Network(P)*
Hiroki Hamashoji, Ken-ichi Tanaka (Meiji University, Japan)

PS1-4 *Embedding of the Confidential Image in a Dithered Color Image*
Keisuke Banba and Kenichi Tanaka (Meiji University, Japan)

PS1-5 *Three Dimensional Images Reconstruction on Computer Generated Holograms the Multiple Regression Analysis*
Kenta Ayabe, Ken-ichi Tanaka(Meiji University, Japan)

PS1-6 *Embedding used binary number conversion into a Color image*
Hiroaki Oguma, Ken-ichi Tanaka (Meiji University, Japan)

Room 304

10:00-11:00 OS5 Analysis of Randomized Algorithms

Chair: Hiroshi Furutani (University of Miyazaki, Japan)

Co-Chair: Ichihi To (University of Miyazaki, Japan)

OS5-1 *Markov Chain Analyses of Random Local Search and Evolutionary Algorithm*
Hiroshi Furutani , Hiroki Tagami, Ichihi To, Makoto Sakamoto
(University of Miyazaki, Japan)

OS5-2 *Runtime Analysis of OneMax Problem in Genetic Algorithm*
Ichihi To, QuinLian Ma, Makoto Sakamoto, Hiroshi Furutani
(University of Miyaaki),
Yu-an Zhang (QinghaiUniversity, China)

OS5-3 *A Model for Low-Frequency Burst in Subthalamic Nucleus Neuron*
S. Kubota (Yamgata Univ., Japan)
J. E. Rubin (Pittsburgh Univ., USA)

13:00-14:40 OS1 Intelligent Control

Chair: Yingmin Jia (Beihang University, P.R.China)

Co-Chair: Chaoli Wang (University of Shanghai for Science and Technology, China)

OS1-1 *Finite-Time Stabilization for Nonholonomic Chained Form Systems with Communication Delay*
Hengjun Zhang, Chaoli Wang
(University of Shanghai for Science and Technology, China)

OS1-2 *Sliding Mode Variable Structure Control for Magnetic Levitation Vehicles*
Juanjuan He, Yingmin Jia(Beihang University, China)

OS1-3 *Variable-Poled Tracking Control of a Two-Wheeled Mobile Robot using Differential Flatness*
Liming Chen, Yingmin Jia (Beihang University, China)

OS1-4 *Adaptive Consensus Control of Multi-Agent Systems with Large Uncertainty and Time Delays*
Dongxu Zou, Weicun Zhang(University of Science and Technology Beijing, China)

OS1-5 *Iterative Learning Control for Overhead Crane Systems*
Xuhui Bu, Fuzhong Wang , Sanyi Li, Fashan Yu (Henan Polytechnic University, China)

15:00-17:00 OS2 Environment Navigation and Localization
Chair: J. M. Lee (Pusan National University, Korea)

OS2-1 *A Hybrid Path Planning Algorithm for UGV by Combining A* and B-spline Curve Equation*
Min-Ho Kim, Hee-Mu Lee, Min-Cheol Lee(Pusan National University, Korea)

OS2-2 *The actuator device design for the implementation of haptic joystick*
Dong-hyuk Lee, Sunkyun Kang, Bo-yeon hwang, Ki-jung Kim and Jangmyung Lee
(Pusan National University, Korea)

OS2-3 *Design of Fuzzy Controller using Variable Fuzzy Membership Function Factors for Inverse Ball Drive Mobile Robot*
Keon-woo Jeong, Shin-nyeong Heo, Seung-Ik Hwang, Han-Dong Yoo and Jangmyung Lee
(Pusan National University, Korea)

OS2-4 *Error Correction of Angular Velocity for Gyroscope using Genetic Algorithm and FIR filter*
Jaeyong Kim, Hyunhak Cho, Sungshin Kim (Pusan National University, Korea)

OS2-5 *Control of Robot Arm with Sterilization System for Ballast Water*
Seung-Hwa Baek, Dong-Hyun Kim, Hee-Je Kim (Pusan National University, Korea)

OS2-6 *Performance Study for Vehicle Infotainment System*
Heung-in Park, Min-gyu Kim, Jeong Hee-In, Kang-il Park, In-uk Lee,
Dong-Ju Lee and Jangmyung Lee
(Pusan National University)

Room305

10:00-11:20 OS8 Software Development Support Method
Chair: Tetsuro Katayama（University of Miyazaki,Japan）
Co-Chair: Makoto Sakamoto（University of Miyazaki,Japan）

OS8-1 *Proposal of a Visualizing Method of Data Transitions to Support Debugging for Java Programs*

Hiroto Nakamura, Tetsuro Katayama, Hisaaki Yamaba, Naonobu Okazaki
(University of Miyazaki, Japan)
Yoshihiro Kita (Kanagawa Institute of Technology, Japan)

OS8-2 *Proposal of a Method to Build Markov Chain Usage Model from UML Diagrams for Communication Delay Testing in Distributed Systems*
Zhijia Zhao, Tetsuro Katayama, Hisaaki Yamaba, Naonobu Okazaki
(University of Miyazaki, Japan)
Yoshihiro Kita (Kanagawa Institute of Technology, Japan)

OS8-3 *Proposal of a Supporting Method for Debugging to Reproduce Java Multi-threaded Programs by Petri-Net*
Shoichiro Kitano, Tetsuro Katayama, Hisaaki Yamaba, Naonobu Okazaki
(University of Miyazaki, Japan)
Yoshihiro Kita (Kanagawa Institute of Technology, Japan)

OS8-4 *Proposal of a Supporting Method to Generate a Decision Table from the Formal Specification*
Kenta Nishikawa, Tetsuro Katayama, Hisaaki Yamaba, Naonobu Okazaki
(University of Miyazaki, Japan)
Yoshihiro Kita (Kanagawa Institute of Technology, Japan)

13:00-14:40 OS3 Facial Expression Analysis, Music Recommendation and Augmented Reality

Chair: Yasunari Yoshitomi (Kyoto Prefectural University, Japan)

Co-Chairman: Masayoshi Tabuse (Kyoto Prefectural University, Japan)

OS3-1 *Development of a Campus Guide System based on Augmented Reality*
M. Tabuse, A. Tada (Kyoto Prefectural University, Japan)

OS3-2 *Facial Expression Analysis While Using Video Phone*
T. Asada, Y. Yoshitomi, R. Kato, M. Tabuse (Kyoto Prefectural University, Japan)
A. Tsuji, N. Kuwahara (Kyoto Institute of Technology, Japan),
J. Narumoto (Kyoto Prefectural University of Medicine)

OS3-3 *Method of Facial Expression Analysis Using Video Phone and Thermal Image*
Y. Yoshitomi, T. Asada, R. Kato, M. Tabuse (Kyoto Prefectural University, Japan)

OS3-4 *Facial Expression Recognition Using Thermal Image Processing and Efficient Preparation of Training-data*
Y. Nakanishi (ITOKI CORPORATION,Japan)
Y. Yoshitomi, T. Asada, M. Tabuse (Kyoto Prefectural University, Japan)

OS3-5 *Music Recommendation System through Internet for Improving Recognition Ability Using Collaborative Filtering and Impression Words*
Y. Yoshitomi, T. Asada, R. Kato, Y. Yoshimitsu, M. Tabuse
(Kyoto Prefectural University, Japan),
N. Kuwahara (Kyoto Institute of Technology, Japan)

J. Narumoto (Kyoto Prefectural University of Medicine)

15:00-16:40 OS4 Applications in Kansei Engineering
Chair: Tetsuo Hattori (Kagawa University, Japan)
Co-chair: Yoshiro Imai (Kagawa University, Japan)

OS4-1 *Kansei Engineering based Evaluation for an e-Learning System with IP-based Network Design and Animation*
Yoshiro Imai, Chiaki Kawanishi, Tetsuo Hattori (Kagawa University, Japan)

OS4-2 *Model Introduced SPRT for Structural Change Detection of Time Series (I)*
Yoshihide KOYAMA, Tetsuo HATTORI (Kagawa University, Japan)
Hiromichi KAWANO (NTT AT, Japan)

OS4-3 *Model Introduced SPRT for Structural Change Detection of Time Series II)*
--- Kansei Channel Corresponding to the SPRT ---
Yoshihide KOYAMA, Tetsuo HATTORI (Kagawa University, Japan)
Hiromichi KAWANO (NTT AT, Japan)

OS4-4 *Automated Color Image Arrangement Method and Kansei Impression*
Yusuke Kawakami, Tetsuo Hattori, Yoshiro Imai, Haruna Matsushita
(Kagawa University, Japan)
Hiromichi Kawano (NTT AT, Japan),
R.P.C. Janaka Rajapakse (Tainan National University of the Arts,Taiwan)

OS4-5 *Novel Transistor by Opt-coupling of LED and PD and its Application to Audio Amplifier*
Junichi Fujita, Tetsuo Hattori, Daisuke Sato, Kensho Okamoto
(Kagawa University, Japan)

January13 (Monday)

9:30- Registration

Room 304

10:00-11:40 GS6 Robotics Ⅰ
Chair: Fengzhi Dai (Tianjin University of Science and Technology, China)

GS6-1 *Study on the pose under complex multiple targets environment for the industrial robot based on machine vision*
Jiwu Wang, Xianwen Zhang, Huazhe Dou (Beijing Jiaotong University, China)
Sugisaka Masanori (Nippon Bunri University, Oita)

GS6-2 *Research and Development of an Intelligent Robot of Medical Assistance Based on Embedded RTOS*
Yuan Li, Fengzhi Dai, Junhong Xi, Binyao Li, Guodong You
(Tianjin University of Science and Technology, China)

GS6-3 *High Acceleration Robotic Arm for Dynamic and Dexterous Manipulation of Deformable Object*
Hiroaki Seki, Hiroki Shibata, Yoshitsugu Kamiya, Masatoshi Hikizu
(Kanazawa University, Japan)
Khairul Salleh Mohamed Sahari (Universiti Tenaga Nasional, Malaysia)

GS6-4 *Switching based controller algorithm design for uncertainties rejection in smooth takeoff / landing and of quad-rotor*
M. Hassan Tanveer, D. Hazry, S. Faiz Ahmed, M. Kamran Joyo, Faizan. A. Warsi, A.T. Hussain
(Universiti Malaysia Perlis) (UniMAP)

GS6-5 *Parameters of Gas-Liquid Flow Distribution Uniformity in Upward Multi-Pass Channels*
Zuradzman M. Razlan, R. Heng, D. Hazry, A.B. Shahriman, Khairunizam WAN,
S. Faiz Ahmed, Nazrul H. ADNAN (Universiti Malaysia Perlis, Malaysia)
M. Hirota, N. Maruyama, A. Nishimura (Mie University, Japan)
H. Hisyam (Lean Applied Pte. Ltd.,Malaysia)

13:00-14:40 OS6 Empirical research on Network and MOT

Chair: T. Ito (Ube National College of Technology)

Co-Chair: S. Matsuno (Ube National College of Technology)

OS6-1 *An Empirical Examination of Inter-firm Capital Relationships in Mazda's Yokokai using the IDE Spatial Model*
M. Sakamoto (University of Miyazaki, Japan),
S. Tagawa, T. Ito, S. Matsuno (Ube National College of Technology, Japan)
R. Mehta (New Jersey Institute of Technology, U.S.A)
V. Berdonosov (Komsomolsk-on-Amur State University of Technology, Russia)
S. Ikeda (University of Miyazaki, Japan)

OS6-2 *A Comparative Study of Transactional Network between Kyohokai and Yokokai*
M. Sakamoto (University of Miyazaki, Japan)
M. Hasama, T. Ito, Y. Uchida (Ube National College of Technology, Japan)
R. Mehta (New Jersey Institute of Technology, U.S.A)
Y. Park (Prefectural University of Hiroshima, Japan)
S. Ikeda (University of Miyazaki, Japan)

OS6-3 *A path analytic model and measurement of the relationships between green supply chain management implementation and corporate performance*
S. Matsuno, M. Hasama, Y. Uchida, T. Ito (Ube National College of Technology, Japan)

OS6-4 *A Study of Open Source Cloud System for Small and Medium Enterprise*
Y. Uchida, S. Matsuno, T. Ito, M. Hasama (Ube National College of Technology, Japan)
M. Sakamoto (University of Miyazaki, Japan)

OS6-5 *Visualization of Patent Strategies in Japanese ICT Companies Based on Text-Mining*
T. Tokumitsu, T. Okada, I. Nakaoka (Ube National College of Technology, Japan)
Y. Park (Prefectural University of Hiroshima, Japan)

Room305

10:00-12:00 OS7 Intelligence in Biological Systems

Chair: Jun Nishii (Yamaguchi University, Japan)

Co-Chair: Hiroshi Matsuno (Yamaguchi University, Japan)

OS7-1 *Dissipative Particle Dynamics of shape changes of vesicle*
Yoshiyuki Oofuji, Naohito Urakami, Takashi Yamamoto (Yamaguchi University, Japan)
Masayuki Imai (Tohoku University, Japan)

OS7-2 *Intelligent mechanisms in E. coli in processing carbon sources*
Zhongyuan Tian, Hiroshi Matsuno (Yamaguchi University, Japan)

OS7-3 *An autonomous propagation of ciliary metachronal wave on elastic surface of Paramecium cells.*
N. Narematsu, Y. Iwadate (Yamaguchi University)

OS7-4 *An analysis of the synergy in a ball throwing task*
Hiroshi KIMURA, Jun NISHII (Yamaguchi University, Japan)

OS7-5 *A low dimensional feedback control model that exploits abundant degrees of freedom*
Jun Nishii, Tohru Hamamura (Yamaguchi University, JAPAN)

OS7-6 *Extension of Genetic Toggle Switch Based on the Effective Search of State Transitions*
M. Sugii, A. Fauré, H. Matsuno (Yamaguchi University, Japan)

13:00-14:20 OS9 Foundation of computation and its application

Chair: Makoto Sakamoto (University of Miyazaki, Japan)

Co-Chair: Tetsuro Katayama (University of Miyazaki, Japan)

OS9-1 *Hierarchy Based on Neighborhood Template about k-Neighborhood Template A-Type Three-Dimensional Bounded Cellular Acceptor*
Makoto Sakamoto, Makoto Nagatomo, Xiaoyang Feng, Tatsuma Kurogi, Tuo Zhang, Satoshi Ikeda, Masahiro Yokomichi, Hiroshi Furutani (University of Miyazaki, Japan)
Takao Ito, Yasuo Uchida (Ube National College of Technology, Japan)
Tsunehiro Yoshinaga(Tokuyama College of Technology, Japan)

OS9-2 *Hierarchy Based on Configuration-Reader about k-Neighborhood Template A-Type Three-Dimensional Bounded Cellular Acceptor*
Makoto Sakamoto, Tuo Zhang, Tatsuma Kurogi, Makoto Nagatomo, Xiaoyang Feng, Satoshi Ikeda, Masahiro Yokomichi, Hiroshi Furutani (University of Miyazaki, Japan)
Yasuo Uchida, Takao Ito (Ube National College of Technology, Japan)
Tsunehiro Yoshinaga (Tokuyama College of Technology, Japan)

OS9-3 *Handicap of Othello Game*
Y. Kato, S. Ikeda, M. Sakamoto (University of Miyazaki, Japan)
T. Ito (Ube National College of Technology, Japan)

OS9-4 *On Parameter Setting in Identifying the Same Languages Involved in Different Language Data*
Ren Wu (Yamaguchi Junior College, Japan)
Hiroshi Matsuno (Yamaguchi University, Japan)

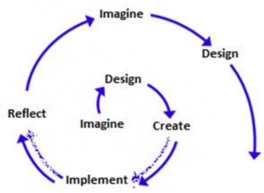
ABSTRACTS

IS

IS1-1 Playware Research – Methodological Considerations

Henrik Hautop Lund (Centre for Playware, Technical University of Denmark, Denmark)

Several sub-disciplines of engineering are driven by the researchers' aim of providing positive change to the society through their engineering. Based on two decades research in developing engineering systems with a societal impact (e.g. in robotics, embodied AI, and playware), in this paper we suggest a cyclic research method based on a mix between participatory and experimental processes. In particular, inspiration from the action research method applied to interdisciplinary technology development becomes a participatory approach characterized by rapid prototyping cycles which allow iterative technology specification and development together with people in their real world environment. With the mixed research method, we suggest that there are cases, where approaches from the positivistic and interpretivistic epistemologies can and should be merged. We exemplify this with health care technology and playware.



IS1-2 ALife for Real and Virtual Audio-Video Performances

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

MAG is an electronic art piece in which an ALife software attempts to "translate" musical expression into a corresponding static or animated graphical expressions. The mechanism at the base of such "translation" is based on artificial learning techniques. MAG use population of neural networks that can use both genetic and reinforcement learning algorithms to evolve appropriate behavioral answers to inputs. The combination of artificial evolution and the flows of a repeated song or different musical tunes allows the software to obtain a special relationship between sound waves and the aesthetics of consequent graphical result. Further, we are exploring the concept of run-time creation of both music and graphical expression. To do that we use MusicTiles app that allows users to remix music by connecting musical building blocks. When combining MusicTiles app and MAG software, we provide the possibility to melt both musical expression and graphical expression in parallel and at run-time, and therefore creating a audio-video performance that is always unique.

IS1-3 Heart-pulse Biofeedback in Playful Exercise using a Wearable device and Modular Interactive Tiles

Tomoya Shimokakimoto[1], Henrik H. Lund[2], Kenji Suzuki[3]
([1] University of Tsukuba, Japan)
([2] Technical University of Denmark, Denmark)
([3] University of Tsukuba/ JST, Japan)

We developed a playful biofeedback system using a wearable device and modular interactive tiles. We suppose that patients could regulate exercise intensity on their own through biofeedback. We propose biofeedback play system called "bioToys" based on exercise with the modular interactive tiles. The system consists of a wearable device and modular interactive tiles. We combine the two systems to provide users with heart pulse biofeedback in playful exercise. We show that using the developed system it is possible for the users to regulate the exercise intensity on their own with biofeedback, and also possible to analyze exercise activity using number of steps on the tiles and heart beat rate.

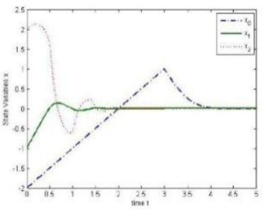
OS

OS1 Intelligent Control

OS1-1 Finite-Time Stabilization for Nonholonomic Chained Form Systems with Communication Delay

Hengjun Zhang, Chaoli Wang
(University of Shanghai for Science and Technology, P.R.China)

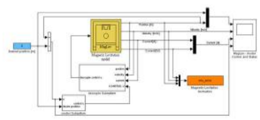
In this paper, the problem of finite-time stabilization is developed for nonholonomic chained form systems with communication delay in the input. The finite-time control laws are presented by utilizing the switching control strategy and the theory of finite-time stability, which can make the states of the nonholonomic chained form systems to converge from any non-equilibrium state to the equilibrium or a given point in a finite time. Finally, the simulation results show the effectiveness of the proposed control approach.



OS1-2 Sliding Mode Variable Structure Control for Magnetic Levitation Vehicles

Juanjuan He, Yingmin Jia
(Beihang University, P.R.China)

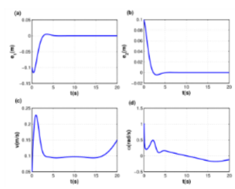
This paper focuses on stability control for the levitated positioning of the magnetic levitation vehicle system. For the nonlinear magnetic levitation system model, the output feedback linearization method is first employed to derive a global linearization error model. However, there exists uncertain item in the error model. To stabilize this error model, the adaptive sliding mode variable structure control method is used to design stability controller. Simulations show that the magnetic levitation system can be stable and track the desired signals quickly under the proposed control scheme.



OS1-3 Variable-Poled Tracking Control of a Two-Wheeled Mobile Robot using Differential Flatness

Liming Chen, Yingmin Jia
(Beihang University, P.R.China)

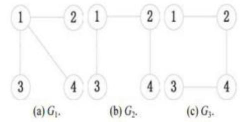
This paper investigates the tracking controller design of a two-wheeled mobile robot in its kinematic model and dynamic model. Differential flatness and linear time-varying(LTV) systems' PD-spectral theory are used. Based on differential flatness, original system is transformed via a state prolongation and state transformation into a normal form to apply feedback linearization. Then using PD-spectral theory, variable poles of tracking error dynamics are assigned to realize trajectory tracking stability. Finally, simulation results are presented to demonstrate the feasibility and effectiveness of the proposed method.



OS1-4 Adaptive Consensus Control of Multi-Agent Systems with Large Uncertainty and Time Delays

Dongxu Zou, Weicun Zhang
(University of Science and Technology Beijing, China)

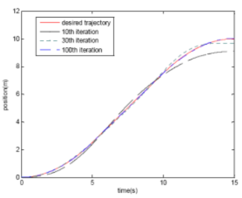
A weighted multi-model adaptive control method is proposed to achieve consensus of multi-agent system with large parameter perturbation and communication delays, in which H^{inf} control is adopted to construct the controller set. Moreover a simple and effective weighting algorithm is also presented. The simulation results demonstrate the effectiveness of the proposed method.



OS1-5 Iterative Learning Control for Overhead Crane Systems

Xuhui Bu, Fuzhong Wang , Sanyi Li, Fashan Yu
(Henan Polytechnic University, P.R.China)

In many factories, overhead crane often transfers the same loads from one place to another following a predefined position and speed diagrams. This unique feature offers overhead crane an opportunity of improving its performance through learning iteratively. In this paper, we apply the iterative learning control approach to overhead crane systems. Based on the linearization dynamic model of overhead crane, an ILC scheme contains both feedforward learning control part and state feedback control is proposed. By providing a 2D formulation, the ILC design for crane systems can be transformed into the problem of state feedback control for 2-D systems described by Roesser models. It is shown that the proposed approach can guarantee the trolley position asymptotically converges to its desired profile with small swing angle. Simulations are illustrated to show the feasibility and effectiveness of the proposed approach.



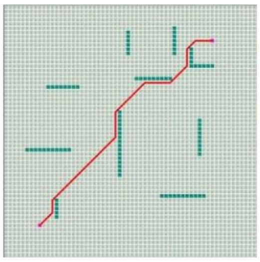
OS2 Environment Navigation and Localization

OS2-1 A Hybrid Path Planning Algorithm for UGV by Combining A* and B-spline Curve Equation

Min-Ho Kim, Hee-Mu Lee and Min-Cheol Lee

(Pusan National University, Busan, South Korea)

This article presents a hybrid path planning algorithm to make a smooth path for an UGV. A main theme of the proposed algorithm is to combine of the A* algorithm and the B-spline curve equation. A* is one of the well-known path planning algorithm which finds the optimal path on the given map by using the heuristic cost function. However, since A* is based on the grid map, the result path consists of straight lines with only 8 directions. It's not suitable of UGV's navigation. Therefore in this paper to overcome this issue, B-spline curve equation is proposed to make a smooth and continuous path with control points which are selected from the A* result path. And the optimal control point selection algorithm is proposed to make the hybrid path. At last, to verify the proposed algorithm, the hybrid path results are compared with A* algorithm by using a developed simulation program.

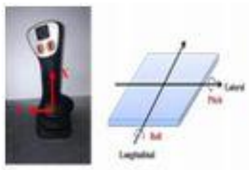


OS2-2 The actuator device design for the implementation of haptic joystick

Dong-hyuk Lee, Sun-kyun Kang, Bo-Yeon Hwang, Ki-jung Kim, Jang-myung Lee

(Pusan National University, Korea)

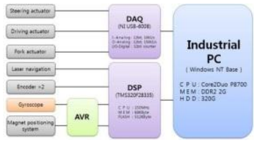
In this paper, we designed the actuator device for haptic implementation in the general joystick. Each linear servo actuator as pillars to support the joystick mounting plate can move up and down. And this movement will generate the inclination of the plate. The haptic can be implemented by the change of the inclination. The actuator device was designed through the performance analysis of actuator device structure. The haptic will represent information from obstacle cognition device by using the designed actuator device. The haptic performance of the actuator device was verified through experiments.



OS2-3 Design of Fuzzy Controller using Variable Fuzzy Membership Function Factors for Inverse Ball Drive Mobile Robot

Keon-woo Jeong, Shin-nyeong Heo, Seung-Ik Hwang, Han-Dong Yoo, Jang-myung Lee
(Pusan National University, Korea)

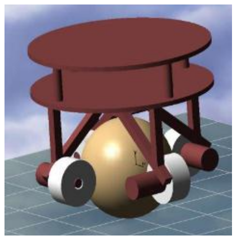
In this paper, a fuzzy controller for a inverse ball drive mobile robot is implemented to have more stable balancing capability than the conventional control system. Fuzzy control structure is chosen for a inverse ball drive mobile robot, and fuzzy membership function factors for the control system are obtained for 3 specified weights using a trial-and-error method. Next a linear Interpolation method is employed to generate fuzzy membership function factors for more stable control performance when the weight is arbitrarily selected. Through some experiments, we find that the proposed fuzzy controller using the neural network is superior to the conventional fuzzy controller.



OS2-4 Error Correction of Angular Velocity for Gyroscope using Genetic Algorithm and FIR filter

Jaeyong Kim, Hyunhak Cho, Sungshin Kim
Pusan National University, Korea

This paper is research which sensitivity of gyro sensor is optimization. Recently, the MEMS-gyroscope is usually used at mobile robots and AGV by development of MEMS technology. However, the MEMS-gyroscope has measurement error, cumulate error, bias drift, etc. To reduce these errors, measurement of angular velocity using MEMES-gyroscope required high precision and exactly sensitivity. The sensitivity, is a constant to change angular velocity form ADC value, is changed by external factors as installed location and gradient because the MEMS-gyroscope measure angular velocity with inertia. Therefore, this paper proposes to reduce variance of ADC value and to optimize the sensitivity of MEMS-gyroscope using genetic algorithm. To verify performance of proposed method, we attached MEMS-gyroscope to automatic guided vehicle (AGV) of fork type. In experimental result, we verified that the sensitivity using the proposed method is accurate than the sensitivity in specification.



OS2-5 Control of Robot Arm with Sterilization System for Ballast Water

Seung-Hwa Baek, Dong-Hyun Kim and Hee-Je Kim

(Pusan National University,Korea)

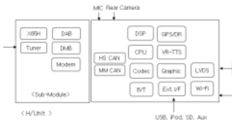
The inadvertent transfer of harmful aquatic organisms and pathogens in the ballast water of ships has been determined to have caused a significant adverse impact to many of the world's coastal regions. Loading and discharge of ballast water is an essential part of a ship's operation, and is fundamental to maintaining safe operations under different onditions of load. However, large vessels require thousands of tonnes of water to ensure stability and manoeuvr ability, and the environmental impacts of this can be considerable. These impacts result from the fact that the ballast water can contain hundreds of different species, many of which can have serious ecological, economic and public health effects if transferred to regions where they are not native. The recognition of these effects has made ballast management increasingly important for protection of the marine environment. As a means to prevent, we developed a system using high voltage- high frequency pulsed power, so called arc discharge, when this discharge happens, the electrodes are oxidized and consumed. Regarding this point, we wanted to apply robot arm to supply continuous electrode when it consumed by arc discharge.

OS2-6 Performance Study for Vehicle Infotainment System

Heung-in Park, Min-gyu Kim, Jeong Hee-In, Kang-il Park, In-uk Lee, Dong-Ju Lee, Jang-myung Lee

(Pusan National University, Korea)

CPU performance, booting time and navigation widget are improved by the new design concepts. These points are providing some solutions that can overcome the risks of open architecture system in the Automotive. The improvement was verified with each of the test results.



OS3 Facial Expression Analysis, Music Recommendation and Augmented Reality

OS3-1 Development of a Campus Guide System based on Augmented Reality

M. Tabuse, A. Tada

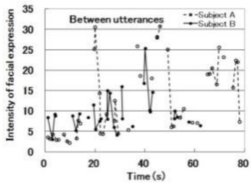
(Kyoto Prefectural University, Japan)

This paper presents a campus guide system based on augmented reality (AR). AR is a technology of displaying CG overlaid on a real scene. AR is a fusion of a real world and a virtual world of human interfaces. We propose a campus guide system based on markerless AR. This system displays information of a university overlaid on a real scene using a note PC and a web camera. In addition, this system presents a variety of events, for example a university festival and a welcome party, as 3D virtual objects overlaid on a real scene. Using this system, visitors can understand a university life more deeply. We have developed a campus guide system based on Parallel Tracking and Multiple Map system and construct 3D virtual objects using Visual Structure from Motion System. And we have confirmed the effectiveness of the system in our campus.

OS3-2 Facial Expression Analysis While Using Video Phone

T. Asada[1], Y. Yoshitomi[1], A. Tsuji[2] , R. Kato[1], M. Tabuse[1], N. Kuwahara[2], J. Narumoto[3]
([1] Kyoto Prefectural University, Japan)
([2] Kyoto Institute of Technology, Japan)
([3] Kyoto Prefectural University of Medicine)

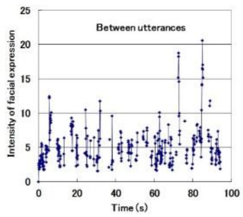
We have proposed a method for analyzing facial expressions of a person while speaking with use of a video phone system (Skype). The video is analyzed using image processing software library (OpenCV) and the newly proposed feature vector of facial expression, which is extracted in the mouth-part area with use of 2D-DCT performed for each domain having 8×8 pixels. The facial expression intensity defined as the norm of difference vector between the feature vector of neutral facial expression and that of observed one can be used for analyzing a chance of facial expression. Combining the video signal obtained from the Skype with the sound signal obtained with the other way, we can distinguish the facial expression with speaking from that without speaking. The judgment of speaking is performed with a threshold of the sound intensity. Some experimental results show the usefulness of the proposed method.



OS3-3 Method of Facial Expression Analysis Using Video Phone and Thermal Image

Y. Yoshitomi, T. Asada, R. Kato, and M. Tabuse
(Kyoto Prefectural University, Japan)

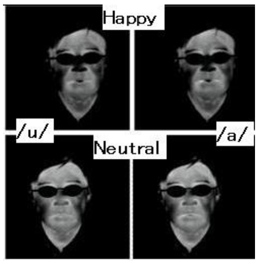
In this study, using thermal image processing, we have proposed a method for analyzing facial expressions of a person while speaking with a video phone system (Skype). The newly proposed feature vector of facial expression, which is extracted in the mouth-part area with use of 2D-DCT performed for each domain having 8×8 pixels. The facial expression intensity defined as the norm of difference vector between the feature vector of neutral facial expression and that of observed one can be used for analyzing a chance of facial expression. Combining the video signal obtained from the Skype with the sound signal obtained with the other way, we can distinguish the facial expression with speaking from that without speaking. The judgment of speaking is performed with a threshold of the sound intensity. Some experimental results show the usefulness of the proposed method.



OS3-4 Facial Expression Recognition Using Thermal Image Processing and Efficient Preparation of Training-data

Y. Nakanishi[1], Y. Yoshitomi[2], T. Asada[2], M. Tabuse[2]
([1] ITOKI CORPORATION, Japan)
([2] Kyoto Prefectural University, Japan)

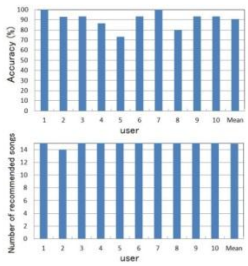
We investigated the influence of training data on the facial expression accuracy using the training data of “taro,” whose first and last vowels are /a/ and /o/, for the three intentional facial expressions of “angry,” “sad,” and “surprised,” and the training data of 25 pairs of vowels for the two intentional facial expressions of “happy” and “neutral.” Using the proposed method, the facial expressions of one subject were discriminable with accuracies of 100%, 70.0%, and 47.2% for “taro,” “koji” (the first and last vowels of which are /o/ and /i/), and “tsubasa” (the first and last vowels of which are /u/ and /a/), respectively, for the three facial expressions of “happy,” “neutral,” and “other” when one of the five intentional facial expressions of “angry,” “happy,” “neutral,” “sad,” and “surprised” was exhibited.



OS3-5 Music Recommendation System through Internet for Improving Recognition Ability Using Collaborative Filtering and Impression Words

Y. Yoshitomi[1], T. Asada[1], R. Kato[1],Y. Yoshimitsu[1], M. Tabuse[1], N. Kuwahara[2], J. Narumoto[3]
([1] Kyoto Prefectural University, Japan)
([2]Kyoto Institute of Technology, Japan)
([3]Kyoto Prefectural University of Medicine)

In our previously reported system, the recommendation process using collaborative filtering was terminated when no users in the reference list have the same preference for recommended music as that of a new user. Based on the scores for impression words, the second recommendation process finds the most similar music to the successfully recommended music among music that has not yet been recommended. In the present study, based on our previously reported system, we propose a system for music recommendation through the Internet using a videophone system (Skype). The effectiveness of the proposed system is verified. The recommendation accuracy of the proposed system was 90.6% for 10 trials by five elderly subjects. The average number of recommended pieces of music for the subjects by the proposed system was 14.9 per trial.

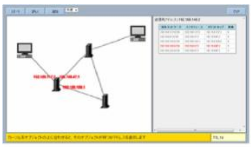


OS4 Applications in Kansei Engineering

OS4-1 Kansei Engineering based Evaluation for an e-Learning System with IP-based Network Design and Animation

Yoshiro Imai, Chiaki Kawanishi, Tetsuo Hattori
(Kagawa University, Japan)

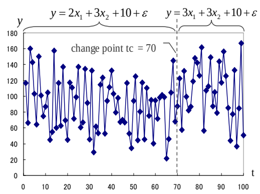
An e-Learning system has been developed, which is designed for network education with facilities of specification of IP-based network topology and demonstration of packet-transferring animation. This system can execute on the major browsers by means of accessing a specific Web server. It is frequently used in some lectures of university from beginners of network to the students of information related course. Kansei Engineering approach is utilized and applied in order to improve and evaluate the e-Learning system.



OS4-2 Model Introduced SPRT for Structural Change Detection of Time Series(I)

Yoshihide KOYAMA[1], Tetsuo HATTORI[1], Hiromichi KAWANO[2]
([1]Kagawa University, Japan)
([2]NTT AT, Japan)

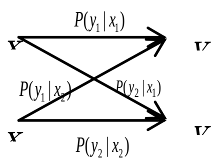
Previously, we have proposed a method applying Sequential Probability Ratio Test (SPRT) to structural change detection problem of ongoing time series data. In this paper, we introduce a structural change model with Poisson process into the system that outputs a set of time series data, moment by moment. Then we can calculate the value of P(Z | H1) that denotes the probability of observing the data Z under the condition that H1(alternative hypothesis) is true. Also, we concretely show the theory of change detection of time series by the model introduced SPRT.



OS4-3 Model Introduced SPRT for Structural Change Detection of Time SeriesII) --- Kansei Channel Corresponding to the SPRT ---

Yoshihide KOYAMA[1], Tetsuo HATTORI[1], Hiromichi KAWANO[2]
([1]Kagawa University, Japan)
([2]NTT AT, Japan)

We present an application of the notions in Information Theory to Kansei Engineering toward a mathematical methodology for the analysis in the Kansei Engineering field. In this paper, first, we propose a definition of correlation using the mutual information in Information Theory. Second, we explain a relation between the Bayes' Updating using the notion of binary channel and Sequential Probability Ratio Test (SPRT). Moreover, we show the Channel Matrix corresponding to the SPRT where a model introduced.



OS4-4 Automated Color Image Arrangement Method and Kansei Impression

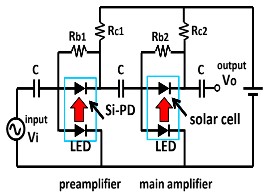
Yusuke Kawakami[1], Tetsuo Hattori[1], Yoshiro Imai[1], Haruna Matsushita[1],
Hiromichi Kawano[2], R.P.C. Janaka Rajapakse[3]
([1]Kagawa University, Japan),
([2]NTT AT, Japan),
([3]Tainan National University of the Arts, Taiwan)

This paper proposes a new color image arrangement method using an elastic transform on some kinds of axes. In this paper, we present the principle of our method using HMGD (Histogram Matching based on Gaussian Distribution). And we describe that the automated method applies the HMGD to input color image only when the image has single-peakedness in its histogram on the focused axis. We also show that the method gives a good Kansei effect in the case of applying the HMGD onto Lightness axis.

OS4-5 Novel Transistor by Opt-coupling of LED and PD and its Application to Audio Amplifier

Junichi Fujita, Tetsuo Hattori, Daisuke Sato, Kensho Okamoto
(Kagawa University, Japan)

This paper presents a simple transistor-like amplification function by the optical coupling of LED (Light Emitting Diode) and Photodiode where the Photodiode current by photo electromotive force is feed backed into the LED, in the same way as the case of common emitter circuit of the conventional Bipolar Transistor. And also, this paper shows that an audio amplifier can be realized in a simple form. Since the proposed audio amplifier provides good sound with very low noise characteristic, we consider that the audio system becomes a promising Kansei product.

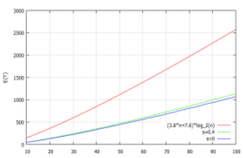


OS5 Analysis of Randomized Algorithms

OS5-1 Markov Chain Analyses of Random Local Search and Evolutionary Algorithm

Hiroshi Furutani, Hiroki Tagami, Ichihi To, Makoto Sakamoto
(University of Miyazaki, Japan)

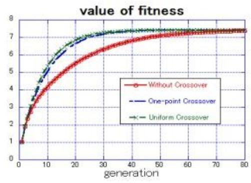
In this paper, we report the computational complexity of algorithms (1+1)EA and Random Local Search (RLS). It has been noted the close resemblance of RLS with the coupon collector's problem (CCP). The CCP has a long history of probabilistic research, and many interesting results are obtained. This study makes use of such results with some modifications. We also show the results for (1+1)EA and RLS using absorbing Markov chain.



OS5-2 Runtime Analysis of OneMax Problem in Genetic Algorithm

Ichihi To[1], QuinLian Ma[1], Makoto Sakamoto[3], Hiroshi Furutani[1],Yu-an Zhang[2]
([1]University of Miyazaki, Japan)
([2] QinghaiUniversity, China)

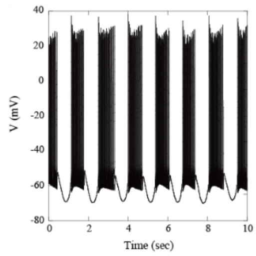
Genetic algorithms (GAs) are stochastic optimization techniques that simulate the biological evolution. Theoretical study of the evolution of the genetic algorithm is very important for the application of GA. We have studied the effects of stochastic fluctuation in the process of GA evolution. We consider the task of estimating the hitting time of the optimal solution in GA for OneMax problem. This task is performed by using the Markov chain representing the behavior of population in the stationary state of evolution.



OS5-3 A Model for Low-Frequency Burst in Subthalamic Nucleus Neuron

S. Kubota[1], J. E. Rubin[2]
([1]Yamgata Univ., Japan)
([2]Pittsburgh Univ., USA)

Bursting oscillation in the subthalamic nucleus (STN) is suggested to be linked to the motor symptoms in Parkinson’s disease. In this study, to investigate the effects of interactions among STN, the cortex, and globus pallidus (GP), on the burst characteristics, we examine a model of STN neuron that contains NMDA conductances activated through the cortical inputs and GABAergic conductances associated with GP activities. To reproduce cortical slow oscillation consisting of two phases, the NMDA conductances are activated by inhomogeneous Poisson spikes whose firing rate is stochastically and periodically modulated. The results show that a higher level of GABA inhibition leads to strong burst with larger coefficient of variation of interspike intervals, while the strengthening of NMDA activation mainly acts to increase the firing rate.



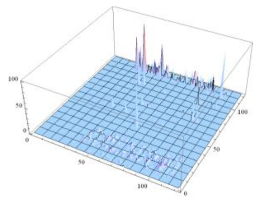
Burst response of the STN neuron model

OS6 Empirical research on Network and MOT

OS6-1 An Empirical Examination of Inter-firm Capital Relationships in Mazda’s Yokokai using the IDE Spatial Model

M. Sakamoto[1], S. Tagawa[2], T. Ito[2], S. Matsuno[2], R. Mehta[3], V. Berdonosov[4], S. Ikeda[1]
([1] University of Miyazaki, Japan)
([2] Ube National College of Technology, Japan)
([3] New Jersey Institute of Technology, U.S.A)
([4] Komsomolsk-on-Amur State University of Technology, Russia

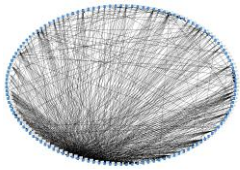
In this paper, we focused on examining the capital relationships in Yokokai, the Mazda’s Keiretsu. Employing regression analysis, three significant indices, influence, degree, and efficiency, were selected from nine indices, including others that include dyadic redundancy, dyadic constraint, effective size, constraint, hierarchy, and density. We propose a new approach, called IDE spatial model, to calculate the strength of the inter-firm’s relationships. In order to ascertain the rational inter-firm relationships, network indices and corporate performance are analyzed. This research suggests a new perspective to examine the rational inter-firm relationship that can be used in any network organization.



OS6-2 A Comparative Study of Transactional Network between Kyohokai and Yokokai

M. Sakamoto[1], M. Hasama[2], T. Ito[2], Y. Uchida[2], R. Mehta[3], Y. Park[4], S. Ikeda[1]
([1] University of Miyazaki, Japan)
([2] Ube National College of Technology, Japan)
([3] New Jersey Institute of Technology, U.S.A)
([4] Prefectural University of Hiroshima, Japan)

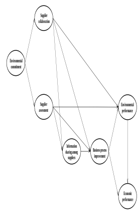
Consistent with the stream of research on the coordination of internal resources and external inter-firm relationships, the purpose of this research is to identify and contrast the best practices of Toyota to Mazda using a comparative approach. This paper reviews relevant studies of network organizations to focus on the differences between Kyouhokai and Yokokai. Specifically, it makes a contribution to the literature by proposing a new perspective to identify the determinants of corporate performance and clarify the difference among the external inter-firm's relationships between Kyohokai and Yokokai, thus ascertaining the rational structure of network organizations.



OS6-3 A path analytic model and measurement of the relationships between green supply chain management implementation and corporate performance

S. Matsuno, M. Hasama, Y. Uchida, and T. Ito
(Ube National College of Technology, Japan)

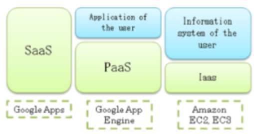
We propose and empirically test a model of the relationships between green supply chain management (GSCM) activities and corporate performance. Five constructs, namely, environmental commitment, supplier collaboration, supplier assessment, information sharing among suppliers, and business process improvement are used to form a structural model explaining the environmental and economic performance. The model was analyzed using the data from a survey of sample of manufacturing firms in Japan. The results suggest that the degree of supplier collaboration has an influence on the environmental performance directly. While, the impact of supplier assessment on the environmental performance is mediated by the information sharing and/or business process improvement.



OS6-4 A Study of Open Source Cloud System for Small and Medium Enterprise

Y. Uchida[1], S. Matsuno[1], T. Ito[1], M. Hasama[1], M. Sakamoto[2]
([1]Ube National College of Technology, Japan)
([2]University of Miyazaki, Japan)

The use of cloud services among SMEs remains low, and R&D and technical support for SMEs is an urgent problem. Even so, there have been almost no academic studies on the relationship between cloud computing and SME information infrastructure, or on the future direction of this. With this background, we considered and designed a cloud system framework suitable for SMEs, built a prototype model and investigated the ease of implementation. Services that deliver functionality equivalent to Amazon EC2 which is representative cloud service but delivered through other technologies are called EC2 clones. OpenStack is not EC2 clone, and is a typical open source project for providing Infrastructure as a Service (IaaS) type cloud services. OpenStack was used in this study to implement a simple cloud service model.



OS6-5 Visualization of Patent Strategies in Japanese ICT Companies Based on Text-Mining

T. Tokumitsu[1], T. Okada[1], I. Nakaoka[1], Y. Park[2]
([1] Ube National College of Technology, Japan)
([2] Prefectural University of Hiroshima, Japan)

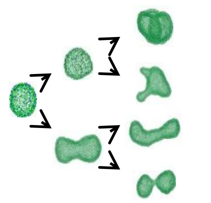
We deduce that a major reason of the continual poor business performances in most Japanese ICT companies is that they have lost their technology innovation strategies. Especially in smartphone market, Japanese ICT companies have lagged behind foreign companies such as Samsung and Apple. Therefore, we conduct longitudinal research of three Japanese large ICT companies, SONY, Panasonic and Sharp, in smartphone market based on text-mining approach covering their patent information. A visualized result of patent strategies by SONY is shown in the figure. We analyze the other two companies, Panasonic and Sharp, in the same way, and reveal characteristics of their patent strategies.

OS7 Intelligence in Biological Systems

OS7-1 Dissipative Particle Dynamics of shape changes of vesicle

Yoshiyuki Oofuji[1], Naohito Urakami[1], Takashi Yamamoto[1], Masayuki Imai[2]
([1]Yamaguchi University, Japan)
([2]Tohoku University, Japan)

Spherical vesicles change to various shapes such as oblate type, prolate type, stomatocyte, etc., according to the osmotic pressure difference between the inner and the outer vesicle. The shape changes of vesicles are very important to understand the living cell activities. In the research, we investigated the process of the shape changes of vesicles by carrying out dissipative particle dynamics simulations. The shape changes of vesicles were reproduced by the variation of two parameters, the number of water molecules inside the vesicle and the difference between the number of lipid forming the inner and the outer leaflets in the vesicle. The vesicle shapes obtained in our simulations were in good agreement with the experiments. The simulation result indicates that the vesicle shapes are determined by two parameters.

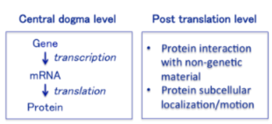


Shape changes of vesicles

OS7-2 Intelligent mechanisms in *E. coli* in processing carbon sources

Zhongyuan Tian, Hiroshi Matsuno
(Yamaguchi University, Japan)

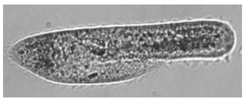
E. coli is "wise" enough to take suitable responding time, and suitable responding behaviors, when facing different lengths or intensities of stimuli. According to the time cost of a respond to a signal, we divide the intracellular processes into 2 levels: the *central dogma level* and *the post translation level*. Firstly, we constructed a systematical network of glucose, PTS, glycogen and chemotaxis system. And this network plays a function as a switch, which controls *E coli*'s motion within the first few seconds. By using this switch, we unveiled a fact that shorter time stimuli result in *the post translation level* reactions for quicker response. But longer time stimuli will activate more time consuming *the central dogma level* reactions. Secondly, different intensities of signals result in different kinds of actions, was illustrated by ppGpp example.



OS7-3 An autonomous propagation of ciliary metachronal wave on elastic surface of *Paramecium* cells.

N. Narematsu and Y. Iwadate
(Yamaguchi University)

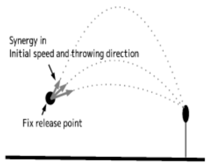
Ciliary movements in protozoa show metachronal coordination so as to maintain a constant phase difference between adjacent cilia. This coordination is called as “metachronal wave”. It is now generally thought that metachronal waves arise from hydrodynamic coupling between adjacent cilia at extracellular fluid. However, under the breakdown the hydrodynamic coupling of ciliary movements at a restricted portion of a *Paramecium* cell, metachronal waves pass over the portion. We will discuss the mediator of propagation of metachronal waves.



OS7-4 An analysis of the synergy in a ball throwing task

Hiroshi KIMURA, Jun NISHII
(Yamaguchi University, Japan)

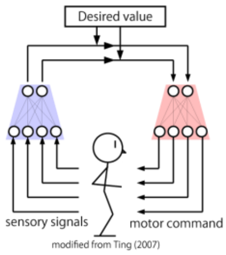
To throw a ball into a given target, we must adequately select the initial state of the ball, the position, speed, and throwing direction at ball release. There are two ways in the choice of these variables. The first solution is to learn a desired value set for these state variables and try to adjust the variables to the desired values. Another solution is to allow the variance of each variable in every trial but compensate for the values each other in a cooperative manner so as to hit the target. The purpose of this study is to find how good throwers select the state variables in throwing and elucidate the knack of the ball throwing task.



OS7-5 A low dimensional feedback control model that exploits abundant degrees of freedom

Jun Nishii, Tohru Hamamura
(Yamaguchi University, JAPAN)

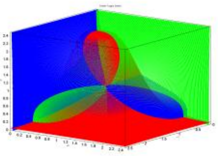
When animals perform skilled motor task, they often adaptively choose a solution depending on the circumstances by utilizing abundant degrees of freedom (DOFs) of their body. For instance, a spinal frog can wipe an irritating stimulus off by its foot even if a leg joint is fixed so as not to be able to bent. In this example, multiple leg joints cooperatively work so as to control the foot position by utilizing available joints. Such compensative control among multiple joints is called joint synergy. We propose a simple neural network model that realizes synergetic control that exploits abundant degrees of freedom.



OS7-6 Extension of Genetic Toggle Switch Based on the Effective Search of State Transitions

M. Sugii, A. Fauré and H. Matsuno
(Yamaguchi University, Japan)

In the context of synthetic biology, artificial genetic circuits are designed in the following way: after setting a biological target phenomenon to be investigated, reaction parameter estimations among related molecules are conducted based on the dynamic analyses with mathematical models. Finally a system of biological reactions is developed with these molecules in vivo or in vitro. We propose a new procedure to effectively design a mathematical model in two steps. The first step is the creation of possible network architectures under a logical formalism. The second step is the creation of dynamic models. We report an extension of a genetic toggle switch from 2-state to 3-state by our proposed procedure.



Nullclines of

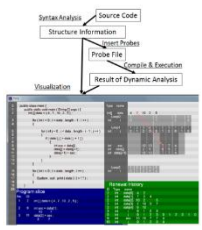
3-state toggles witch

OS8 Software Development Support Method

OS8-1 Proposal of a Visualizing Method of Data Transitions to Support Debugging for Java Programs

Hiroto Nakamura[1], Tetsuro Katayama[1], Yoshihiro Kita[2], Hisaaki Yamaba[1], Naonobu Okazaki[1],
([1]University of Miyazaki, Japan)
([2]Kanagawa Institute of Technology, Japan)

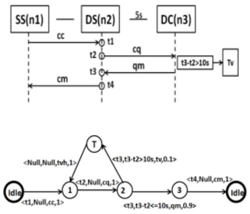
Finding the cause of a bug needs to comprehend a flow and data transitions in executing programs. It is difficult to grasp behavior in executing the programs whose behavior is unexpected by a bug. We propose a visualizing method of data transitions to support debugging for Java programs in order to improve efficiency of debugging by supporting to find the cause of a bug. We have implemented TVIS in order to show efficiency of the proposed method. The data transitions diagram is the most characteristic function of TVIS which shows the data transitions in executing programs as a table. It can show visually abnormal behavior: no data renewed at all, data abnormally renewed, and so on. Because abnormal behavior is detected in the data transitions diagram at first glance, it is useful for programmers in finding the cause of a bug.



OS8-2 Proposal of a Method to Build Markov Chain Usage Model from UML Diagrams for Communication Delay Testing in Distributed Systems

Zhijia Zhao[1], Tetsuro Katayama[1], Yoshihiro Kita[2], Hisaaki Yamaba[1], Naonobu Okazaki[1]
([1]University of Miyazaki, Japan)
([2]Kanagawa Institute of Technology, Japan)

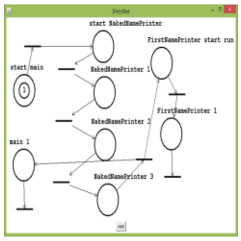
As the growth of network technology with high parallelism and high reliability of distributed systems, they have been widely adopted in the enterprise and society. But, how to design for testing the real-time or a communication delay of distributed systems is not been discussed much. This paper proposes a new method to automatically build Markov Chain Usage Model. The proposed method establishes the time points and builds a new UML diagram called De-sequence diagram by combining the deployment diagram and sequence diagram which can test the communication delay of a distributed system to improve its reliability. We have confirmed the usefulness of the proposed method to adapt it to practical examples.



OS8-3 Proposal of a Supporting Method for Debugging to Reproduce Java Multi-threaded Programs by Petri-Net

Shoichiro Kitano[1], Tetsuro Katayama[1], Yoshihiro Kita[2], Hisaaki Yamaba[1], Naonobu Okazaki[1]
([1]University of Miyazaki, Japan)
([2]Kanagawa Institute of Technology, Japan)

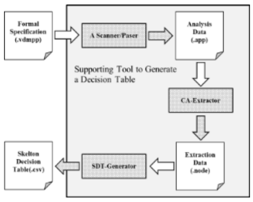
In multi-threaded programs, it is difficult to discover a cause of bugs. This paper proposes a supporting method using Petri-net for debugging so that programers can discover cause of bugs in multi-threaded programs written in Java language. Specifically, in order to give reproducibility to multi-threaded programs, the proposed method generates a data file for an execution path of a multi-threaded program, and simulates the behavior of the program by Petri-net based on the data file. However, ordinal Petri-net cannot express the behavior of multi-threaded programs written in Java completely. Therefore, we have extended Petri-net so that we can treat with the behavior of multi-threaded Java programs. In addition, we have confirmed validity of our method by having implemented a supporting tool.



OS8-4 Proposal of a Supporting Method to Generate a Decision Table from the Formal Specification

Kenta Nishikawa[1], Tetsuro Katayama[1], Yoshihiro Kita[2], Hisaaki Yamaba[1], Naonobu Okazaki[1]
([1]University of Miyazaki, Japan)
([2]Kanagawa Institute of Technology, Japan)

As a means for writing specifications strictly, the formal methods are proposed. By the way, as one of the test design techniques, the decision table is proposed in the testing process of the software cycle. However, it takes much time and effort to extract test items and understand contents written on specifications in designing manually the decision table. This paper proposes a supporting method to generate a decision table from the formal specification in order to improve efficiency of the test design with formal methods. We have implemented a supporting tool to generate a decision table. It automatically generates a skeleton decision table from the formal specification. By using the tool, it is considered that the efficiency of the test design is improved.

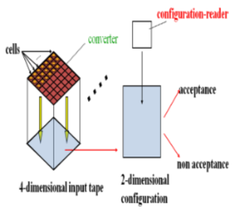


OS9 Foundation of computation and its application

OS9-1 Hierarchy Based on Neighborhood Template about k-Neighborhood Template *A*-Type Three-Dimensional Bounded Cellular Acceptor

Makoto Sakamoto[1], Makoto Nagatomo[1], Xiaoyang Feng[1], Tatsuma Kurogi[1], Tuo Zhang[1], Takao Ito[2], Yasuo Uchida[2], Tsunehiro Yoshinaga[3], Satoshi Ikeda[1], Masahiro Yokomichi[1],Hiroshi Furutani[1]
([1]University of Miyazaki, Japan)
([2]Ube National College of Technology, Japan)
([3]Tokuyama College of Technology, Japan)

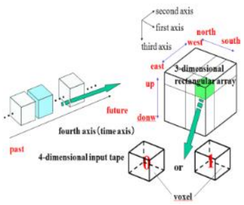
In this paper, we introduce a new computational model, k-neighborhood template *A*-type three-dimensional bounded cellular acceptor on four-dimensional tapes, and discuss some basic properties. This model consists of a pair of a converter and a configuration-reader. The former converts the given four-dimensional tape to three-dimensional configuration. The latter determines whether or not the derived three-dimensional configuration is accepted, and concludes the acceptance or non-acceptance of given four-dimensional tape. We mainly investigate how the difference of the neighborhood template of the converter affects the accepting powers of this computational model. It is well known that three-dimensional digital pictures have 6- and 26-connectedness. However, we include the remarkable picture in neighbor.



OS9-2 Hierarchy Based on Configuration-Reader about k-Neighborhood Template *A*-Type Three-Dimensional Bounded Cellular Acceptor

Makoto Sakamoto[1], Tuo Zhang[1], Tatsuma Kurogi[1], Makoto Nagatomo[1], Xiaoyang Feng[1], Yasuo Uchida[2], Takao Ito[2], Tsunehiro Yoshinaga[3], Satoshi Ikeda[1], Masahiro Yokomichi[1], Hiroshi Furutani[1]
([1]University of Miyazaki, Japan)
([2]Ube National College of Technology, Japan)
([3]Tokuyama College of Technology, Japan)

Question of whether or not processing four-dimensional digital patterns is more difficult than processing two- or three-dimensional ones is of great interest from both theoretical and practical standpoints. Thus, the study of four-dimensional automata as the computational models of four-dimensional pattern processing has been meaningful. From this point of view, we are interested in four-dimensional computational model, and we first proposed four-dimensional automata as computational models of four-dimensional pattern processing in 2002. In this paper, we introduce a new computational model, k-neighborhood template A-type three-dimensional bounded cellular acceptor on four-dimensional tapes, and mainly investigate how the difference of configuration-reader affects the accepting powers of this model.

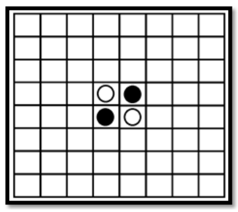


OS9-3 Handicap of Othello Game

Y. Kato[1],S. Ikeda[1] T. Ito[2], M. Sakamoto[1]
([1]University of Miyazaki, Japan)
([2]Ube National College of Technology, Japan)

In the game of Go, it is known experientially that Black (=the first move) is advantageous. Thus, in the game of Go, White (=the defensive hand) has received a handicap for this compensation. There is Othello game well known as the Go. However, the handicap of White in Othello game has been hardly considered until now. This research considered quantitatively the handicap in Othello game from a viewpoint of points. As our conclusions of Othello game is the followings.

- The advantage of the defensive hand on board of 4 × 4 is -0.641 points on average.
- The advantage of the defensive hand on board of 4 × 6 is 0.778 points on average.



OS9-4 On Parameter Setting in Identifying the Same Languages Involved in Different Language Data

Ren Wu[1], Hiroshi Matsuno[2]
([1]Yamaguchi Junior College, Japan)
([2]Yamaguchi University, Japan)

We have proposed a method based on tree structure and string alignment technique for identifying the same languages involved in two language classification trees provided by different linguists. Several kinds of similarity measure, such as language name similarity and language general similarity etc., and an algorithm are proposed in this method. Several unknown parameters are used there and need to be set to constants first for calculating the similarity values. This paper aims to determine all the values of these parameters and then get the identification results in order to confirm that our proposed similarity measures and the algorithm are useful and effective. We obtained better values for the parameters throughout the experiments.

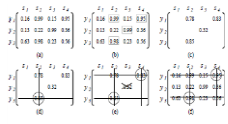


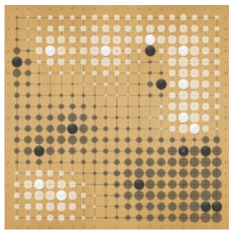
Fig. 4. Process of finding same languages

GS

GS1 Artificial intelligence & Complexity
GS1-1 Cognitive approach to Computer Go programming

N. Tenys, I. Tanev, K. Shimohara
(Doshisha University, Japan)

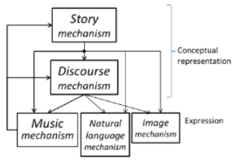
Go (Igo) is an ancient board game that is still one of the more difficult challenges for Artificial Intelligence research. While computers have defeated human champion in chess 16 years ago, to this day best computer Go programs play only at a high amateur level. Humans play Go using experience, heuristic values, pattern recognition and whole-board evaluations that computers are unable to match. Today’s most successful Go programs use Monte Carlo search algorithm, but while this approach often finds good local moves on the board, due to its random nature it often fails to see the whole board position as well as human does. Our approach instead is implementation of human Go professional’s “thinking” into a computer program that is, creating a system, where computer uses algorithms and data constructs that mimic human’s mind.



GS1-2 Circulative Narrative Generation Based on the Mutual Transformation between Narrative Conceptual Structures and Music in the Integrated Narrative Generation System

Taisuke Akimoto, Takashi Ogata
(Iwate Prefectural University, Japan)

We have proposed a new framework of narrative generation which generates narratives through the mutual transformation according to various pathways between narrative (conceptual representation) and music. This paper will continuously incorporate this framework into the integrated narrative generation system which is a comprehensive architecture for our narrative generation study and develop an experimental system combined with the integrated system. This new framework aims to apply various musical methods for composition and variation to narrative generation as a way to connect different types of media. An experiment of the mutual generation between music and narrative will show various transformation pathways or routes between music and narrative.



GS1-3 Dynamic Analysis of Tensegrity Systems Subject to Arbitrary Joint Constraints

Youngsu Cho and Joono Cheong
(Korea Univ., Korea)

This paper provides a dynamics formulation of tensegrity systems using non-minimal coordinates subject to arbitrary joint constraints. Advantages of non-minimal dynamics including the simplicity of formulation and scalability are enjoyed. We start from a matrix form of dynamics and then transform the dynamics into a vector form so that arbitrary constraints in the joints can be handled with ease. Due to the generality of the formulation, we can create a software module that can carry out dynamic simulation of numerous tensegrity systems. To illustrate the effectiveness of the proposed dynamics formulation, we show simulation results for a deployable tensegrity structure and a multi-stage tensegrity prism under various types of joint constraints such as pin, revolute, and planar types.

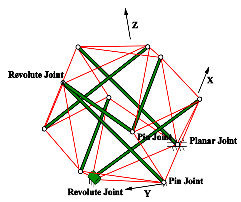


Figure 1. Tensegrity system with constraints

GS1-4 Road map generation from Smartphone data

Masao KUBO, Chau viet DAN, Hiroshi SATO, Akira NAMATAME
(National Defense Academy of Japan)

Smartphones become widespread quickly. This type of mobile phone has a built-in GPS, motion sensor, digital compass and large memory and battery. It is good for recording human activity. In this paper, we introduce a method to make occupancy grid map when the user drive outside. Usually, the vehicle behavior is analyzed by a time series of accelerations sensed by customized sensors installed on vehicle. The data is high quality for understanding but it is difficult to collect data from the public. The data recorded by a smartphone is lower accuracy and the pose is unknown. Our method uses GPS data mainly.

GS1-5 Zigzag-Perceptually Important Points for Financial Time Series Indexing

Chawalsak Phetchanchai, Ali Selamat , Md Hafiz Selamat
(University Teknologi Malaysia, Malysia)

In this paper, we propose a zigzag based m-ary tree (ZM-Tree) to index financial time series. The index is done by mean of zigzag-perceptually important points (ZIPs). The tree represents the zigzag feature of financial time series which benefits in reversal pattern searching, and wave counting. We also propose a method of identifying ZIPs in multi-resolution. The identification process uses the vertical distance as a measurement to evaluate the point importance. The experiments demonstrate the performance of ZM-Tree in tree building, tree retrieval, tree pruning, and dimensionality reduction

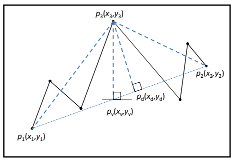


Fig. 2. Measurement of PIP-ED, PIP-PD, and PIP-VD

GS2 Reinforcement Learning System & Genetic programming

GS2-1 Development of phrase and music search engine by humming

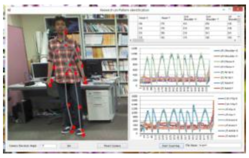
Kiminori Satou, Eiji Hayashi
(Kyushu Institute of Technology, Japan)

In recent years, the internet has become ubiquitous and is changing dramatically every aspect of how people live their lives. Users not only can purchase the CD at home but also can download the music files without leaving home using the internet, it is very easy and convenient for users who can find the lyrics. But the music's names are not able to be efficiently searched if users forget the music's names. Then, enter the voice, do a search using that voice's length and pitch, find the rhythm and pitch similar phrases, and thought that it is possible to determine the name of the song. Therefore, we developed a phrase and music search engine that searches a tune from various tunes with one phrase of voice data.

GS2- 2 Human Recognition based on Gait feature and Genetic Programming

Dipak Gaire Sharma, Ivan Tanev, Katsunori Shimohara
(Doshisha University, Japan)

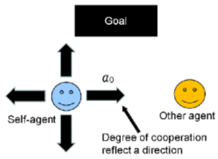
Human walking has always been the curious field of research for different disciple of Social and Information Science. The study of human walk or human gait in association with different behaviors and emotions has not only fascinated social science researcher, but its uniqueness has also attracted many computer scientist to work in this arena for the quest of uncovering reliable mechanism of biometric identification. In this research we use new method for human identification based on human gait features and genetic programming (GP) approach.



GS2-3 An Action Selection Method Using Degree of Cooperation in a Multi-agent System

Masanori Kawamura, Kunikazu Kobayashi
(Aichi Prefectural University, Japan)

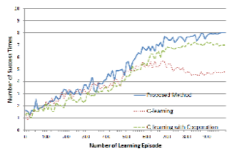
In recent years, a concept of a dividual is proposed to interact properly with another person. To construct a model of the dividual, the degree of cooperation is assigned to the corresponding dividual. By introducing the degree of cooperation into multi-agent systems, we evaluate what kind of changes appears in the agent behavior. In addition, we propose an action selection method by introducing the degree of cooperation into the soft-max method in multi-agent systems. Using the proposed method, we confirm whether the cooperative action is promoted or suppressed through computer simulations.



GS2-4 Cooperative Action Acquisition Based on Intention Estimation Method in a Multi-agent Reinforcement Learning System

Tatsuya Tsubakimoto, Kunikazu Kobayashi
(Aichi Prefectural University, Japan)

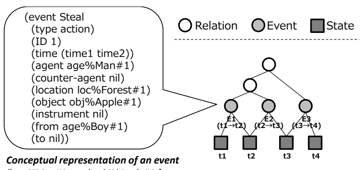
In this paper, we propose a method that can acquire cooperative action to reach an appropriate goal without controlling reward to realize cooperative action by designers. We assume every action of other agents as a process to achieve an ultimate goal and then treat it unimportant. In order to confirm the effectiveness of the proposed method, we carried out computer simulations. The simulation results show that the proposed method is superior to a standard Q-learning method and a Q-learning method with cooperation in terms of the number of successful cooperation.



GS2-5 Integrating the Event Generation Mechanism in the Propp-based Story Generation Mechanism into the Integrated Narrative Generation System

Shohei Imabuchi, Takashi Ogata
(Iwate Prefectural University, Japan)

In the Propp-based story generation mechanisms and the integrated narrative generation system which we have been developing, the former functions as a module in the latter. These systems have used respective event generation mechanisms to generated events which are the most important units in a narrative structural representation. A common event generation mechanism needs to be used in these two systems towards the complete blending. This paper will present the first tentative attempt as a prototype to be revised in the future.



Conceptual representation of an event
"age%Man#1 steals obj%Apple#1 from age%Boy#1 at loc%Forest#1."

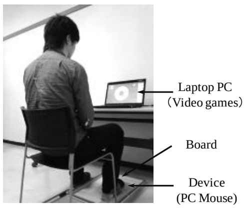
Conceptual structure of a story

GS3 Human-machine cooperative systems & Human welfare robotics

GS3-1 Development of Measurement System for Quantitative Evaluation of Skillfulness of Lower Extremities

Kazunori Yamazaki[1], Donggun Kim[1], Yoshifumi Morita[1], Hiroyuki Ukai[1]
Kenji Kozakai[2], Satoru Shibata[2], Shigenori Onishi[2], Akihiro Ito[2], Daisuke Mizuno[2]
Tatsuya Hirai[3], Haruna Takeda[3], Yuka Sugiura[3]
([1]Nagoya Institute of Technology, Japan),
([2]Sanyo Machine Works, Ltd., Japan),
([3]Tanakakai Nishio Hospital, Japan)

In this study, we developed a quantitative evaluation method of the skillfulness of the lower extremities. For this purpose, we used the measurement system which we have developed in our previous work and redesigned the target trajectories of foot movement. Moreover, we verified the reliability of the proposed quantitative evaluation method by calculating the intra-subject reliability for two trials ICC(1,1) of the evaluation results. We developed a measurement system comprising the straight course, the circular course and the star-shaped course. Because the ICC(1,1) of the circular course and the star-shaped course are higher than 0.7, the quantitative evaluation method of lower extremities with the circular course and the star-shaped course has sufficient reliability.



GS3-2 Glove-Based Virtual Interaction for the Rehabilitation of Hemiparesis Stroke Patient

Khairunizam WAN, Aswad A.R, Rashidah Suhaimi, Nazrul H. ADNAN, D. Hazry, Zuradzman M. Razlan, Syed Faiz Ahmed
(Universiti Malaysia Perlis, Malaysia)

This paper proposes the rehabilitation of a stroke patient by using dataglove called *GloveMAP*. *GloveMAP* is a low cost hand glove developed by using flexible bending sensor and accelerometer attached at the position of the fingers and wrist, respectively. The rehabilitation exercises are conducted in the virtual environment. In the studies, 22 fundamental movements of arm are evaluated and the results are employed to design the virtual environment. The combination of these fundamental movements focuses on Hemiparesis sufferers which are the most common in the stroke and have the inability to move one side of the body. The experimental results show that by combining several fundamental arm movements, glove based-virtual interaction supports the stroke patient do the rehabilitation by them without assisting from physiotherapist.

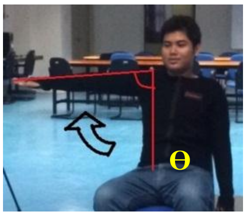
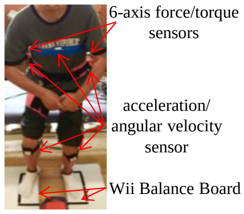


Figure. 1 Performing the fundamental arm movement

GS3-3 Measuring system of therapist's guiding motion for standing up training of patient with hemiplegia after stroke and analysis results of patient's motion

JungTang Wang[1], Kazunori Yamazaki[1], Yoshifumi Morita[1], Noritaka Sato[1], Hirofumi Tanabe[2]
([1]Nagoya Institute of Technology, Japan)
([2]Self-Defense Force Central Hospital, Japan)

In this paper we developed a measuring system to analyze the therapist's guiding motion during the standing up training of a hemiplegic stroke patient. To confirm the usefulness of the developed measuring system, we measured the motion of a hemiplegic stroke patient during the standing up training with/without the therapist's guiding motion. By comparing the results of the patient's left/right weight balance during the standing up motion with/without the therapist's guiding motion, the effectiveness of the therapist's guiding motion was confirmed. Moreover, the measured patient's motion with/without the therapist's guiding motion agrees with the therapist's observation. Therefore, the usefulness of the measuring system was confirmed.



GS3-4 Image Segmentation of Coronary Artery Plaque Using Intuitionistic Fuzzy C-Means Algorithm

Zahra Rezaeia, Ali Selamata, Mohd Shafry Mohd Rahima, Mohammed Rafiq Abdul Kadirb
(Universiti Teknologi Malaysia)

Every year, hundreds of thousands of people die because of Coronary Heart Disease (CHD) in all over the world..Coronary Artery Disease (CAD) as a cardiovascular illness causes blood vessels narrowing that supply blood and oxygen to the heart . Atherosclerosis is known as the deadliest type of heart disease, which is caused by soft or "vulnerable" plaque (VP) formation in the coronary arteries. Acute Coronary Syndrome (ACS) is recognized as the first coronary atherosclerosis indicator which identifies high-risk plaques. Intravascular ultrasound (IVUS) can be applied for characterization of plaque and segmentation of vessel's walls borders. Recently, Virtual Histology as a new approach based on spectral analysis of IVUS has been introduced. In this work, we applied a clustering method based on Intuitionistic Fuzzy C-means (IFCM) in order to automatic segmentation of Coronary Artery plaque using VH-IVUS images.

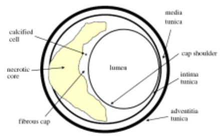


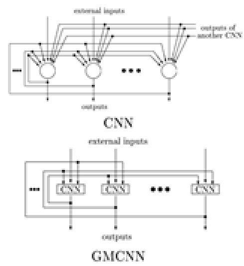
Figure 1. Arterial plaque geometry. Cross-sectional view [2].

GS4 Neural Networks

GS4-1 The recollection characteristics of Generalized MCNNusing different control methods

Shun Watanabe[1], Takashi Kuremoto[1], Kunikazu Kobayashi[2], Shingo Mabu[1], Masanao Obayashi[1]
([1]Yamaguchi University, Japan)
([2]Aichi Prefectural University, Japan)

As an auto-association memory model, chaotic neural network (CNN) proposed by Adachi et al. is able to recollect multiple stored patterns dynamically. Kuremoto et al. proposed a multi-layer chaotic neural network (MCNN) combined multiple controlled CNN layers to realize mutual association of plural time series patterns. However, recollection simulation of MCNN was limited in a two-layer model and the recollection characteristics concerning with the different external inputs and the control methods were not investigated. In this paper, we extend the MCNN to be a general form (GMCNN) with more layers and use control method using particle swarm optimization (PSO). And the recollecting characteristics of different GMCNNs using different control methods were invested by computer simulations.



GS4-2 Revealing Terrorism Contents form Web Page Using Frequency Weighting Techniques

Thabit Sabbah, Ali Selamat, Md Hafiz Selamat
(Universiti Teknologi Malaysia, Malaysia)

Intelligence and security informatics plays an important role in revealing terrorism in web content. However, the extremist groups are exploiting Internet facilities intensively to stimulate violence, hatred, and terrorism. Accurate classification of web content will increase the opportunities for the effective use of intelligence and security informatics in the early detection of terrorist activities. In this paper, we propose a feature selection method based on term weighting techniques TF, DF, and TF-IDF for revealing terrorism in web content. Firstly, we compare the performance of the individual term weighting techniques in revealing terrorism. Then, we present a method for enhancing the accuracy of classification. The method was tested with selected dataset from Dark Web Portal Forum. The experimental results show that classification using

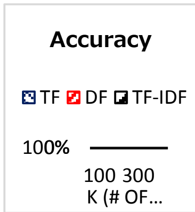
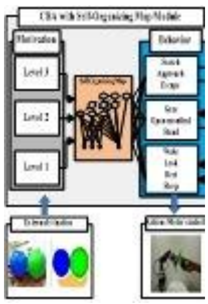


Fig. 1. Classification accuracy for individual term weighting techniques

GS4-3 Application of Self Organizing Map for analyzing of robotic arm's action with Consciousness-Based Architecture module

Wisanu Jitviriya, Eiji Hayashi
(Kyushu Institute of Technology, Japan)

Our research has been focused on developing human-robot interactions, The robot need to perform a high level of intellectual activity and user compatibility. Therefore, we considered the primary structure of a conscious human/animal action, which can be represented by the sequence process, Perception → Motivation → Action. We have improved a hierarchical structure model, which defines the relationship between the consciousness field and the behavior module. This model is called Consciousness-Based Architecture (CBA). Furthermore, the robot should select the action itself, we have investigated the application of brain-inspired technology so we introduced a Self Organizing Map (SOM) neural network that is trained using unsupervised learning to classify its behavior according to the motivation value in the CBA module. In this paper, we attempt to describe the integration of a Self Organizing Map (SOM) method into CBA module in order to classify and select autonomous behavior. We confirmed the effectiveness of the proposed system by the experimental results.



GS4-4 Study on Time Synchronization Algorithm of Wireless Sensor Networks Based on Maximum Likelihood Estimation

Ai Dongmei, Zeng Guangping, Tu Xuyan, He Di
(University of Science and Technology Beijing, China)

Wireless Sensor Network (WSN) integrates technologies of wireless communication, sensing, network interconnection, and distribution. Comprised of a large amount of micro sensor networks, it can collect real-time information, monitor objects within a specific range, and send back usable information at the first time to observers. Unlike general computer networks, WSNs face the issue of constraints on energy. Hence, the time synchronization mechanism, a supporting technology for application of WSNs, isn't suitable when extensively applied to traditional network. This paper proposes High-Accurate Energy-Efficient Time Synchronization (HAEE) algorithm based on TPSN algorithm, which significantly improves the accuracy of time synchronization and reduces energy consumption, as proven by software simulation results.

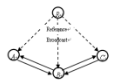


Fig. 1: Message Exchange in RBS

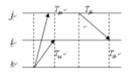


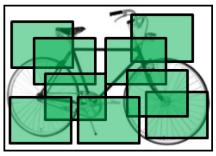
Fig. 2: RBS Synchronization Algorithm Message Transmission

GS5 Pattern recognition

GS5-1 Recognizing a Road Environment Using Multiple-window Bag of Features

Shou Morita, Joo Kooi Tan, Hyoungseop Kim, Seiji Ishikawa
(Kyushu Institute of Technology, Japan)

In this paper, we propose a method of recognizing a road environment employing bag of features considering positional information. This technique works by making multiple windows in an image. The proposed method makes a codebook in each window and computes histograms of local features found inside of the window. The original bag of features method disregards positional information in an image It is, however, important to recognition. The proposed method considers positional information in an image by introducing multiple windows. For recognition, we employ the support vector machine. Experimental results show satisfactory performance of the proposed method.



Multiple windows on an image.

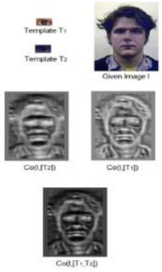
GS5-2 Recognized Face Tracking for CONBE Robot

Sakmongkon Chumkamon, Eiji Hayashi
(Kyushu Institute of Technology)

In our research we develop the robot to combine with the consciousness and behavior which is the conscious and behavior robot (CONBE). In this paper we present the basis system in the CONBE robot that consist of the two main sections as the face recognition and robot gaze tracking. This system uses the face recognition using Fisherfaces that also refer to linear discriminant analysis (LDA). For the robot gaze tracking, we implement with the head of the CONBE robot that has a camera, two displays used for robot eyes and the actuators of 2 degrees of freedom. The actuators use to control the robot gaze for tracking the face which the robot can recognize. In this paper, we experiment and present the results of the face recognition system such as frame rate, error rate; and the tracking control system such as the tracking error and time.

GS5-3 Eye detection Using Composite Cross-Correlation form Face Images in Varied Illumination

Kutiba Nanaa, Mohamed Rizon, Mohd Nordin Abd Rahman
(Universiti Sultan Zainal Abidin (UniSZA), Malaysia)

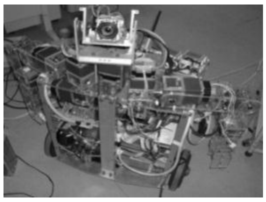


Facial feature detection is essential process in the field of face recognition. Template matc h based approach is widely used to implement the issue of facial detection.However, man y obstructions adversely affect in template based implementation such as varied illumination, size, face pose and the status of eye in either open case or close case. the objective of this paper is to detect left eye and overcome the obstruction of low illumination in given image using composite cross-correlation. Two single template form the proposed composite template. First single template is created from an eye image in ordinary illumination while the second single template is created from an eye image in low illumination. The experimental results is concluded by applying our t emplates on the PICS Database whose varied illumination images and show thatapplying composite templates gives a better result than applying the both single templates individu ally

GS5-4 Development of an autonomous-drive personal robot "An object recognition system using BoF and SOM"

Keisuke Ito, Eiji Hayashi (Kyusyu Institute of Technology, Japan)

In the near future, autonomous self-driving robots are expected to provide various services in human living environments. The ability to work autonomously and accurately recognize surrounding objects are required for the autonomous robot. In previous research, the robot has enabled the determination of the angle and type of the object by a matching system based on the Speeded Up Robust Features (SURF). In this paper, the authors developed the system in order to determine the type of the object by using the Self-Organizing map (SOM) and SURF features. This system can visualize the relationship of the object and determine the type of the object.

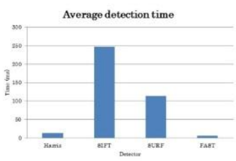


Please attach one typical figure in your paper here.

GS5-5 Comparison of Feature Detectors for Obstacles Detection

Shaohua Qian[1], Joo Kooi Tan[1], Hyoungseop Kim[1], Seiji Ishikawa[1]
Takashi Morie[1], Takashi Shinomiya[2]
([1]Kyushu Institute of Technology, Japan)
([2]Japan University of Economics, Japan)

Detection of obstacles in a video sequence is a basic task in autonomous collision avoidance systems of intelligent vehicle. We have already proposed an obstacles detection method using a video taken by a vehicle-mounted monocular camera. In this method, accurate obstacles detection depends on whether we can accurately detect and match feature points. In order to improve the accuracy of obstacles detection, in this paper we make a comparison between four most commonly used feature detectors: Harris, SIFT, SURF and FAST detectors. The experiments are done using our obstacles detection method. The experimental results are compared and discussed, and then we find the most suitable feature point detector for our obstacles detection method.



GS6 Robotics Ⅰ

GS6-1 Study on the pose under complex multiple targets environment for the industrial robot based on machine vision

Jiwu Wang [1], Xianwen Zhang [1], Huazhe Dou [1], Sugisaka Masanori [2]
([1] Beijing Jiaotong University, China)
([2] Nippon Bunri University, Japan)

Effectively obtaining multiple targets pose is the key element to improve flexible applications of industrial robots under complicated working environments. Machine vision technology is the main method to solve this problem. In order to effectively reduce influence of changing conditions (such as the illumination, angle and scale etc.) on target pattern recognition, the fusion algorithms of scale invariant feature transform (SIFT) and geometric moment, is developed in this paper. To achieve accurate identification of multi-targets' pose when their position and orientation change in the two-dimensional space, the optical-section method is used to determine the depth information of some feature points of the target. Moreover, that data is used to correct the pose recognized by the traditional pattern technology. Finally, using SCARA-type industrial robots as a platform, the relative algorithms are verified with corresponding experiments.

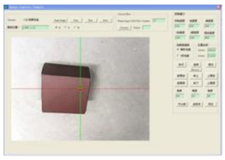


Fig.5. The industrial robot recognition system

GS6-2 Research and Development of an Intelligent Robot of Medical Assistance Based on Embedded RTOS

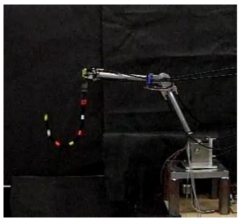
Yuan Li, Fengzhi Dai, Junhong Xi, Binyao Li, Guodong You
(Tianjin University of Science and Technology, P. R. China)

For practical purpose, this paper focuses on design and implementation of a wheeled intelligent robot that could be applied in hospitals or nursing homes for medical assistance. The hardware part includes two microcontrollers, MC9S12XS128 and STC89C52, assisted by a laptop as the control center of the system, various sensors and other robotic pieces. The software part is basically achieved on CodeWarrior 5.0, Keil μVision4 and Visual C++ 6.0 with OpenCV library, porting μC/OS-II real-time operating system to MC9S12XS128 in order for improving stability and flexibility of the system. The robot has functions of infrared remote control, fuzzy rule-based path tracking, obstacle avoidance, real-time LCD display and voice broadcast, color recognition and robot arm grasping.

GS6-3 High Acceleration Robotic Arm for Dynamic and Dexterous Manipulation of Deformable Object

Hiroaki Seki[1], Hiroki Shibata[1], Yoshitsugu Kamiya[1], Masatoshi Hikizu[1], Khairul Salleh Mohamed Sahari[2]
([1]Kanazawa University, Japan)
([2]Universiti Tenaga Nasional, Malaysia)

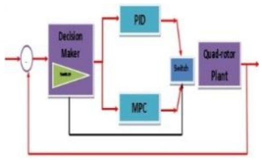
Deformable object manipulation is very important for home service robots. This paper discusses the development of a high acceleration robotic arm capable of performing dexterous manipulation of deformable object such as string. A mass-spring model of the string is first developed to study the design requirement of the robot as well as to determine the required motion of the robot to smartly manipulate the string. The developed system comprising of the string model and robotic arm is then used to manipulate unknown string. Based on some simple basic routine, the parameters of the string can be approximated based on the string model. Once the parameters are known, the string can be smartly manipulated by the robotic arm.



GS6-4 Switching based controller algorithm design for uncertainties rejection in smooth takeoff / landing and of quad-rotor

M. Hassan Tanveer, D. Hazry, S. Faiz Ahmed, M. Kamran Joyo, Faizan. A. Warsi
(Universiti Malaysia Perlis (UniMAP) Malaysia)

This article evaluates the proposed control technique for altitude controlling of Quad-rotor type unmanned Aerial Vehicles under different uncertainty conditions. Smooth takeoff and landing i.e. altitude controlling of Quad-rotor UAV under perturbed conditions such as high wind burst and system or sensor noises has been a challenging research domain for the researchers. In this paper a Hybrid controller is purposed for batter controlling of quad-rotor's altitude through MPC and PID controller. This hybrid controller works in such a way that quad-rotor's sensor and system noises problem are handled by MPC controller and in the occurrence of external disturbances, auto-tuned PID controller is activated and well manage the situation. Simulations are done on MATLAB and results shows the effectiveness of the proposed hybrid method and suggest it as a promising way for real time implementation in altitude stabilization for quad-rotor UAV

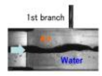


Proposed Hybrid Controller:

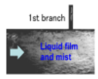
GS6-5 Parameters of Gas-Liquid Flow Distribution Uniformity in Upward Multi-Pass Channels

Zuradzman M. Razlan[1,], R. Heng[1], D. Hazry[1], A.B. Shahriman[1], Khairunizam WAN[1], S. Faiz Ahmed[1], Nazrul H. ADNAN[1], M. Hirota[2], N. Maruyama[2], A. Nishimura[2], H. Hisyam[3].
([1] Universiti Malaysia Perlis, Malaysia)
([2] Mie University, Japan)
([3]Lean Applied Pte. Ltd., Malaysia)

The gas-liquid flow distributions in multi-pass upward parallel channels that simulate the evaporator for an electric vehicle air-conditioner system were examined experimentally. In this paper, the attentions are to discover the most influenced parameter to the flow distribution uniformity by using design of experiment method. Experiments were conducted in an isothermal air-water flow system. In the mist-flow inlet, the water distribution was insensitive to the backpressure conditions and its uniformity was improved in comparison with that in the stratified-flow inlet. The flow distribution uniformity for gas phase is influenced mostly by superficial air velocity, and the flow distribution uniformity of liquid phase is mostly influence by 2-way interaction of parameters which are flow pattern and superficial air velocity.



(a) Stratified-flow inlet



(b) Mist-flow inlet

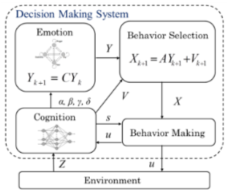
Fig. 1 Flow-inlet conditions

GS7 Robotics Ⅱ

GS7-1 Decision Making System of Robots Introducing a Re-construction of Emotions Based on Their Own Experiences

Shogo Watada, Masanao[1] Obayashi[1], Takashi Kuremoto[1], Shingo Mabu[1], Kunikazu Kobayashi[2]
([1]Yamaguchi University, Japan)
([2]Aichi Prefectural University, Japan)

We already proposed a decision making system for autonomous robots, which is based on a Markovian emotional model proposed by Banik. In our proposed system, decision making of the robot is designed automatically by two processes. The first is a construction of emotions based on learning of the input stimuli. The second is optimization of behavior selection policy for emotions by optimization of system parameters which define behavior selection probability of each emotion. In our above previous studies, a construction of emotions based on learning of the input stimuli was processed using learning samples of input stimuli on offline in advance. In this study, we introduce an experiences-based reconstruction of emotions to our decision making system as an approach to an automatic construction of decision making online learning.



GS7-2 Visual Servoing and Sound Localization in a Surveillance Robot

So-Yeon Park, Yeoun-Jae Kim and Ju-Jang Lee
(KAIST, Korea)

The purpose of this paper is to explain the control and implementation details of the visual servoing and sound localization in a tank shaped surveillance robot. With an eye in hand camera configuration, the key point of the visual servoing is to measure the depth to the target object. In this paper, we suggested the depth estimation technology by moving the camera in a lateral plane. In sound localization, the hardware and software details are fully explained to implement the sound localization technology. Especially the electronic filter design is implemented by utilizing the genetic algorithm.

GS7-3 Study on the improvement of flexibility for an industrial robot based on machine vision

Jiwu Wang[1], Xianwen Zhang[1], Weining Zhang[1], Sugisaka Masanori[2]
([1] Beijing Jiaotong University, China)
([2] Nippon Bunri University,)

In order to improve the applications for an industrial robot, it is necessary to increase its flexibility and control accuracy. Many algorithms and control strategies are developed, in which machine vision gets more and more attention. This is because the CCD camera can provide more information than other sensors, which makes the whole control process simple. But it requires that the target position should be extracted accurately and robustly. Moreover, the frequency of image processing should be satisfied with the control requirements. Here an industrial robot arm is designed and set up for experiment simulation with machine vision. The target extraction method with fusion of color space and moment invariants is developed. The experiments results showed that the developed image processing algorithms are robust, and the flexibility of an industrial robot can be improved by machine vision.

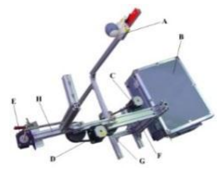
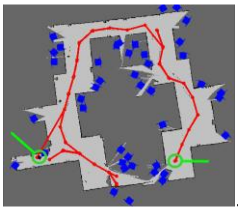


Fig.1. The illustration of the structure of the arm robot

GS7-4 Improved Map Generation by Addition of Gaussian Noise for Indoor SLAM using ROS

Khairul Salleh Mohamed Sahari, Barry Loh Tze Yuen
(Universiti Tenaga Nasional, Malaysia)

Rao-Blackwellized Particle Filter (RBPF) is used in this paper to solve the Simultaneous Localization and Mapping (SLAM) problem. RBPF algorithm uses particle filter where each particle carries an individual map of the environment. With the usage of Robot Operating System (ROS), GMapping package was used as a basis for map generation and SLAM. To improve the map generation, Gaussian noise was introduced to the data from laser range finder and also the odometry from the robot Pioneer P3AT's pose. The introduced algorithm was successful in decreasing the uncertainty as well as increased the knowledge of each particle in the estimation of the robot's pose, proven through practical experiment. Exploration experiments were also carried out to test the performance of P3AT based on our proposed method.

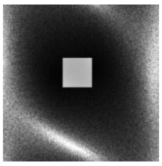


PS

PS1-1 Relationship input object position and optimal error diffusion coefficients for Kinoform using error diffusion method

Daisuke Kashima. Ken-ichi Tanaka.
(Meiji Univ. Japan)

We performed simplify defining optimal error diffusion coefficients for Kinoform. In order to define optimal error diffusion coefficients simply, we performed searching relationship input object position and formulating the relationship to approximation. In order to prove that approximation is true, this paper show simulation result of making reconstructing by error diffusion method.



PS1-2 Optimization of dither matrix by hybrid of Genetic Algorithm and Simulated Annealing

Kohei Kato, Ken-ichi Tanaka
(Meiji University)

Halftoning is used by ink-jet printer, color copier and so on. Dither method is the method of halftoning. Bayer method that is one of the dither methods is able to obtain a superior evaluation value compare with other dither methods, though it is the decided matrix. Several investigations of Optimization of the dither matrix are reported. However, there is no investigation from point of view of the combinational optimization problems. In this paper, we search the optimum dither matrix by Genetic Algorithm (GA) and Simulated Annealing (SA) and hybrid of GA and SA(GA/SA).We weigh conventional method against these methods.

PS1-3 Inverse Halftoning using Multi-Layer Feed-Forward Neural Network

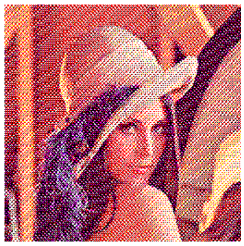
Hiroki Hamashoji, Ken-ichi Tanaka
(Meiji University)

In this paper, we proposed novel inverse halftoning technique, which use Multi-Layer Feed-Forward Neural Network (MLFFNN) and Gaussian filtering. This method reconstruct continuous tone image from any halftoning methods. As a result, a high quality image can reconstructed from binary image compared with conventional method. Especially the proposed method image quality is good both visually and Peak Signal-to-Noise Ratio (PSNR) value, compared with reconstructs inverse halftone images by Look-up table (LUT) based inverse halftoning.

PS1-4 Embedding of the Confidential Image in a Dithered Color Image

Keisuke Banba, Kenichi Tanaka
(Meiji University, Japan)

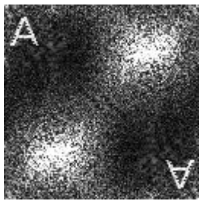
This paper proposed embedding information in dithered color image. We have developed a method of encryption using computer-generated holograms (CGH) embedded in dithered color image. First, confidential information is converted into a CGH. Next, the CGH data undergo two separate dithering processes in parallel: one corresponding to CGH white pixels and one corresponding to CGH black pixels. The results from both processes are used to form a dither matrix for creating the final dithered and encoded image. In this way, confidential information can be embedded into the image. The dither method is algorithm used for digital halftone processing. We evaluated image quality of a dithered color image, and evaluation value of the image is superior to one of conventional method.



PS1-5 Three Dimensional Images Reconstruction on Computer Generated Holograms the Multiple Regression Analysis

Kenta Ayabe, Ken-ichi Tanaka
(Meiji University, Japan)

In this paper, for a Computer Generated Hologram (CGH) reconstructing a three-Dimensional (3D) object, we present the improvement of the reconstruction image, reducing of the calculation time using error diffusion method and a multiple regression analysis. The multiple regression analysis predicts one response variable with plural explanatory variables. The purposes of this paper are reducing calculation time and improving reconstruction images. Therefore we move 26 pieces of images from A to Z at various positions. We prepare for 650 pieces of images in total and perform GA in each.We expect diffusivity for a multiple regression analysis from the result and regenerate it..



PS1-6 **Embedding used binary number conversion into a Color image**

Hiroaki Oguma, Ken-ichi Tanaka
(Meiji University, Japan)

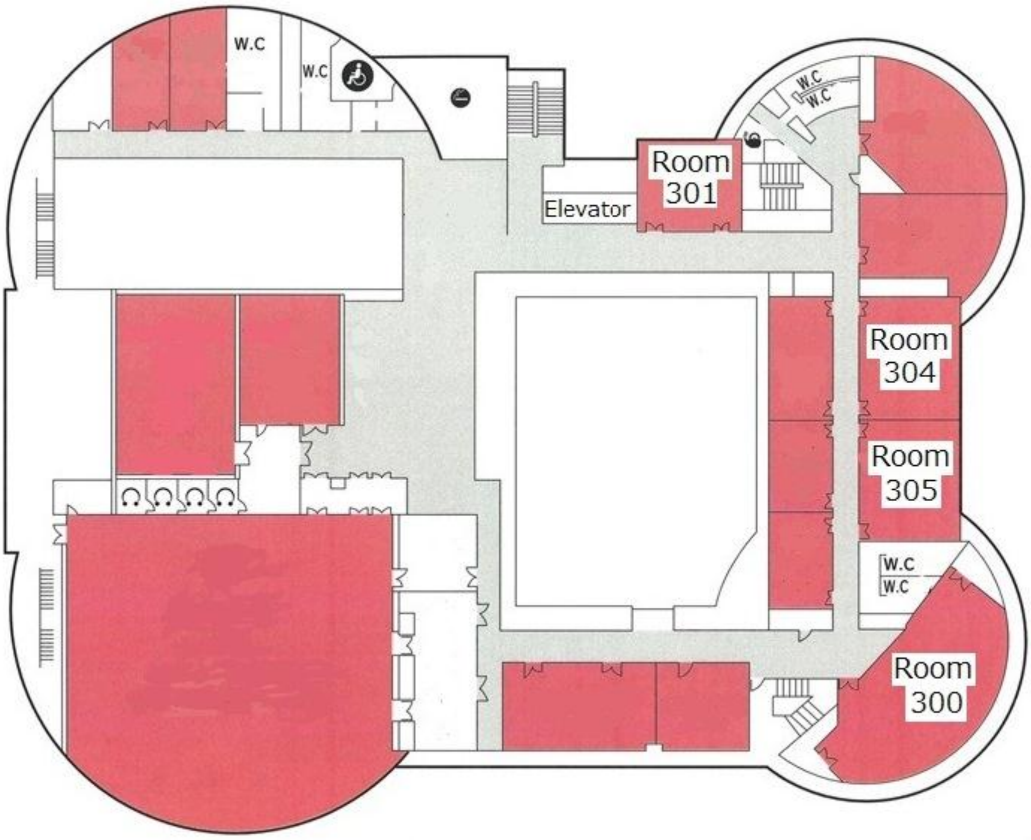
Digital watermarking is widely used as security of a binary format image. The purpose of this paper is extension to a color image from grayscale image. Then, it thought as important giving many embedding information to a color image. A color image is divided into RGB, the concentration value of a pixel is made into a binary number, and last 3 digits are changed to arbitrary numbers. It embeds by carrying out by repeating this, and secure many information. 10x19660 times of embedding are able to be performed to the picture of 256x256 pixels. Moreover, the signal to noise ratio is about 40 db. Therefore, it seems that there is little influence which it has on a picture.

AUTHORS INDEX

Notation of session name

IS: Invited session,OS: Organized Session, GS: General Session, PS: Poster Session

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


W.C
W.C
Room
301
Elevator
W.C
W.C
Room
304
Room
305
W.C
W.C
Room
300

Playware Research – Methodological Considerations

Henrik Hautop Lund

Centre for Playware, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

hhl@playware.dtu.dk
www.playware.dk

Abstract: Several sub-disciplines of engineering are driven by the researchers' aim of providing positive change to the society through their engineering. These researchers are challenged by the traditional research method of experimental research with a waterfall model which demands clearly defined project definition and functional requirements, and impose a sequential processes leading to the final system evaluation, which may lead to solutions which work in the lab, but have little impact in the messy real world. Based on two decades research in developing engineering systems with a societal impact (e.g. in robotics, embodied AI, and playware), in this paper we suggest a cyclic research method based on a mix between participatory and experimental processes. In particular, inspiration from the action research method applied to interdisciplinary technology development becomes a participatory approach characterized by rapid prototyping cycles which allow iterative technology specification and development together with people in their real world environment. With the mixed research method, we suggest that there are cases, where approaches from the positivistic and interpretivistic epistemologies can and should be merged. We exemplify this with health care technology and playware.

Keywords: Playware, Research Method, Synthesis, Rapid Prototyping, Modular Technology.

I. INTRODUCTION

Development of different research methods have often been neglected in the engineering discipline, since the discipline has followed the traditional natural science research methods of performing experimental research of a positivist nature. This involves rigor in controlling parameters keeping specific parameters constant and others variable in empirical testing. Hence, the research method arises from a reductionist belief that systems can be described fully by its components and their interaction with a fixed and constant environment. This "ideal" view of the world often leads the engineer to a conventional development approach termed as the waterfall model [1] with clearly identifiable and separated sequential processes of 1) Project definition, 2) Functional requirements, 3) Functional design, 4) Implementation design, 5) System assembly, and 6) System evaluation.

However, engineering is also about creating and bringing technological solutions to people and the society. Several sub-disciplines of engineering are driven by the researchers' aim of providing positive change to the society through their engineering. These researchers are challenged by the traditional research method. Insight from social science, humanities, and arts tells us that people and the society may not be as rigorously controllable as lab conditions, industrial factory halls, etc. where the conventional engineering research method and the waterfall model has been applied with success. In the wild, messy real-world, there may be a need for another research method for researchers who aim at developing technological solutions that have a deep impact on people in their daily life and on the society.

Here we suggest a cyclic action research method based on a mix between participatory and experimental processes. In particular, inspiration from the action research method applied to interdisciplinary technology development becomes a participatory approach characterized by rapid prototyping cycles which allows iterative technology specification and development together with people in their real world environment.

II. MIXED RESEARCH METHOD

With the mixed research method, we suggest that there are cases, where approaches from the positivistic and interpretivistic epistemologies can and should be merged. Health care technology is one such case. Health care technology aims at providing new technological solutions for the public health care sector, e.g. for hospitals, for rehabilitation and care centres, and for care in the private home. Developing such solutions demand a broad knowledge about health and disease, engineering competencies, and interdisciplinary competencies to develop health care solutions. The mixed action research method is therefore particularly well suited for performing such research in health care technology.

Traditional action research methods for addressing the needs and practice in care centers and private homes are clearly different from traditional research methods in medicine of performing experimental research e.g. effect studies through a randomized controlled trial. However, both issues are crucial to include in research which aims at having an impact on society. Without an iterative participatory development, it is unlikely that users will see the technology fit into their lives and practice, whereas it is unlikely that health authorities will accept the technology without experimental research according to the acknowledged protocols within the field of medicine. There is little chance that our society will accept technological health care solutions without the rigorously controlled experiments that provide evidence of effect and potential collateral effects. Hence, a mixed research method is a necessity to ensure impact in the society. Indeed, it would be clearly against the aims of action research of creating society improvements not to include such rigorously controlled experiments in sub-processes of the research method, since these experiments are the cornerstone for the society to accept the health care technology. Hence, in such a mixed action research, some sub-processes in the iterations can be performed in a participatory manner, whereas other sub-processes must be performed through controlled experiments.

Other examples include contextualized IT training and community-based rehabilitation, which by their nature are close to action research and participatory methods. Nevertheless, in the case of contextualized IT training there may still be a need to perform experimental research on technology and educational outcomes in order to obtain impact in society, and in the case of community-based rehabilitation, there is the need to perform rigorous effect studies similar to those for other health care technologies.

III. APPLYING THE METHOD

As an example of applying this mixed research method, let us look at the research field of playware [2, 3]. Typically, playware research has a core technology research activity focusing on research into modular playware technology and its supporting fundamental research areas of modern AI, adaptivity, modular robotics, and tangible interfaces. Engineering researchers perform fundamental research in these areas to develop the basis for understanding and creating user-interactive technological systems. To operationalise this knowledge and develop prototypes with users, the research method takes its point of departure from the technological discipline *and* the humanistic discipline investigating, understanding and exploiting these scientific fields (modern artificial intelligence, modular robotics, and tangible interaction) in combination with an understanding of play and play dynamics, and in more general terms play culture and human motivation. The knowledge is *combined in synthesis* to develop design principles, prototypes, and demonstrators of playware technology, and a choice of these demonstrators are selected for (e.g. industrial) refinement to become modular playware technology products. The resulting modular playware technology is systematically studied and investigated in the messy real-world environments in order to guide the next cycle of synthesis, demonstrator and prototype development. Hence, the research method is based upon the interdisciplinary research on how playware in the form of intelligent hardware and software creates play and playful experiences amongst the users in their environment.

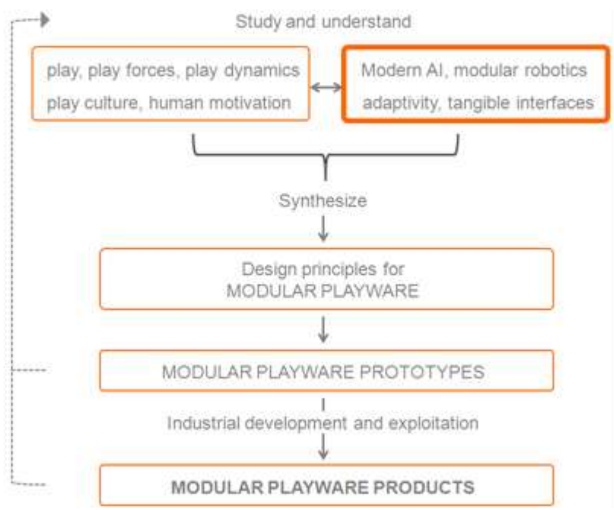


Figure 1. The combination of several research disciplines in synthesis in an iterative working process

Within the individual cycle, the systematic study and investigation of prototypes and products in the messy real-world environments can follow an experimental research method and/or an interpretive research method. Indeed, often it may be necessary to perform a number of studies with a plurality of methods, for instance in order to be able to investigate both the technological and the practical feasibility in the environment. There are issues regarding technological stability and robustness which lends itself best to experimental research of a positivist nature, whereas issues regarding human interaction may lend themselves best to participatory research of an interpretivist nature. The combined knowledge from the two research directions' investigations will be the guide for the next cycle of synthesis, demonstrator and prototype development. Combining and weighing the knowledge from the two research directions is challenging and by no means a trivial task. Therefore, the research method demands that participants build shared knowledge and language about the environment, and research focuses on how results from the two directions are to be combined. The research method facilitates this combination by focusing on synthesis in the iterative process.

IV. ITERATIVE APPROACH

The example of applying the mixed research method to playware research points to an iterative approach in which knowledge is built from iterations of synthesis and application in the environment. Indeed, Lewin [4] (1946) originally pointed to a spiral of steps in action research, and the research method has been known as cyclic action research. In general, the cyclic action research involves identification of a practical problem, making a solution, and reflections about the solution, which then leads to the next iteration of identification, solution, and reflection, and so on. Indeed, in 1978, Susman and Evered [5] developed the action research cycle to include the five stages of diagnosing, action planning, action taking, evaluating, and specifying learning, which are then iterated.

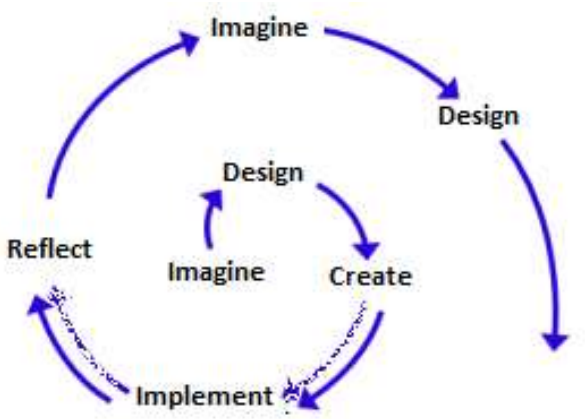


Figure 2. The iterative working process used to create and combine interdisciplinary knowledge in synthesis. There may be parallel processes studying implementation with different scientific methods (double arrows), and the knowledge is combined in synthesis of the subsequent cycle.

For the development of solutions in the mixed research method, performing such iterations is viewed as beneficial to both understanding the environment (users, use cases, practice, etc.) and to specify, develop, and refine the technological solution, since the iterative approach provides repeated observations to detect clear patterns. Further, it is noteworthy and important that users and environments change with the introduction of the technological solutions, as studied through the iterations.

In order to perform these iterations to build knowledge and to perform successful synthesis, rapid prototyping provides a method for ensuring grounding of the technology development in human interaction reality and grounding of the human interaction analyses in the technological reality.

Rapid prototyping is characterized by making iterations of prototypes in a fast manner, where the quick development of prototypes allows constant interaction and testing in the environment (e.g. with the users). This serves both as development of the system specifications themselves and as grounding in reality. Since the rapid prototyping approach recognizes that all requirements cannot be specified a priori, in contrast to the waterfall approach, the requirements are learned and satisfied incrementally. Hence, small improvements are constantly implemented and aligned with users/environment.

According to the ASTM Standard [6], in contrast to the waterfall approach of preparing "requirements and design documents that describe the needed system, rapid prototyping methods concentrate on preparing a working prototype. Users and developers learn the functional requirements and an appropriate system design by interacting with a series of prototypes, each of which is rapidly produced from a starting framework or from an earlier version. A prototype can evolve into an operational system, it can serve as an exact behavioral specification of an operational system, or it can be used to explore the feasibility of a new idea or design which can be incorporated in a larger system."

Such rapid prototyping is attractive to perform synthesis, not only because it creates the technological prototypes and advances, but also because it provides all participants (researchers, professionals, users) a common tool to meet around, share, and interact with, thus creating a common language and understanding, which is essential for the interdisciplinary research. The concrete hands-on experience with the prototypes in the environment grounds abstract concepts in concrete actions, which facilitates the creation of a common language and understanding amongst the participants. The common understanding is further refined fast in the rapid prototyping cycles.

This kind of cyclic action research method proposes flexibility in the research design to meet the real needs in the environment. Cunningham stated this in 1976 as: "The action research process makes it highly unlikely that the investigator will know exactly, or in advance, the design of the inquiry. Since every execution has to be evaluated and judged as to how effectively it meets the plan, revisions to fit new needs will be necessary. As hypotheses are validated or invalidated by the interim results, the problem may be redefined and the hypotheses and research methods modified." [7] p. 218.

V. DISCUSSION

Interestingly, the spiral of steps in action research as proposed originally by Lewin [4] seems widely to have influenced other spiral models, which have become known extensively in other engineering and scientific disciplines such as software development [8, 9], design, and constructivist education [10]. For instance, in the creative thinking spiral, people imagine what they want to do, create based on their ideas, play with their creations, share their ideas and creations with others, and reflect on their experiences—all of which leads them to imagine new ideas and new creations.

The iterative process needs to start somewhere. Susman and Evered [5] suggest that the cycle starts with the diagnosing phase. Similarly, the imagine phase provides the start in the creative thinking spiral [10]. This can often be a difficult and abstract phase for researchers and participants. It is our belief that the process can be kick-started in such a diagnosing or imagining phase by the introduction of technology suitable for diagnosing, imaging and creating prototypes. Such technology can ideally have a modular expression to allow any user to easily *construct, combine and create* prototypes from those modules. Based on a deep engineering and computer science knowledge on modular technology and embodied artificial intelligence, it is possible to develop technological modules that any user can easily understand and construct with within a minute [e.g. 2, 11, 12]. This provides an important hands-on involvement in the first phase, which makes it easier for participants to ground the diagnosing and imagination in both the world reality and the technological reality.

ACKNOWLEDGEMENT

The author would like to thank colleagues from the Center for Playware, DTU for experimental research, interdisciplinary collaboration, and methodological discussions which led to the formulation of the presented research method.

REFERENCES

[1] W. Royce. "Managing the Development of Large Software Systems", Proceedings of IEEE WESCON 26 (August): 1–9, 1970.
[2] H. H. Lund, T. Klitbo, and C. Jessen, C. "Playware Technology for Physically Activating Play", *Artificial Life and Robotics Journal, 9:4*, 165-174, 2005.
[3] H. H. Lund, and P. Marti. "Designing Modular Robotic Playware". In Proc. of 18th IEEE Int. Symposium on Robot and Human Interactive Communication (Ro-Man 2009), IEEE Press, 115-121, 2009.
[4] K. Lewin. "Action Research and Minority Problems," *Journal of Social Issues 2* (1946): 34-46.
[5] G. I. Susman, and R. D. Evered. "An Assessment of the Scientific Merits of Action Research", *Administrative Science Quarterly*, (23) 1978, pp. 582-603.
[6] ASTM International, "Standard Guide for Rapid Prototyping of Information Systems", E1340 − 05, PA, USA, 2010.
[7] B. Cunningham. "Action Research: Toward a Procedural Model," *Human Relations* (29:3), 1976, pp. 215-238.
[8] B. Boehm. "A Spiral Model of Software Development and Enhancement", ACM SIGSOFT Software Engineering Notes, ACM, 11(4):14-24, August 1986
[9] B. Boehm. "A Spiral Model of Software Development and Enhancement", *IEEE Computer*, IEEE, 21(5):61-72, May 1988
[10] M. Resnick. "Sowing the Seeds for a More Creative Society." *Learning and Leading with Technology*, December 2007
[11] H. H. Lund. "Modular Playware Technology – A Brief Historical Review". In Proceedings of 17th International Symposium on Artificial Life and Robotics, ISAROB, Japan, 2012.
[12] H. H. Lund. "Anybody, Anywhere, Anytime – Robotics with Social Impact through a Building Block Approach". In Proceedings of IEEE Advanced Robotics and Its Social Impact (ARSO 2011), IEEE, 2012

ALife for Real and Virtual Audio-Video Performances

Luigi Pagliarini[1,2] Henrik Hautop Lund[1]

[1] Centre for Playware, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[2] Academy of Fine Arts of Macerata, Via Berardi, 6, 405111 Macerata, Italy

luigi@playware.dtu.dk hhl@playware.dtu.dk
www.playware.dk

Abstract

MAG (i.e.: an Italian acronym which stands for Musical Genetic Algorithms) is an electronic art piece in which a multifaceted software attempts to "translate" musical expression into a corresponding static or animated graphical expressions. The mechanism at the base of such "translation" consists in a quite complex and articulated algorithm that, in short, is based on artificial learning. Indeed, MAG implements different learning techniques to allow artificial agents to learn about music flow by developing an adaptive behaviour. In our specific case, such a technique consists of a population of neural networks – one dimensional artificial agents that populate their two dimensional artificial world, and which are served by a simple input output control system – that can use both genetic and reinforcement learning algorithms to evolve appropriate behavioural answers to an impressively large shapes of inputs, through both a fitness formula based genetic pressure, and, eventually, a user-machine based feedbacks.

More closely, in the first version of MAG algorithm the agents' control system is a perceptron; the world of the agents is a two dimensional grid that changes its dimensions accordingly to the host-screen; the most important input artificial agents get (i.e. not necessarily the only one) is the musical wave that any given musical file produces, run-time; the output is the behavioural answer that agents produce by moving, and thereby drawing on to a computer screen, therefore graphical. The combination of artificial evolution and the flows of a repeated song or different musical tunes make it possible for the software to obtain a special relationship between sound waves and the aesthetics of consequent graphical results.

Further, we started to explore the concept of run-time creation of both music and graphical expression. Recently, we developed a software by which it is possible to allow any user to create new song versions of popular music with the MusicTiles app simply by connecting musical building blocks. This creation of musical expression can happen as a performance (i.e. run-time). When connecting the MusicTiles app to the MAG software, we provide the connection and the possibility to melt both musical expression and graphical expression in parallel and at run-time, and therefore creating a audio-video performance that is always unique.

Introduction

In this paper we present the "MAG" [1] an application, an ALife software based on and derived from our previous work on the ALife models and work.
MAG (i.e. Musical Genetic Algorithm, in Italian) is an electronic art piece that throughout an automatism translates sound expressions into graphical ones. The mechanism behind such "translation" is a quite complex algorithm that, synthetically, relies on artificial learning. Indeed, the MAG software implements a population of artificial agents – living in an artificial world and furnished with a simple sensory-motor system – that genetically evolves to coherently respond to different inputs, thanks to the most classical selective pressure [2].
In the MAG software the agents control system is a Perceptron [3], the simplest version of a Neural Network [4]; The world in which such perceptrons are placed is a two dimensional grid of adaptive dimensions and with a toroidal shape; The main input which the artificial organisms respond to is a sound wave; the artificial agents' answer to inputs is a behavioural one where the behavior is traced and graphically impressed onto the computer screen.
The combination of artificial evolution and the execution of a song or many different songs make it possible to achieve at any software execution a special and unique relationship between incoming sound waves and the consequent emerging graphical expression.
After having made all possible efforts to reach highly aesthetical results, and once obtained such a relationship

between music and graphics, we developed the concept of applying the MAG algorithm to our MusicTiles application [5, 6] - a software conceived to open up music remixes and composition to ordinary people and professional users. All of that, because our final intention is to research for a mechanism by which any kind of user could use a tool for creativity to live performing both music and graphics, sound and images.

Technology

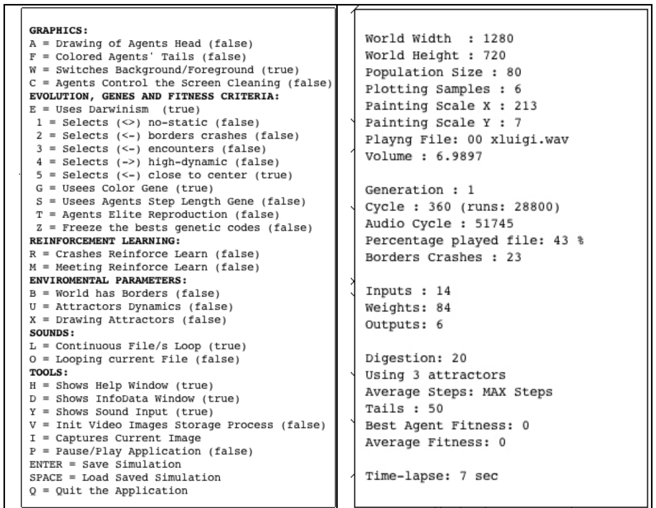
The underlying technologies of the MAG algorithm belong to the classic simulative logic with its roots in Artificial Life [7], and which we already used for artistic means [8, 9]. Within an artificial planet (i.e. a two dimension matrix, a chessboard with a toroidal spatial geometry) a population of artificial organisms - whose N shapes according to the screen size – starts evolving and by doing so draws on an “electronic canvas” a continuously-transforming image.

Artificial organisms, that are evolved thanks to a Genetic Algorithm [10], occupy a Cartesian point of such world (i.e. are mono-dimensional) and their central nervous system consists of a Perceptron that allows to establish a specific relationship between a sensory input and a motor output.

The main input we chose to stimulate such perceptrons is a given portion (i.e. of parameterized dimensions) of a sound wave, extrapolated run-time from any file or direct sound running on the sound board. To such input, that numerically speaking is the main source of information to all the Perceptrons, the user could easily add a different input, such as the organism’s distance from one or more landmarks. Such landmarks - which are positioned into the world either randomly or logically (i.e. at user choice) - create a referring point(s) around which the organism can navigate or maybe, as in our specific case, “dance”.

To these, that under all circumstances can be considered real inputs, we added different simulation options and tunable algorithmic parameters (see Figure 1) that enrich the behavioral scenario and, consequently, the graphic one, transforming each single instance of the simulation.

Indeed, as we can easily imagine observing the Control Panel, shown in Figure 1*A*, there are quite many parameters that contribute to define the specificity and evolutionary characteristics of the running simulation. Some of them act on the agent’s general behavior - for example the static or dynamics of the landmarks – and others influence the aesthetical result – for example the phenotypic expressions of the agents as the tail or the color can be.



A. Control Panel *B*. Data Panel

Figure 1. The Control and Data Panels

All the variables shown in the Control Panel can be changed run-time, meaning that the user can shape each single simulation, making it more or less unrepeatable. Indeed, opening up such a set of variables to users’ intervention corresponds to acting on the computed fitness gradient - since the fitness is the index by which the adaptation coefficient of an artificial agent is calculated. By modulating such computation we act on the reproductive power of each single organism and, therefore, we can change the effective history of the running simulation. Each simulation can be monitored through a Data Panel – see Figure 1*B* for an example.

Music is at the center of the MAG software, that’s is why each simulation starts as soon as an audio stream appears (e.g. at the beginning of a song) and produces a *generational transition* as soon as silence interrupts it (e.g. at the end of the song). By *generational transition* we mean that the ordinary GA’s process of best individual’s fitness selection and reproduction. In our algorithm, the reproduction routine applies both mutations and cross-over. All of that is repeated for each song (i.e. a high significant discontinuity in the amplitude of the streaming audio, or silence) and such process is reiterated up to when the user consider it sufficient, as an aesthetical result. Indeed, as the simulative process goes on, one can observe and more and more well defined behavioral expression of the artificial agents where, most of the times, agents learn to “dance” on the sound wave that goes across themselves.

As shown in the Control Panel (Figure 1*A*), besides the inter-generational selective pressure, it is made possible to apply to agent behavior two different forces of

Reinforcement Learning [12], since we were seeking for a force able to change the intra-generational behavior of our perceptrons.
The two Reinforcement Learning routines we implemented:

A. *Crash.* We apply the RL each time an agent crashes against the world borders (i.e. we might reward agents that stand in the center of the “dance-floor”);
B. Meeting. We apply the RL each time an agent crashes against another agent (i.e. we might reward agents that do “not smash other agents feet”);

Both RLs are used and seem to work despite of the fact that agents are not always given the possibility of sensing the world contours or the mates’ presence. This is also because the two behaviors described above start to be computed in to the Fitness Formula as soon as they are selected as operative.
In synthesis, to summarize the whole simulation flow:

1. As soon the MAG is started, the users are allowed to choose amongst few basic and starting parameters (e.g. number of landmarks);
2. A population of N Perceptrons (according to screen size) is initiated;
3. As soon as the sound board is activated, agents population is started;
4. Agents, one after the other, receive as input a portion of run-time executed sound wave (i.e. plotting samples, which is itself a variable)
5. For technical reasons, and since the sound process is always faster than simulative computation, each single agent is fed with a different portion of sound and, therefore, has a specific and unique input;
6. Agents calculate their behavioral reaction to sound and, eventually, other factors (e.g. distance to landmarks);
7. Agents calculates their potential movement in the Cartesian space and, therefore the fitness of such (RL takes place here, eventually);
8. All agents draws their movement;
9. The simulators opens up to any user interaction, either regarding functional parameters (i.e. Genetics, Fitness Formula, Reinforcement Learning, Environmental Parameters, Sound Input Choice and Analysis) and non-functional ones (i.e. Graphical Output, Feedback Tools, Run-Time Data Description, Simulation Controls, etc.);
10. As soon as the sound board is deactivated (i.e. end of files or silence) agents population is submitted to a generational transition, which means selection, reproduction and mutation;

This process is re-iterated at the users’ will.

Aesthetical Outcome

The simulation leader decides if, and to what extends, to control the graphical outcome of our population of agents, leave that decision to emerge from specific expressions of agents’ artificial genes or, since the whole process can be modified run-time, both. For the graphical outcome we intend both the colour and the shapes of the drawings. The MAG genes act in such a way that part of the phenotypic expressions of the output are controlled by the single agent, while other ones (e.g. the background color) might be controlled by all agents together. For example, a single agent’s gene controls the length of its stepping, while in each cycle the algorithm weights up the outputs of the whole population and determines the background color (i.e. when a specific option has been activated by the users). Therefore, the chromatic basis of the agents’ “drawings” can either be determined by the user or rely on a genetic mechanism. The resulting aesthetical outcomes perform a quite impressive range of graphical solution and, moreover, is capable of big variations even within one single simulation run (see Figure 2 for few examples).

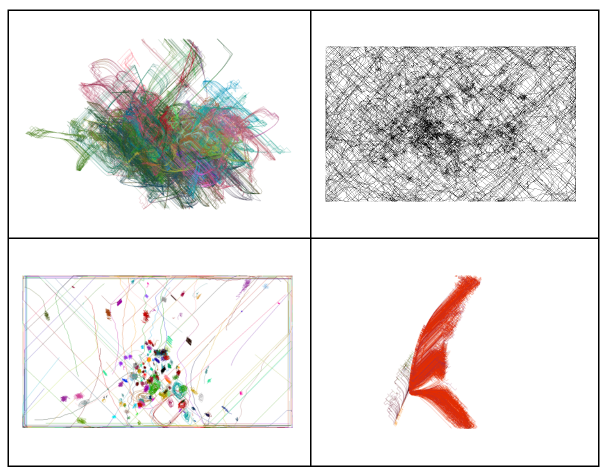


Figure 2. Some still images from different evolutions of different songs.

Aesthetics is (or can be) modified across the time and changes accordingly to the interactive process enacted the users with genetic components, landmarks, reinforcement learning and the other variables but, overall, it is influenced and transfigurated by the choice of different music flows, that stands for the user decision of the sound input which can be determined at any given moment along the evolutionary paths initiated by the population of artificial agents.

The MAG software has already found its artistic positioning under the form of a preassembled video (Gabriele 2012) whose projection rights have been sold to different galleries and festivals, all over the world. Despite of that, the software itself is under constant technical evolution, and in the early future we will release an advanced version where the agents population will:

1. Incorporate three layers Neural Networks, instead of the actual Perceptrons (i.e. to reach much more complex and articulated movements);
2. From simple input-output organisms agents' brain will be furnished with a neural memory mechanism (i.e. to better follow the bit);
3. Each single population will be split into different "races", or different subsets of populations (i.e. to push agents specialization on different frequencies of sound analysis).

MusicTiles Interface

The MusicTiles [5] (former MagicTiles [6]) music app, derived from the RoboMusic concept [13] and that can be used by users of different musical competencies - from precipitants to musician to expert composers - to both remix songs or experiment with new tunes.

The MusicTiles app that runs on iPads and iPhones, is made of a small number (i.e. more than 3 less than 10) of virtual music tiles. Each single tile embodies a single instrument and each instrument-tile is capable of expressing seven different musical solution (i.e. loops) or variation of the same one. Every instrument of a MusicTiles song is wrapped inside a single tile and, therefore, we can have a tile for vocals, one for the rhythm guitar, one for the bass, one for the piano, etc.

At the start, the user is presented with all the tile-instruments of a given song inside the tiles board with the tile-instruments being disconnected from each other (see Fig. 3).

Figure 3. MagicTiles for iPad.

Since a tile is disconnected from other tiles, it is not active and will not play. Once the user drags and attaches a tile to another (see Fig. 5 left) the two tiles get synchronized and are able to play together.

To actually start the music, the user has to press the play/stop icon, positioned at the bottom-right side of the screen (see Fig. 3). At this point, the music will keep playing and the remaining music tiles can be connected to the former two enriching the music scenario of the song running. In the same way, by dragging the tiles around and rearranging their geometry, the user changes the music sequence and flow.

Figure 4. Left: Connecting two tiles. Right: 2-sides connection example.

Every tile has 7 music loops, one for each side (north side, east side, south side and west side) plus 3 extra music loops which will play when a tile is connected on, respectively, two, three or four sides. In this way, as in Fig. 5 left, the Drum tile will play the East Side loop and the Guitar tile will play the West Side loop, and as in Fig. 5 right, the Drum tile will play the *2 Sides* loop, the Guitar tile will play the West Side loop and the Choir tile will play the North Side loop.

Likewise, as in Fig. 5 the Drum tile will play *3 Sides* loop, the Choir tile the North Side loop, the Guitar tile will play the West Side loop, and the Bass tile the East one.

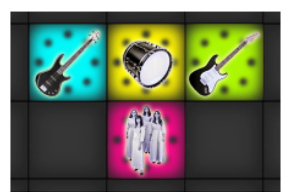


Figure 5. A *3 Sides* connection example.

Growing the number of instruments/tiles impressively increase the number of possible combinations and out of these simple rules, one obtains an extraordinary number of musical possibilities.

MAG and MusicTiles App Integration

Besides all of that, we are at the beginning of a process of integration of MAG software with MusicTiles music app to explore the concept of run-time composition of both music and graphical expression. The idea is to allow any user to create musical expression through the MusicTiles (i.e. which can happen run-time as in a performance run-time) and, when connecting the MusicTiles app to the MAG software, providing he/she the possibility to melt both musical expression and graphical expression in parallel therefore creating a audio-video performance that is always unique.

Discussion and Conclusion

In this paper, we presented both a new application – MAG, which aims to create graphical expression by following musical paths - and a previous one - MusicTiles, which aims at exploring a new approach to music for run-time composition and/or remixing. Through such applications any user could create and perform both Graphics and Music, on one side, by manipulating simulation parameters and, on the other, by assembling a number of interactive virtual tiles on the screen. After our past efforts in building physical interfaces, we decided to conceive a cross-media platform to head to cross-modal and cross-sensory artistic result. Further evolution of such an approach might lead to a combination of both virtual interface (i.e. MAG and MusicTiles) with a physical interface. In short, it seems that we instantiated a research process in which we are trying to combine different input and output methodologies, systems and tools that might lead to a broader vision of software and robotic systems with an articulated, fluid and bidirectional flow between the physical and virtual environment. Indeed, we believe that such a fluid and multifaceted representation of a single activity may widely enhance the user immersion into a *one reality* that combines multi-sensory activation on either a physical and a virtual level.

References

[1] L. Pagliarini. *Algoritmi Genetici Musicali (MAG).* In Rivista di Psicologia dell'Arte, Nuova Serie, Anno XXXIII, n° 23, Dicembre 2012. Pag. 103-108.

[2] D. E. Goldberg. *Genetic Algorithm in search, optimization and machine learning.* New York, Addison-Wesley. 1998.

[3] F. Rosenblatt. The Perceptron--a perceiving and recognizing automaton. Report 85-460-1, Cornell Aeronautical Laboratory, 1957.

[4] D. E. Rumelhart, J. L. McClelland. *Parallel Distributed Processing. Explorations in the Microstructure of Cognition.* MIT Press, Cambridge, MA. 1986.

[5] Peter Gabriel's Latest App is So 2.0, Fast Company, 7 January 2013: http://www.fastcompany.com/3004537/peter-gabriels-latest-app-musictiles-so-20

[6] L. Pagliarini and H.H. Lund: MagicTiles. ALife for real and virtual Robot Music. Proceedings of the 17th International Symposium on Artificial Life and Robotics (AROB'12), ISAROB, Japan, 19-21 Jan, 2012.

[7] C. G. Langton. *Artificial Life*. In C. G. Langton, editor. *Artificial Life*, Volume VI of *SFI Studies in the Sciences of Complexity*, pages 1-47, Addison-Wesley, Redwood City, CA, 1989.

[8] A. Spina, L. Pagliarini, Globalization: tra arte scienza e società. In *Rivista di Psicologia dell'Arte* Anno XII, n.12, 87-92. Roma, Dec. 2001.

[9] L. Pagliarini. Una Tela Viva. Rivista di Psicologia dell'Arte, Nuova Serie Anno XXIII, n°13, Dicembre 2002. Pag. 33-36

[10] M. Mitchell. *An introduction to genetic algorithms*. MIT Press. 1997.

[11] A. Gabriele, Ogni Dove Edizioni, 2012 http://www.ognidove.eu/ode2/.

[12] R. S. Sutton, Temporal Credit Assignment in Reinforcement Learning, Ph.D. thesis, University of Massachusetts, Amherst, MA, 1984

[13] H. H. Lund, and M. Ottesen. RoboMusic – A Behavior-Based Approach, *Artificial Life and Robotics Journal, 12: 1-2*, pp. 18-23, 2008.

Heart-pulse Biofeedback in Playful Exercise using a Wearable device and Modular Interactive Tiles

Tomoya Shimokakimoto[1], Henrik H. Lund[2] and Kenji Suzuki[1,3]

[1]Artificial Intelligence Laboratory, University of Tsukuba, Japan
[2]Center for Playware, Technical University of Denmark, Denmark
[3]Japan Science and Technology Agency, Japan

shimokakimoto@ai.iit.tsukuba.ac.jp, hhl@playware.dtu.dk, kenji@ieee.org

Abstract: We developed a playful biofeedback system using a wearable device and modular interactive tiles. In this approach we suppose that patients could regulate exercise intensity on their own through biofeedback. We propose biofeedback play system called “bioToys” based on exercise with the modular interactive tiles. The system consists of a wearable device which measures heart-pulse via ear-mounted sensor, and modular interactive tiles which are used for physical rehabilitation exercise through playing a game. The wearable devise enables detection of heart pulse in real-time and therefore provides heart beat rate during playful activities, even if the heart pulse wave have motion artifacts. The tiles are designed to build flexible structures and to provide immediate feedback based on the users’ physical interaction with the tiles. We combine the two systems to provide users with heart pulse biofeedback in playful exercise. We show that using the developed system it is possible for the users to regulate the exercise intensity on their own with biofeedback, and also possible to analyze exercise activity using number of steps on the tiles and heart beat rate.

Keywords: **playware, biofeedback, exercise, rehabilitation**

I. INTRODUCTION

In order to improve the health and quality of life (QOL) of elderly people and patients with metabolic disorders, stroke and cardiac diseases, physical training and exercise is required in daily life. Heart Beat Rate (HBR) is usually used as a standard for regulating the exercise intensity [1]. Thus, there are some commercial wearable devices of measuring HBR. And some exercise systems that control the exercise intensity based HBR are worth consideration [2][3].

For practical implementation of wearable sensing technologies in physical exercise, there is a need for methods to simplify measurement of bio-signals, as well as signal processing techniques to reduce motion artifacts in order to provide stable biofeedback. Systems such as those in [4][5] try to take those considerations into account.

In physical exercise, it is not easy to maintain participants’ motivation for joining a sport [6]. Recent research also utilized games as means for exercise, showing a number of cases where gaming may be beneficial for health.

In this paper, we propose the hypothesis that patients could regulate exercise intensity on their own through biofeedback during playful activity, and we propose a biofeedback play system called “bioToys” based on exercise with a wearable device and the modular interactive tiles.

II. SYSTEM CONFIGURATION

1. Wearable Device

We developed a wearable device to measure heart pulse and generate biofeedback in real-time. The wearable device consists of an ear-mounted sensor (ear sensor, Combi Wellness Corp.), a processor (LPC1752FBD80, NXP), a light emitting diode for biofeedback and Bluetooth module for communication with PC. Fig. 1 shows the overview of the wearable device for heart pulse measurement and biofeedback.

Heart pulse measurement via ear-mounted sensor is prone to noise resulting from body movements, thus it is important to process the heart pulse waveform in real-time to generate a stable biofeedback output. In our device we implemented an algorithm based on linear prediction and an adaptive template matching method to overcome noise in the waveform [5]. The algorithm is capable of detecting actual edges in the heart pulse waveform in real-time, even if the pulse waveform include motion artifacts similar to heart pulse. This method utilizes human physiological features to generate a correct cyclic pulse from noisy signals. This algorithm is mainly constructed by two equations (see equation 1 and 2). The equation 1 represents a differential equation of linear prediction and the equation 2 means a template matching.

In equation 2, $P(t)$ is the pulse waveform vector measured by the A/D converter with the mean value normalized to zero. When P includes a cyclic motion artifact such as those resulting from locomotion or jumping, a mere thresholding method is not adequate to distinguish such artifacts from heart pulse. The developed algorithm, by linear prediction, avoids detecting artifacts similar to heart pulse. In the linear prediction method we designed a Kalman filter according to equation 1, which is a differential equation model for calculating an estimate time interval of the heart pulse $\widehat{\tau_c}(m)$[s] depending on previous detected time intervals. This linear prediction method is capable of approximately predicting the next heart pulse interval according to the pattern of the heart pulse[7]. Therefore, an edge of heart pulse waveform is detected after $\widehat{\tau_c}(m) + \delta_t$ [s], which is the sum of estimated time and difference between the previous estimated time and actual detected pulse interval.

In addition, for reducing noise effect around an edge of heart pulse waveform, this algorithm uses ideal pulse waveform vectors for heart rates (τ_s[bpm]) that don't include noise: $I(t;\tau_s)$. Then, using the correlation value between I and P; $M(\tau_s)$ in equation 2, when the value of $M(\tau)$ is higher than a threshold, the system detects an edge of heart pulse waveform.

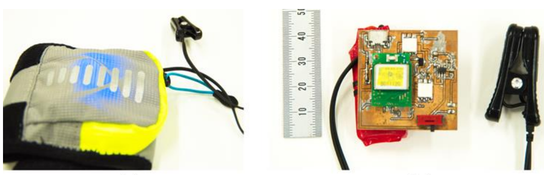
Using the linear prediction and template matching methods, this wearable device is able to detect edges of heart pulse and calculate HBR stably. LEDs on the device blink in synchronization with the detected heart pulse, and the HBR data is transmitted to PC via Bluetooth in real-time.

$$\widehat{\tau_c}(m+1) = \widehat{\tau_c}(m) + \gamma(\widehat{\tau_c}(m) - \widehat{\tau_c}(m-1)) \quad (0 < \gamma \leq 1) \tag{1}$$

$$M(\tau_s) = \frac{1}{N}\sum_{i=0}^{N-1} I(t-n\tau)P(t-n\tau) \tag{2}$$

2. Modular Interactive Tiles

We developed a system for playful biofeedback based on modular interactive tiles. The tiles attach to each other to form the overall system, and they are designed to be flexible for different arrangements. The tiles are also motivating to the user since they provide immediate feedback based on the users' physical interaction with the tiles, which is a design principle of modular playware [8]. Fig. 2 shows a scene of playing with the tiles.



(a) (b)

Fig.1. (a) Overview of the wearable device in a sports wrist pocket (b)the circuit board and heart pulse sensor

Each tile comprises a sensor and transmitter for communicating with other tiles, has a quadratic shape measuring 300mm*300mm*33mm, a force sensitive resistor (FSR) and eight RGB LEDs that emit light with equal spacing in between on a circle shape for game feedback.

On the circuit board of each tile there are also connectors to mount an XBee radio communication add-on board, including the MaxStream XBee radio communication chip. Hence, there are two types of tiles, those with a radio communication chip (master tiles) and those without one (slave tiles). The master tile may communicate with the game selector box and initiate the games on the built platform. Every platform has to have at least one master tile if communication is needed to game selector box or a PC. The tile becomes a master tile only when it has XBee radio communication connected to it.

The modular interactive tiles can be easily set up on floor. They can be simply attached to each other as a jigsaw puzzle, and there are no wires. The modular interactive tiles can register whether they are placed horizontally or vertically, and they can adapt the software games to behave accordingly.

3. Real-time biofeedback for exercise system

The games on modular tiles are designed for physical exercise. Basically when users touch or step on the tiles, the tiles generate feedback through LEDs' color and intensity according to the game's rules that are installed in tiles. We developed a playful biofeedback system by combining a wearable device for measurement of heart pulse rate with the modular

Fig.2. Modular interactive tiles used for playful physiotherapy with feet

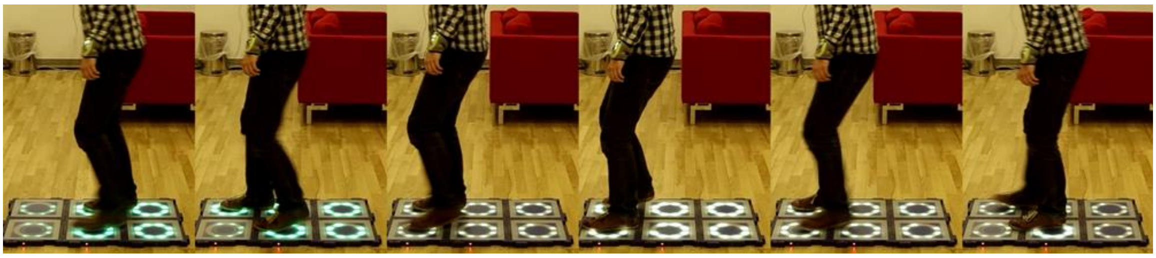


Fig.4. The scene in the playing using a wearable device and modular interactive tiles

interactive tiles. In this paper, we modified the games installed in modular interactive tiles [9] for some new purposes as will be illustrated.

The wearable device transmits the calculated HBR to PC via Bluetooth as the wearable device detects users' heart pulse. After the PC receives data of HBR, data is logged in the PC and transmitted to the master tile wirelessly via XBee. When the master tile receives the heart beat data, it interprets it and light the LEDs according to the game rules. This system has six colors for biofeedback. Fig. 3 shows an example of color assignment based on HBR in reference to ratings of perceived exertion scale, which is a scale of perceived physical exertion relating to HBR [10].

For biofeedback, HBR is assigned by color and the game level or score point. Modular interactive tiles have games based on counting steps or touches, or based on time constraints. For example, there are "color race game" and "final countdown". Color race is based on counting steps. When the user steps on an activated or illuminated tile, user gets score. Finally when the user acquires an adequate score the game finishes successfully. By adding the wearable heart beat device, the user is able to know his heart pulse and HBR and share it with audience.

Final countdown game is based on time constraint. In this game, all tiles are activated and the eight LEDs on each tile count down independently. When any tile finishes to count down it will send a "game over" message to the master tile and the game will finish. When the user steps on a tile, count down time in this tile will reset and the eight LEDs on this tile are lit again. For this game, in normal case, count down speed parameter is fixed depending on game difficulty. For regulating the exercise, we designed a game where the user keeps playing as long as proper HBR is observed. In this case, we designed the countdown parameter to

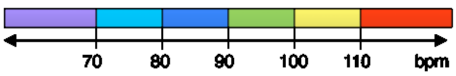


Fig.3. Example of a color assignment based on HBR

change depending on the HBR. At the first level which is under 70 bpm (color is purple), the counts down interval is 2s. Then for the following levels (light blue, blue, green) the countdown time decreases 0.32s respectively. Considering the fourth level (green, 100bpm) as the desired level, count down timer will increase for later levels (yellow and red) also by 0.32s respectively, in order to keep the user at the desired level.

III. EXPERIMENT

1. Regulating exercise

In order to evaluate the performance of the proposed system, we conducted an evaluation experiment to verify the feasibility of the proposed system as a playful biofeedback system for regulating exercise. In this experiment, we used six tiles and final countdown game. We asked four healthy male participants to exercise for two minutes with the proposed system and to go up to the third level (green color level) and keep this level. We logged heart beat pulse by 120Hz, HBR, the number of steps and also recorded these experiments on video for analysis.

2. Result and analysis method

Fig. 4 shows a scene during the experiment. We confirmed that the proposed system provide color feedback dependent on HBR during the game. Fig. 5 shows all participants' results of this experiment. Here

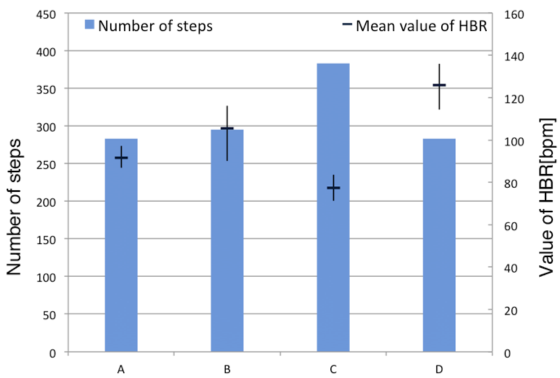


Fig.5. The result of all participants

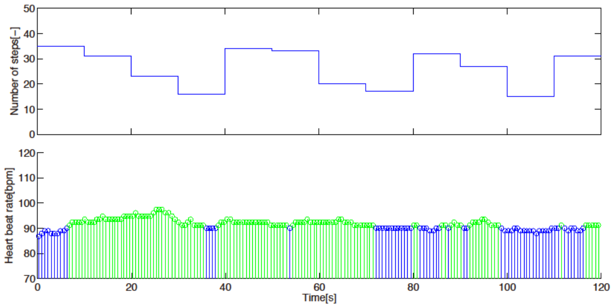


Fig.6. The result of participant A in two minutes exercise

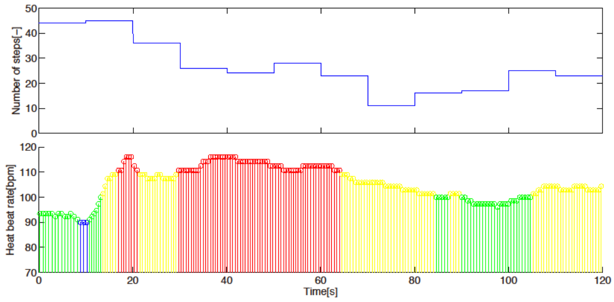


Fig.7. The result of participant B in two minutes exercise

we compare the number of steps and HBR among participants. In this experiment, the results of participant A and B were similar. Participant C (who exercises on regular bases) did not get tired during the experiment. On the other hand, participant D who does not usually exercise was tired from the point of HBR.

Then we focused on the time series of steps and HBR. Fig. 6 and 7 show the results of two participants (A and B). The top graph shows the events of detected heart pulse, the HBR is expressed in the color of biofeedback. The bottom graph shows the sum value of number of steps for each 10 seconds.

In Fig.6, when participant A reached green level, he intentionally decreased his pace to the desired HBR. In participant B's case, first he overshoot to the red level for around 20s, then he decreased his pace to keep the desired HBR. Finally, he was able to find the proper pace to stay between green and yellow levels.

The proposed system showed the possibility for the users to regulate the exercise intensity during playful activity on their own with biofeedback.

IV. CONCLUSION

In this paper we proposed a biofeedback play system called "bioToys" based on exercise with modular interactive tiles and a wearable device. We consider that games as a mean for exercise are beneficial for health, and for maintaining the motivation for exercise.

We developed a wearable device which measures heart pulse via ear-mounted sensor, and we implemented an algorithm for biofeedback in the real time which detects heart pulse in intense activity. We also developed a system for playful feedback game based on modular interactive tiles. Then we combined the two systems and developed games to provide users with heart pulse biofeedback in playful exercise.

We devised an evaluation experiment to verify the feasibility of the proposed system for regulating exercise. We showed that it is possible to use the developed system to regulate the exercise intensity on one's own with biofeedback, and it is also possible to analyze exercise activity using number of steps on the tiles and HBR.

REFERENCES

[1] Karvonen, J., and Vuorimaa, T., (1988), Heart Rate and Exercise Intensity During Sports Activities., *Sports Medicine*, **5**(5): 303–311.
[2] Su, S., W., and Wang, L., and Celler, B., G., and Savkin, A., V., (2005), Heart rate control during treadmill exercise. *Proc. of Annual Intl. Conf. of the IEEE EMBS*, pp. 2471–2474.
[3] ElSamahy, E. and Genedy, A. and Abbass, M.A. and Gaddallah, M. (2011), A Computer–based System for Safe Physical Fitness Evaluation, *Proc. of Annual Intl. Conf. of the IEEE BMEI, 2011*, pp. 1443–1447.
[4] Watanabe, M., and Tsukamoto, S., and Hoshino, H., (2008), A handle electrode system for measuring heart rate while riding a bicycle, *Proc. of Annual Intl. Conf. of the IEEE BMEI*, 2008, pp. 542–545.
[5] Shimokakimoto, T., and Ayuzawa, S., and Suzuki, K., (2013), Real-time Pulse Detection for Physiothrapy and its Application toWearable Device, *IPSJ Journal*, **54**(4): 1480–1488 (in Japanese)
[6] Dishman, R.K., (1988), Exercise adherence: Its impact on public health, *Human Kinetics Publishers*.
[7] Kobayashi, M., and Musha, T., (1982), 1/f Fluctuation of Heartbeat Period, *IEEE transactions on biomedical engineering*, **29**(6): 456–457.
[8] Lund., H., H., and Marti, P., (2009), Designing modular robotic playware, *the IEEE Int. Workshop Robots Human Interactive Commun Toyama*, Japan. Sep. 27-Oct. 2., IEEE Press, pp. 115–121.
[9] Lund., H., H., (2010), Modular Interactive Tiles for Rehabilitation: Evidence and Effect, *Proc. of the 10th WSEAS Intl. Conf. on ACS*, 2010, pp. 520–525.
[10] Borg, G., (1973), Perceived exertion: a note o n "history" and methods., *Med. Sci. Sports.*, **5**(2): 9 0–93.

Development of Measurement System for Quantitative Evaluation of Skillfulness of Lower Extremities

Kazunori Yamazaki*, Donggun Kim*, Yoshifumi Morita*, Noritaka Sato*, Hiroyuki Ukai*, Kenji Kozakai**, Satoru Shibata**, Shigenori Onishi**, Akihiro Ito**, Daisuke Mizuno**, Tatsuya Hirai***, Haruna Takeda***, and Yuka Sugiura***

**Dept. of Computer Science and Engineering, Nagoya Institute of Technology, Japan*
*** Sanyo Machine Works, Ltd., Japan*
**** Tanakakai Nishio Hospital, Japan*
Tel : +81-52-735-5412; Fax : +81-52-735-5412
20517513@stn.nitech.ac.jp

Abstract: The final goal of our research is to develop a measurement system for quantitative evaluation of the skillfulness of lower extremities. In our previous work, we developed the measurement system of the foot movement on the floor with the sitting posture. The subject moves the training device attached to the subject's foot on the training board. The pointer corresponding to the position of the training device is displayed on the monitor. The subject moves the pointer along the target trajectory by moving the training device. In this study, we developed a quantitative evaluation method of the skillfulness of the lower extremities. For this purpose, we used the measurement system which we have developed in our previous work and redesigned the target trajectories of foot movement. Moreover, we verified the reliability of the proposed quantitative evaluation method by calculating the intra-rater reliability for two trials *ICC*(1,1) of the evaluation results. We developed a measurement system comprising the straight course, the circular course and the star-shaped course. Because the *ICC*(1,1) of the circular course and the star-shaped course are higher than 0.7, the quantitative evaluation method of lower extremities with the circular course and the star-shaped course has sufficient reliability.

Keywords: Rehabilitation system, Quantitative evaluation, Intra-rater reliability, Lower extremities, Skillfulness

I. INTRODUCTION

Sensorimotor integration function and proprioceptive control in the lower extremities are necessary for walking ability and skillful movement. The elderly and patients sensorimotor integration function and proprioceptive control has become impaired. However, evaluation of therapeutic effect is only done subjectively by experienced doctors or therapists. Quantitative evaluation of therapeutic effect has not, been established until now. Therefore, the quantitative evaluation method of therapeutic effect is desired strongly in clinical sites of the rehabilitation. Moreover, it is absolutely imperative that the effective training method for improvement of walking ability and skillful movement.

Thus, the final goal of our research is to develop a measurement system and robot system for quantitative evaluation of the skillfulness of lower extremities. It has been desired in clinical sites of the rehabilitation. The skillfulness of the lower extremities is the ability to adjust the posture and the motion of the body to achieve his/her desired action.

We think that the skillfulness can be evaluated by the accuracy and the speed of the motion when his/her foot is moved to follow the given target trajectory. In this study, we developed a measurement system comprising three courses that have different difficulty levels about the target trajectory following in order to evaluate the skillfulness of the lower extremities.

On the other hand, the quantitative evaluation results should have high reliability. This reliability is Intra-rater reliability and Inter-rater reliability [1]. The intra-rater reliability is the degree of agreement among multiple repetitions of a diagnostic test performed by a single rater. Moreover, the inter-rater reliability is the degree of agreement among raters. The intra-class correlation (*ICC*) has been widely used to evaluate the reliability in rehabilitation medicine.

In this study, we developed a quantitative evaluation method of the skillfulness of the lower extremities. For this purpose, we used the measurement system which we have developed in our previous work and redesigned the target trajectories of foot movement. Moreover, we verified the reliability of the proposed quantitative evaluation method by calculating the intra-rater reliability for two trials *ICC*(1,1) of the evaluation results.

II. Measurement system for quantitative evaluation of skillfulness of lower extremities

1. Hardware and operation procedure

We have developed a rehabilitation support system as shown in Fig.1. The hardware consists of a laptop PC, a training device and a training board [2]. A PC wireless mouse is installed in the training device. The mouse is used as a low-cost position sensor. The pointer corresponding to the position of the training device is displayed on the monitor. And, we designed the system so that the motion of the pointer on the monitor and the motion of the foot on the training board might become the same size.

In the training, the subject moves the pointer displayed on the monitor by moving the training device in the sitting posture. During this the training, the subject watches the monitor of the laptop PC and does not watch their lower limbs directly. The subject can perform the training by playing the video game displayed on the laptop PC monitor. The therapist can evaluate the motion by using the acquired data during the training.

There are few risks of patient's overturning because this system is performed in the sitting posture. Moreover, there are no dangers by the runaway training device because the system does not have any actuators. Therefore, this system can be introduced at the early stage of the rehabilitation because of high safety of this system.

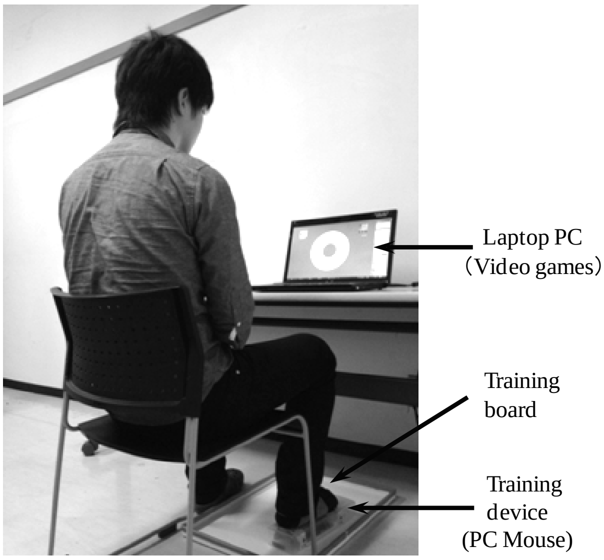


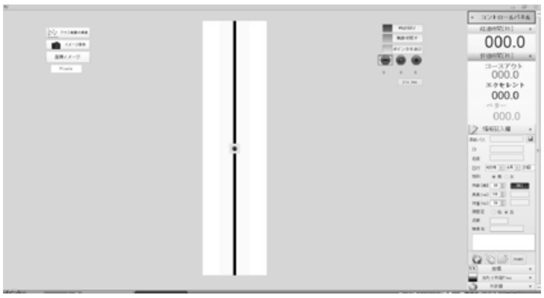
Fig.1. Measurement system for quantitative evaluation of skillfulness of lower extremities

2. Software

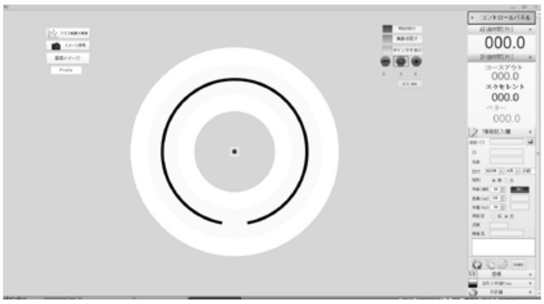
We developed three courses as the target trajectories; the straight course, the circular course and the star-shaped course as shown in Fig.2. The parameters of the courses were designed so that the difficulties of the three courses were different. The description of the software is shown in Fig.3.

The subject moves the pointer displayed on the monitor quickly and accurately along the target trajectory. The evaluation results of the skillfulness of lower extremities are expressed as the speed and the accuracy. The speed t [s] is given by the travelling time at each course. The accuracy S [pixel] is calculated by accumulating the error between the pointer position and the target position in each sampling time at each course.

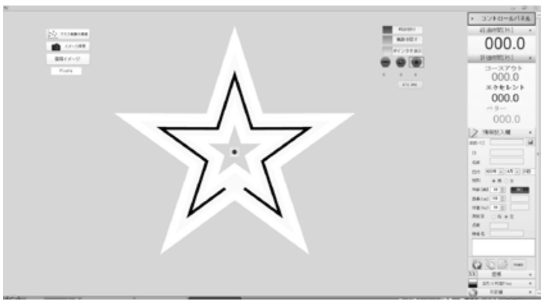
The extension of the knee joint is mainly used, in order to follow the target trajectory in the straight course. On the other hand, in the circular and star-



(a) Straight course



(b) Circular course



(c) Star course

Fig.2. Three courses for evaluation

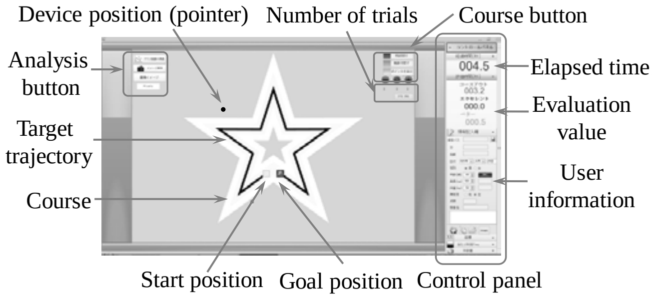
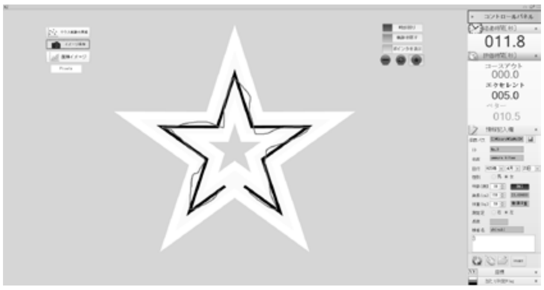
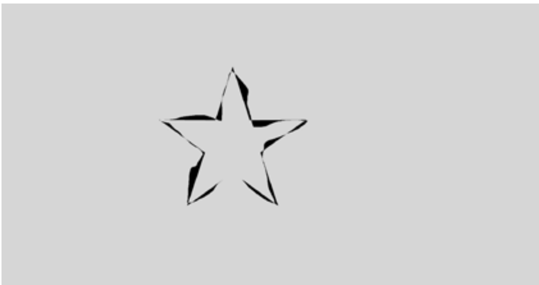


Fig.3. Explanatory drawing of courses



(a) Pointer trajectory and target trajectory



(b) Area which is surrounded by subject's trajectory and the target trajectory

Fig. 4. Evaluation results

shaped course, the complex motion which is mixture of the extension of the knee joint and the internal the rotation of the hip joint is necessary. In addition, in the star-shaped course, more time of switching of the direction of the motion is necessary than in the circular course. Therefore, each course is designed to be different in the difficulty.

When the software starts, it will display a control panel on the monitor. In the control panel, the therapist inputs the subject's information (user name, ID, height, weight, BMI, dominant foot, remark) and the location on HDD to save the data as a CSV file, at first. If not input, only measurement data is saved to the desktop of the PC. A time, a X & a Y coordinate of the pointer is measured and saved at each sampling. The sampling frequency is 100 [Hz]. After inputting the subject's information and the location to save data, the therapist selects the course. When the trial, one course running, is finished, if the course is desired to run continuously, the system counts the time of trials and displays it.

In order to check the results during/after the trials, the system shows the spent time from the start on each trial and the time which is counted up when the pointer is out of the course, as the feedback to the subject. There information can be eliminated from the control panel, if the therapist wants. In addition, the system can display subject's trajectory overlaid on the target trajectory after the training as shown in Fig. 4(a). And, the system can also display the area which is surrounded by subject's trajectory and the target trajectory as shown in Fig. 4(b).

III. Verified the intra-rater reliability

1. Method

In order to verify the inter-rater reliability of the measurement system we perform experiments. The subjects are fourteen healthy young persons. The age was 22±2.4 years old. The height was 166±8.3 cm. The dominant foot was right.

The subject was seated so that the distance between the subject's eyes and the PC monitor was 0.7 [m]. The training device was attached to the subject's left foot. The height of the seat surface was adjusted so that the knee joint angle was 90 [deg] in the sitting position. The subjects performed the tasks consisting of three courses. The subjects were instructed to move the left foot along the black line (the desired trajectory) quickly and precisely as possible, not to move the body, and to move the left foot smoothly on the training board. The subjects perform the tasks two times for each task. After finishing each task, the result was not informed to the subject.

We used the data of three courses and two times for each person. The $ICC(1,1)$ of the speed t and the accuracy S were calculated by the following equation [1].

$$ICC(1,1) = \frac{BMS - WMS}{BMS + (k-1)WMS} \tag{1}$$

$$BMS = \frac{\sum_{i=1}^{N} k(\bar{x}_i - \bar{x})^2}{N-1}, \quad WMS = \frac{\sum_{i=1}^{N}\sum_{j=1}^{k}(x_{ij} - \bar{x}_i)^2}{N(k-1)}$$

where *BMS* is the variance between the subjects, *WMS* is the variance in each subject, k is the number of measurements, N is the number of subjects, $\bar{x}$ is the mean value of data of N subjects and k times, $\bar{x}_i$ is the mean value of data of k times of Subject i, x_{ij} is the data of Subject i and j time. In this study, k = 2 and N = 14.

The criterion methods of *ICC*(1,1) were reported [3][4][6]. In Ref. [3], it was reported that 0.00-0.20: Slight, 0.21-0.40: Fair, 0.41-0.60: Moderate, 0.61-0.80: Substantial, and 0.81-1.00: Almost perfect. On the other hand, in Ref. [4], it was reported that more than 0.6: Possible, more than 0.7: Normal (OK), more than 0.8: Good, and more than 0.9: Great. From the previous studies, if the *ICC*(1,1) is more than 0.7, then the reliability is defined as good.

2. Results and discussion

The *BMS* and the *WMS* of the attainment time and the accuracy in the three courses are shown in Table 1. From these values, the *ICC*(1,1)s were calculated. The *ICC*(1,1)s of the attainment times were 0.951, 0.982 and 0.984 in the straight course, the circular course and the star course, respectively. The *ICC*(1,1)s of the accuracy S were 0.265, 0.924 and 0.711 in the straight course, the circular course and the star-shaped course, respectively.

From these results the reliabilities of the speed t and the accuracy S were confirmed in the circular course and the star-shaped course. On the other, the reliability of the accuracy S was not confirmed in the straight course. The reason was examined from Table 1and Eq.(1). In the case of the accuracy S in the straight course, the ratio of *BMS* and *WMS* was small as compared with the other cases. This means that the variability between the subjects was small and the variability in each subject was large. It is guessed that the reason was that the task of the straight course was easy for the healthy subjects.

As a result, it was confirmed that the developed measurement system has the sufficient reliability in the circular course and the star-shaped course.

Table 1. Results of *BMS* and *WMS*

		Three evaluating courses		
		(a)Straight course	(b)Circular course	(c)Star course
t [s]	*BMS*	4.01×10^{1}	2.10×10^{2}	2.97×10^{2}
	WMS	1.01×10^{0}	1.95×10^{0}	2.42×10^{0}
S [pixel]	*BMS*	1.17×10^{7}	1.29×10^{8}	6.01×10^{7}
	WMS	6.79×10^{6}	5.07×10^{6}	1.01×10^{7}

IV. Conclusion

We developed a measurement system of the skillfulness of lower extremities with tasks of the straight course, the circular course and the star-shaped course. The quantitative evaluation method of lower extremities in the circular course and the star-shaped course has sufficient reliability.

The future work is to verify the reliability of the proposed quantitative evaluation method to elderly persons and subjects with reduced skillfulness of lower extremities.

REFERENCES

[1] Shrout PE, Fleiss JL (1979), Intraclass Correlations: Uses in Assessing Rater Reliability, Psychol Bull, Vol.86:420-428

[2] Yamazaki K, Morita Y, Sato N, Ukai H, Nagaya M, Hayashi T, Ito T , Sakai Y (2013), Quantitative Evaluation System of Sensorimotor Integration Function and Proprioception of Lower extremities -Verification Effectiveness for surgical patients with intramedullary spinal cord tumor-, The Japan Society for Welfare Engineering, Vol.15, No.2: 44-49 (Article in Japanese)

[3] Landis JR, Koch GG (1977), The measurement of observer agreement for categorical data, Biometrics, Vol.33:159-174

[4] Kuwabara Y, Saito T, Inagaki Y (1993), Evaluation of intra- and inter-observer reliability, Kokyu To Junka, Vol.41, No.10:945-952 (Article in Japanese)

[5] Portney LG, Watkins MP (1993), Foundations lf clinical research-Applications to practice, Appleton & Lange, USA:505-516

Glove-Based Virtual Interaction for the Rehabilitation of Hemiparesis Stroke Patient

Khairuniam Wan, Aswad A.R, Rashidah Suhaimi, Nazrul H. Adnan , D. Hazry , Zuradzman M. Razlan and Syed Faiz Ahmed

Centre of Excellence for Unmanned Aerial Systems (COEUAS),
School of Mechatronic, Universiti Malaysia Perlis,
Kampus Pauh Putra, 02600 Arau, Perlis, MALAYSIA
(Tel: 604 9798333, Fax: 604 9798334)

khairunizam@unimap.edu.my

Abstract: This paper proposes the rehabilitation of a stroke patient by using dataglove called *GloveMAP*. *GloveMAP* is a low cost hand glove developed by using flexible bending sensor and accelerometer attached at the position of the fingers and wrist, respectively. The rehabilitation exercises are conducted in the virtual environment. In the studies, 22 fundamental movements of arm are evaluated and the results are employed to design the virtual environment. The combination of these fundamental movements focuses on Hemiparesis sufferers which are the most common in the stroke and have the inability to move one side of the body. The experimental results show that by combining several fundamental arm movements, glove based-virtual interaction supports the stroke patient do the rehabilitation by them without assisting from physiotherapist.

Keywords: Dataglove, VR interaction, Rehabilitation, Stroke.

1 INTRODUCTION

Strokes have been one of the major causes for long term disability. Death cases from stroke have decreased considerably but the survivors usually have to live with some form of disability. Given stroke cause damage to the brain by forming blood clot in the brain artery, the complexity of the brain may affect the stroke patient in various forms.

Virtual Reality (VR) is a digital environment which can be fully manipulated by the respective party. VR has its various function in today’s development in medical area which most have been used by the doctor for surgery simulation and also for psychological therapy for the mind. As it stand VR has serve its purpose in the development of this project for rehabilitation of the stroke patient [1-4].

This paper proposes the impairment of the stroke patient to rehabilitate by using VR approach. The study includes the research in investigation of muscle activities where the rehabilitation takes place.

This research paper organized as subsequent; Section 2 encompasses literature review of the related research and approach toward VR based stroke therapy. Section 3 presents the methodologies of applied procedures. Section 4 is divided into 2 sections, first section states about experiment setup where second section demonstrates the results of experiments. Finally section 5 expresses the conclusions over current research.

2 LITERATURE REVIEW

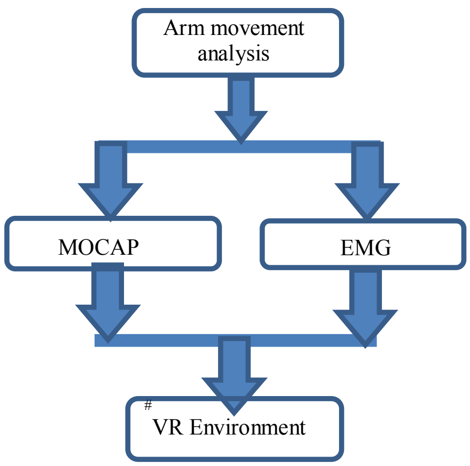
Worldwide stroke is the second leading cause of death, responsible for 4.4 million or equivalent to 9 % of the total 50.5 million deaths each year, and only 10 percent recover almost completely while some other survivor have to live with few disability or impairment. Thus this stroke patient needs special care and rehabilitation to gain back their body function. Most people over 55 years of age and with a male gender of family history of stroke could be one of many vital factor risk of getting a stroke.

Nowadays exists various application of the VR which have make rise to the development of interfacing devices such as dataglove, head-mounted display (HMD) and mechanical suit [3,5]. The application has various usages in the development therapy for the disability patient such as stroke patient. In this modern age lots of equipment have been develop for this VR applications. Though this technique requires a high cost to develop thus a lot improvement can still be made for this VR application.

Therapy by using VR is widely used in the society by integrating perceptions and experience within a simulation for the phobic patient. Rehabilitation can be made easy for the stroke patient as VR can help in attaining portable rehab center by remotely control the virtual exercise for the patient. Furthermore VR creates an environment where all objects can be created digitally allowing the user or patient to interact with the virtual object by using external hardware.

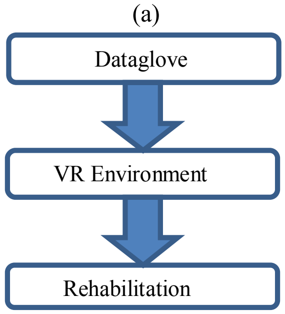
In this research, the dataglove called *GloveMAP* is used by the stroke patient to interact with the VR environments [6-8]. The most challenging part will be the designation of the virtual object installed in the virtual environments. The consultations from the medical doctor and the stroke therapist are required in order to design VR based stroke rehabilitation environments.

3 METHODOLOGIES



#Consultation from the stroke therapy required

(a)



(b)

Fig. 1. (a) Designing the VR environment (b) Rehabilitation teraphy

Fig. 1 shows the flow chart of the proposed methods. The proposed works are divided into 2 parts, which are the design of the Virtual Reality (VR) environments and the rehabilitation therapist [9]. Several preliminary experiments are conducted to design the VR environments. The purpose is to investigate the best sequence of arm movements for the rehabilitation purpose. In the rehabilitation process, the stroke patient has an option to choose the VR environments, which is suitable to their stage of the stroke.

2.1 Hemiparesis Stroke

Hemiparesis stroke attacks when blood flow to the brain stopped. The oxygen and blood that supply to the brain will stop which result in the damage of the brain which can affect the important body parts that control daily activities. Hemiparesis sufferer will experience one side of weakness. They often experience balance impairments which cause difficulties to perform movements such as grabbing objects, drinking from a glass, eating and using bathroom. The loss of abilities of movements due to stroke depends on the brain injury area.

2.3 Fundamental Movements in Stroke Rehab

2.3.1 Rotational Angles of Human Arm

Human arm consist of 7 DOF which are located at the shoulder, elbow and wrist. Shoulder part has 3 DOF for pitch, yaw and roll. While the elbow only has 1 DOF which make it can bend up and down. Finally the wrist that has 3 DOF for bending up and down, side to side deviation and twisting which make a total of 7 DOF for human arm [10].

Anatomical movements for hemiparesis stroke consist of several general movements that involve the shoulder, elbow, wrist and fingers. The movement carry out must be in correct range of motion to avoid injury during rehabilitation. Muscle is the tissues that provide movement for human body structure and the movement are classified by the direction of movement by the affected structures. Movements beyond the normal limit will be added with the prefix hyper- such as hyper flexion and hyper extension. Such movements are very important in medical field. General anatomical motion with its opposites includes flexion, extension, abduction, adduction, internal rotation, external rotation, elevation and depression.

2.4 Arm kinematics

Flexion and extension are the motion of adjusting angle between two parts by bending or straightening the parts. While, abduction and adduction are adjusting relation to midline of the body by pulling the body structure away or toward the midline, the abduction and adduction of the wrist are called as radial and ulnar deviation. Internal and external rotations are rotating the body parts inwards or outwards. The movements of adjusting the elevation in superior or inferior direction are elevation and depression. Special motions for hands and feets are pronation and supination which are the rotation of the forearm.

Human arm structure shows the rotational angles of shoulder and elbow define as yaw, pitch and roll. 3 DOF at the shoulder define the pitch angle β which is the rotation around y-axis for pronation and supination, yaw angle α which is the rotation around z-axis for abduction and adduction, while flexion and extension stand for roll angle γ which rotate around x-axis. The DOF at the elbow must be considered as two which are the pitch and roll angle, however pronation and supination will not change the hand position. A single transformation matrix can be formed by calculating the combination of eulerian angles in the order of Y-X-Y and Z-X-Z and the angles can be computed [11].

2.5 Virtual Reality (VR) software

Since this project needs to do the virtual rehabilitation therapy, the proposed software is Quest3D which is a platform of the VR system. Quest3D is the software for developing real-time 3D Microsoft Windows applications. It's consists of only a few high level software tools and almost all tasks are performed identical to the hardware. By using the software development kit, users can build their own components for Quest3D and build support for specific hardware likes dataglove. The Virtual Reality Peripheral Network (VRPN) will be used to interface the dataglove with the VR environment and Head Mounted Display (HMD). VRPN is a zero cost, open source tool that can handle many VR devices.

4 EXPERIMENTS

4.1 Experimental setup

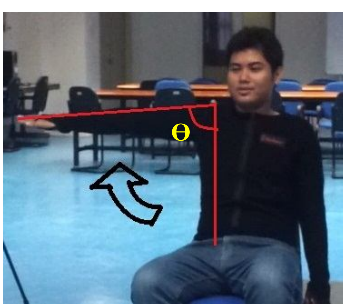


Fig. 2. The subject perform the fundamental movement

Table 1. The location of marker placement

No.	Location	ID Number
1	Shoulder	P1
2	Elbow	P2
3	Wrist	P3
4	Trapezium	P4
5	Index	P5
6	Middle	P6

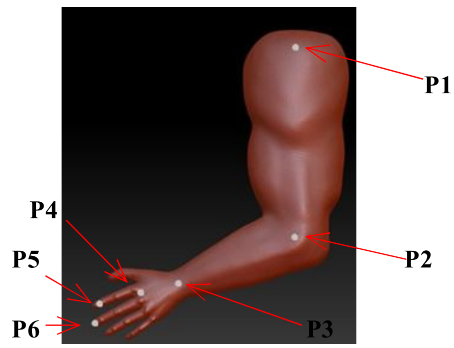
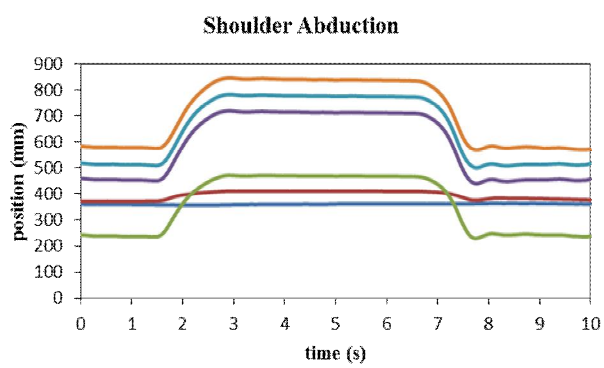


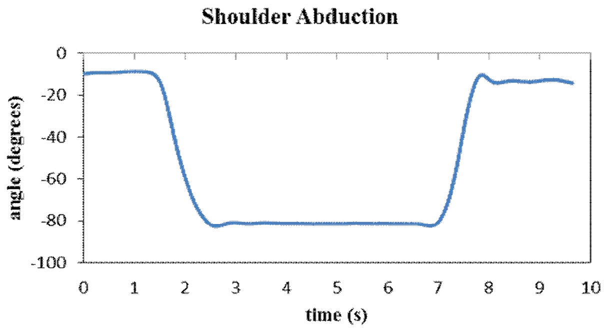
Fig. 3. Position of the 6 markers

In the experiments, 1 subject was selected to perform the fundamental movements of the arm as shown in **Fig. 2**. The movements were captured by using the motion capture system Qualisys Track Manager (QTM) software. The QTM was calibrated each time in the experiment to reduce the tolerance of the 3D measurement. Six markers were attached to the right arm of the subject, which was at the location of shoulder, elbow, wrist, and fingers (Index and middle) as shown in **Fig. 3**.

4.1 Experimental results



(a)



(b)

Fig. 4. Experimental results (a) Position for shoulder abduction (b) angle of shoulder abduction

Fig. 4(a) shows the markers at the position of shoulder (P1) and elbow (P2) experienced very little movement while the markers at the position of wrist (P3), index finger (P5) and middle finger (P6) show significant movement which about 250 mm of the vector magnitude.

5 CONCLUSIONS

The paper proposes the rehabilitation therapy of the stroke patient by interacting dataglove with the VR environments. Preliminary experiments are conducted to investigate the motion sequence of the arm by using MOCAP and EMG sensors for the purpose of designing the VR environments. The stroke patient does picks and place activities of the virtual objects in the VR environments. Based on these preliminary results, the portable stroke rehabilitation therapy device could be developed in the future. As the future work, the functional movements of the arm will be further investigated.

REFERENCES

[1] Holden M (2005), Virtual Environments for Motor Rehabilitation: Review. CyberPsychology and Behavior Vol. 8, No. 3, pp. 187-211

[2] Forducey P. et. al (2006), Tele-Rehabilitation using the Rutgers Master II glove following Carpal Tunnel Release surgery. Virtual Rehabilitation, International Workshop, pp. 88-93

[3] Deutsch J., Latonio J. and Burdea G (2001), Rehabilitation of Musculoskeletal Injury Using the Rutgers Ankle Haptic Interface: Three Case Reports. Eurohaptics'01, pp. 11-16

[4] Sveistrup H., Mccomas J. and Thornton M. (2003), Experimental Studies of Virtual Reality-Delivered Compared to Conventional Exercise Programs for Rehabilitation. CyberPsychology & Behavior, Vol. 6, No. 3, pp. 245-249

[5] Rose FD et. al (1996), Experimental Studies of Virtual Reality-Delivered Compared to Conventional Exercise Programs for Rehabilitati Virtual Reality: an Assistive Technology in Neurological Rehabilitation. Current Opinion in Neurology, Vol. 9, No.6, pp. 461-467

[6] Nazrul H. ADNAN et. al (2013), Principal Component Analysis For The Classification Of Fingers Movement Data Using DataGlove "GloveMAP". International Journal of Computer Engineering & Technology (IJCET), Vol. 4, No.2, pp. 79-93

[7] Nazrul H. ADNAN et. al (2013), PCA-based Finger Movement and Grasping Classification using DataGlove "GloveMAP". International Journal of Innovative Technology and Exploring Engineering (IJITEE), Vol. 2, No.3, pp. 66-71

[8] Nazrul H. ADNAN et. al (2013), Classification Of Finger Grasping By Using PCA Based On Best Matching Unit (BMU) Approach. International Journal of Advanced Research In Engineering And Technology (IJARET), Vol. 4, No.2, pp. 92-105

[9] Aswad A.R et. al (2014), Glove Based Virtual Reality (VR) Interaction for the Purpose of Rehabilitation. Australian Journal of Basic and Applied Sciences, To be published.

[10] A. A. Ali and A. H. Miry (2011), Human Arm Inverse Kinematic Solution Based Geometric Relations and Optimization Algorithm. International Journal of Robotics and Automation (IJRA), Vol. 2, No.4, pp. 245

[11] Khairunizam Wan and Hideyuki Sawada (2007), 3D Motion prediction of human upper body by tracking reflective markers on a moving body. Malaysia Japan International Symposium on Advanced Technology (MJISAT2007)

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Measuring system of therapist’s guiding motion for standing up training of patient with hemiplegia after stroke and analysis results of patient's motion

JungTang Wang *, Kazunori Yamazaki *, Noritaka Sato *, Yoshifumi Morita *, Hirofumi Tanabe**

**Dept. of Computer Science and Engineering, Nagoya Institute of Technology, Japan*
***Self-Defense Force Central Hospital, Japan*
Tel : +81-52-735-5412; Fax : +81-52-735-5412
24417627@stn.nitech.ac.jp

Abstract: Although a patient with slight hemiplegia after stroke can stand up by using both his/her legs, he/she stands up by using mainly his/her healthy leg rather than his/her paralyzed leg. If this motion is repeated as the self training, the patient standing up motion cannot be restored to normal. Therefore, a therapist guides a patient motion to be able to stand up by using both his/her legs well in balance or by putting more weight on the paralyzed leg. The final goal of our research is to develop a rehabilitation support robot of standing up training for patients with hemiplegia after stroke to restore normal standing up motion. For this purpose, we analyze the therapist’s guiding motion, and to design and manufacture a rehabilitation support robot of standing up training on the basis of the analysis results of the therapist’s guiding motion.

In this paper we developed a measuring system to analyze the therapist’s guiding motion during the standing up training of a patient with hemiplegia after stroke. To confirm the usefulness of the developed measuring system, we measured the patient's motion during the standing up training with and without the therapist's guiding motion. The usefulness of the measuring system was confirmed by comparing the results of the patient's left/right weight balance during the standing up motion and the therapist’s objective of guidance.

Keywords: Standing up training, Hemiplegia after stroke, Measuring system, Weight balance, Guiding motion

I. Introduction

In Japan the number of stroke patients are increasing, and 1.235 million stroke patients are living in 2012 [1]. The rehabilitation is important for the patient's motor function to be restored to a normal condition and then for the patient to return to daily life. The corresponding increase in the workload of therapists has become a social issue. Therefore, the patients usually perform the standing up training by themselves at present. During the self-training, the patient feels anxious about the use of a paralyzed leg in the standing up motion. For this reason, the patient stands up by using mainly the healthy leg rather than both the legs. In this case the patient' standing up motion cannot be restore to normal. If the therapist guides the patient's standing up training as shown in Fig.1, the recovery effect is improved for normal standing up motion. Therefore, a rehabilitation support robot which can simulate the therapist’s guiding motion is greatly desired. The final goal of our research is to develop a rehabilitation support robot of standing up training for patients with hemiplegia after stroke to restore normal standing up motion. For this purpose, we analyze both the patient's guided motion and the therapist’s guiding motion in the patient's standing up training. And then we design and manufacture a rehabilitation support robot of standing up training on the basis of the analysis results of the therapist’s guiding motion.

Fig.1 Standing up training guided by therapist

In this paper we developed a measuring system to analyze the therapist’s guiding motion and the patient’s motion guided by the therapist during the standing up training of the patient with hemiplegia after stroke. In order to confirm the usefulness of the developed measuring system, we analyzed the patient's motion during the standing up training with and without the therapist's guiding motion.

II. Standing Up Training of Patient with Hemiplegia after Stroke

In order to recover motor function of a paralyzed leg, standing up training is the one of important rehabilitation training. Although a slight stroke patient can perform the

(a) Phase 1 (b) Phase 2 (c) Phase 3

Fig.2. Three phases of standing up motion.

standing up training by himself/herself, the self-training may not be effective in the case of repetition of abnormal standing up motion.

2.1. Characteristic of standing up training

It was found form the therapist's interview that a patient uses mainly his/her healthy leg as compared with his/her paralyzed leg when standing up from the sitting position. There are two reasons. One is that the muscle strength of the paralyzed leg is insufficient for normal standing up motion. The other is that the patient feels anxious for putting weight on the paralyzed leg and fear of falling down. In this case the self-training is not effective for restoring normal standing up motion.

2.2. Therapist’s guiding motion

The therapist guides the patient’s standing up motion by pulling or pushing the patient’s waist during the standing up training. There are two types of the therapist's guiding motions for two training objectives. One is the training of the paralyzed leg by putting more weight on the paralyzed leg as compared with the healthy leg. The other one is the training of the paralyzed leg by making the patient stand up in good balance with both the legs.

2.3. Analyzing of standing up motion

In order to analyze the standing up motion, it is necessary to divide the standing up motion on the basis of the features. In the previous research [2][3][4], it was found that the standing up motion of healthy persons can be divided into four phases. Phase I is the motion from the initial position to leaving the chair. Phase II is the motion from leaving the chair to reaching the ankle’s angle to the maximum dorsiflexion. Phase III begins after Phase II and ends when the hip angular velocity reaches 0 deg/sec. Phase IV begins after Phase III and ends when all motions are stable. Since the standing up motion of a patient with hemiplegia after stroke is different from that of a healthy person, it is difficult to distinguish Phase II and Phase III. For this reason, we divided the patient's standing up motion into three phases, namely Phase 1, Phase 2 and Phase 3. Phase 1 are Phase 3 are equal to Phase I and Phase IV, respectively. Phase 2 is composed of Phase II and Phase III.

In this study, we used the left/right weight balance to evaluate the ability of the patient’s standing up motion. Let W_L, W_R denote the weights which are applied to the left and the right Wii Balance Boards, respectively. Let R_L and R_R denote the left and right ratios of weight balance of the left and right legs, respectively. R_L and R_R can be calculated by Eqs. (1) and (2).

$$R_L = \frac{W_L}{W_L + W_R} \quad (1)$$

$$R_R = \frac{W_R}{W_L + W_R} \quad (2)$$

$$R_L + R_R = 1 \quad (3)$$

The relationship between the weight balance and the ratio of weight balance are shown in Table 1.

Table 1 Relationship between WB(weight balance) and RWB(ratio of weight balance)

Balance state	WB	RWB
Normal	$W_L=W_R$	$R_L=0.5$ ($R_R=0.5$)
Inclination to the right	$W_L<W_R$	$R_L<0.5$ ($R_R>0.5$)
Inclination to the left	$W_L>W_R$	$R_L>0.5$ ($R_R<0.5$)

III. Motion Measuring System

In standing up training the therapist guides the patient's motion while paying attention to the stability of the trunk, the synchronized movement of both the legs, and the weight balance of both the legs. In order to analyze the patient's guided motion and therapist’s guiding motion during the standing up training, we constructed the measuring system on the basis of the therapist’s interview.

3.1 Sensors for patient's guided motion

In order to measure the patient's guided motion, we used the following sensors. Three wireless acceleration/angular velocity sensors (ATR-Promotions, WAA-006) were used to measure the attitude and the movement of the left and right lower limbs and the trunk. The two sensors were mounted on the patient's tuberositas tibiaes with the Velcro belts, and the other one was mounted on the xiphoid process with the Velcro belt as shown in Fig.3 (a) The pressure sensitive sensor was attached to the seat surface of the chair to detect a timing at which the buttocks leaves the seat surface.

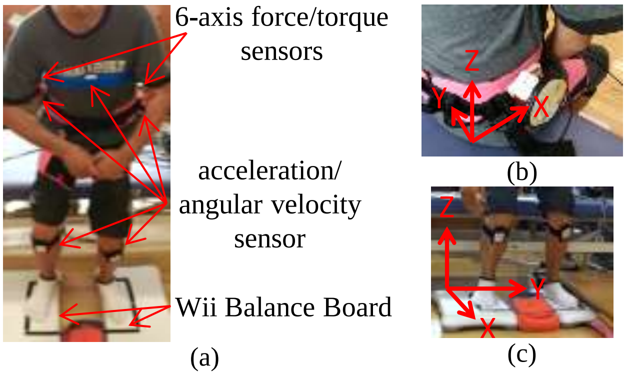


Fig.3. Subject with sensors

The two Wii Balance Boards were used, because the therapist has to put therapist's foot between the patient's legs in order to hold the patient's waist during the standing up motion. The two Wii Balance Boards measure the weight balance and center of pressure on the foot.

3.2 Sensors for therapist's guiding motion

In order to measure the therapist's guiding motion, we constructed the equipment for measurement which is fixed to the patient's waist. The equipment for measurement is shown in Fig.3(b). The two 6-axis force/torque sensors (WACOH-TECH Inc., WEF-6A200-4-UG5) were fixed to the equipment for measurement to measure the therapist's force and torque applied to the left and right sides of the patient's waist. The force/torque sensors were located on the iliac crests. The handle for the therapist to pull and push the patient was fixed to the force/torque sensor. The two wireless acceleration/angular velocity sensors were fixed to the force/torque sensor to measure the attitude of the force/torque sensor. The appearance of the patient mounting the sensors was showing in Fig.3 (a) (b) (c).

3.3 Measuring system

The two digital video cameras (Microsoft Corp., Kinect) were set up to capture the movement of the sagittal plane of the patient and the movement of the diagonal from the rear of the therapist. The data were collected from the sensors and cameras in synchronization. The sampling frequency is 30Hz. The appearance of the patient's standing up motion without the therapist's guiding motion is shown in Fig.4. The inclination of the trunk, the lower limbs and the force/torque sensors are calculated from Eqs.(4) where θ is the angle of inclination, a_y and a_z are the y and z axes components of acceleration that are measured by the acceleration/velocity sensors.

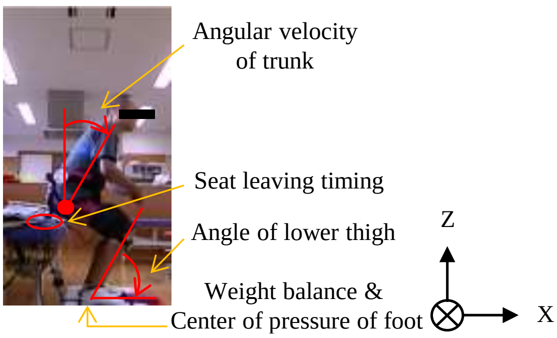


Fig.4. Measuring quantity in subject's motion

$$\theta = \tan^{-1}\left(\frac{a_y}{a_z}\right), \qquad (0 \le \theta \le 180°) \tag{4}$$

IV. Experiments

In order to verify the usefulness of the developed measuring system, we analyzed the standing up motions of a healthy person and a patient with hemiplegia after stroke with/without the therapist's guidance. For this purpose, we measured the standing up motions by using the developed measuring system and calculated the ratios of weight balance of the standing up motion.

4.1. Experimental conditions

The patient with hemiplegia after stroke was a male with paralysis of the left half of the body. The Brunnstrom stage was IV. The patient stands up by using mainly the healthy (right) leg. The height of the seat surface of the chair was 42 cm. The standing up speed was not instructed in the self-standing up training. The experiment was performed two times. The first time was the standing up training without the therapist's guiding motion, which means the self-standing up training. The second time was the standing up training with the therapist's guiding motion.

4.2. Standing up motion without guiding motion.

The patient's self-standing up motion was measured with the developed measuring system and analyzed. The standing up motion was divided into three phases. The time histories of the self-standing up motion are shown in Fig.5. The time histories of the ratios of weight balance are shown in Fig.6. It's seen that the patient's weight was inclined to the healthy side in Phase 2 and the patient can standing well in balance in Phase 3. As the comparison study, we measured the standing up motion of the healthy person. The time histories of the ratios of weight balance are shown in Fig.7. The ratios of weight balance were almost equal to 0.5 during standing up motion. This means that the healthy person can stand up well in balance.

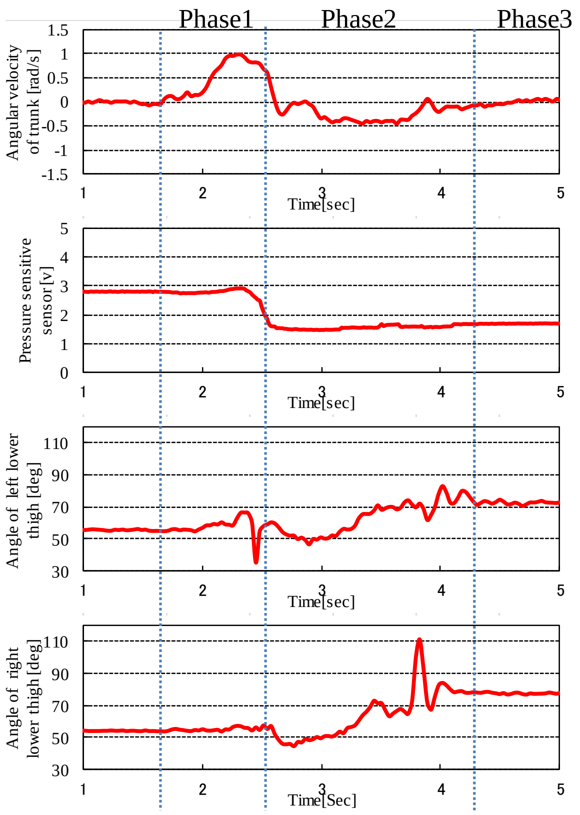


Fig.5. Time histories of standing up motion without the therapist's guiding motion.

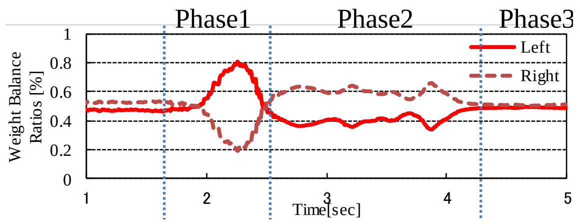


Fig.6. Time histories of left/right weight balance ratios without the therapist's guiding motion.

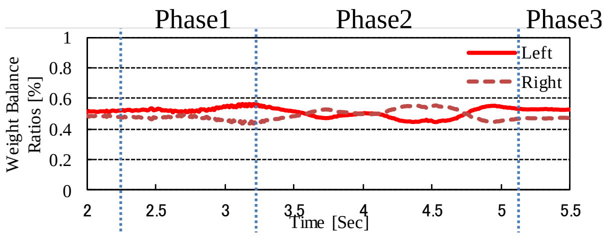


Fig.7. Time histories of left/right weight balance ratios of the healthy person.

4.3. Standing up motion with therapist's guiding motion

The patient's standing up motion with the therapist's guiding motion was measured with the developed measuring system and analyzed. The time histories of the ratios of weight balance are shown in Fig.8. It's seen that the patient's weight was inclined to the paralyzed side in Phase 2. The therapist guided the patient's standing up motion so that the patient uses mainly the paralyzed leg for restoration of the motor function of the paralyzed leg. It was confirmed that the measuring results agreed with the therapist's training objective.

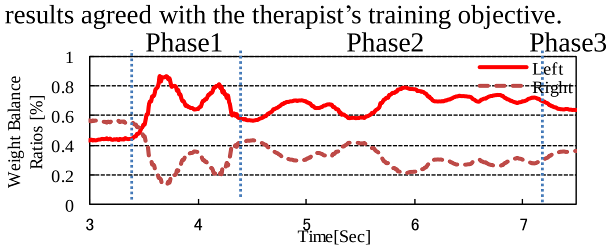


Fig.8. Time histories of left/right weight balance ratios with therapist's guiding motion.

V. CONCLUSION

The usefulness of the measuring system was confirmed from the analysis results of patient's standing up motion with and without the therapist's guiding motion.

The future work is to analyize the therapist's guiding motion and force/torque during the patient's standing up motion, and to design and manufacture a rehabilitation support robot of standing up training on the basis of the analysis results of the patient's and the therapist's motions.

REFERENCES

[1] Ministry of Health, Labour and Welfare (2012), "Overview of the 2011 patient survey," http://www.mhlw.go.jp/toukei/saikin/hw/kanja/11/index.html.

[2] M. Schenkman, et al. (1990), "Whole-Body Movements During Rising to Standing from Sitting", Physical Therapy, Vol.70, No.10, pp.638-648.

[3] D. Chugo, et al. (2006), "Force Assistance System for Standing-Up Motion", Proc. of The 9th International Conference on Climbing and Walking Robots (CLAWAR2006), Brussels, Belgium, pp.65-70.

[4] T. Takayuki and K. Seiki (2005), "A Study of Control Mechanism for of Standing-up Achieving a Motion Task; Biomechanical Motion Analysis of Standing-up", Rigakuryoho Kagaku Vol.20, No.4, pp.303-307.

Image Segmentation of Coronary Artery Plaque Using Intuitionistic Fuzzy C-Means Algorithm

Zahra Rezaei[a,1], Ali Selamat[a,2], Mohd Shafry Mohd Rahim[a,3], Mohammed Rafiq Abdul Kadir[b,4]

[a]K-Economy Research Alliance & Faculty of Computing, Universiti Teknologi Malaysia
[b]Faculty of Biomedical Engineering & Health Sciences, Universiti Teknologi Malaysia
[1]Rzahra.rezaei@gmail.com, [2]aselamat@utm.my, [3]shafry@utm.my, [4]rafiq@biomedical.utm.my

Abstract : Every year, hundreds of thousands of people die because of Coronary Heart Disease (CHD) in all over the world. Coronary Artery Disease (CAD) as a cardiovascular illness causes blood vessels narrowing that supply blood and oxygen to the heart . Atherosclerosis is known as the deadliest type of heart disease, which is caused by soft or “vulnerable” plaque (VP) formation in the coronary arteries. Acute Coronary Syndrome (ACS) is recognized as the first coronary atherosclerosis indicator which identifies high-risk plaques. Intravascular ultrasound (IVUS) can be applied for characterization of plaque and segmentation of vessel’s walls borders. Recently, Virtual Histology as a new approach based on spectral analysis of IVUS has been introduced. In this work, we applied a clustering method based on Intuitionistic Fuzzy C-means (IFCM) in order to automatic segmentation of Coronary Artery plaque using VH-IVUS images.

Keywords – Vulnerable plaque; Intravascular ultrasound; optical coherence tomography, Fuzzy c-means

1 INTRODUCTION

Coronary Artery Disease (CAD) is the result of accumulation of plaques within the coronary arteries. The plaque is made up of fat, cholesterol, calcium, necrotic core along with fibrous shell , flat muscle cells, foam cells and T-leukocytes components (Fig. 1). This condition is called Atherosclerosis. Atherosclerotic plaques affect the wall of coronary artery and result in luminal stenosis. It is contribute to cause of death when the degree of stenosis is over a significant level (75%) [1, 2]. Approximately, 60% of plaque rupturing occurring in ‘‘shoulders’’ or at cap–intima junctions as a result of circumferential stresses and the remaining 40% are away from the junctions [2].

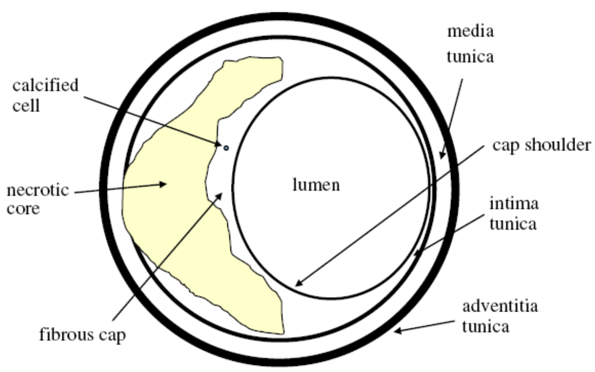


Figure 1. Arterial plaque geometry. Cross-sectional view [2].

Over time, plaque become harden or rupture where the break start to occur. If the plaque ruptures, a blood clot is formed on its surface. Vulnerable Plaque (VP) is likely to rupture, and leads to acute coronary events, such as unstable angina, heart attack, myocardial infarction (MI), stroke and sudden death. Therefore, early detection of high-risk plaque is a critical task in medical science along with its major influences on treatment and prevention of heart event. Recently, characterization of VP is growing in order to acute coronary event prevention [3]. Studies show that cardiac events happen as a result of fat plaque burden , tiny lumen region and thin-cap fibroatheromas (TCFA). It has been widely demonstrated that acute coronary syndromes (ACS) are related to rupture and acute thrombosis over a mildly stenotic plaque, rather than to a slow growth with final occlusion of a plaque encroaching the lumen [4]. The wide clinically practical invasive and the non-invasive imaging techniques to identify ”VP” has been introduced. Non-invasive techniques such that : X-ray angiography, Ultrafast Computed Tomography (UCT), nuclear scintigraphy, Magnetic Resonance Imaging (MRI) and tnvasive techniques: intravascular angioscopy intravascular thermography, spectroscopy, Intravascular Ultrasound (IVUS), IVUS-Palpography, IVUS-VH and Optical Coherence Tomography (OCT) [1]. Virtual Histology-Intravascular Ultrasound (VH-IVUS) image has been applied as invasive clinical technique to reliable identification, quantification, and characterization of vulnerable plaques [1]. In addition, virtual histology (VH-IVUS) as sophisticated tissue characterization technique using RFA, can display tissue map as color-coded way. Plaque components in VH-IVUS is classified into four basic tissue types such as : fibrous (dark-green pixels), fibro-fatty (light-green pixels), necrotic core (red pixels), and dense calcium (white pixels) [1, 6 ,7]. Furthermore, VH-IVUS can be applied in quantitative cross-sectional lumen area, assessment of thickness and vessel wall remodelling as well as progression of plaques monitoring [8].The providing regional observations to study predictors of events in the coronary tree study or PROSPECT study organization [9] is an international multicentre prospective study that is responsible for adverse conditions of coronary events by evaluating plaque types in VH-IVUS image. Recently, optical coherent tomography (OCT) has introduced as a new medical imaging technology that have a higher resolution compare to the IVUS[10].

This paper is organized into three sections as follows: section 2 describes literature review, section 3 defines the concepts of vulnerable plaque and the details of segmentation method based on Intuitionistic Fuzzy C-means (IFCM) is introduced in section 4.

2 LITERATURE REVIEW

Gradient vector flow (GVF) as a new method of segmentation for IVUS image has been proposed by Zhu et al. in [11].They used nonlinear filtering for decreasing the critical points and variation in curves structure along with the balloon snake. Their method solves the local convergence problem. According to [12], multiple k-nearest neighbor (MkNN) has been proposed and applied for tissue characterization of coronary artery plaque in IVUS image. As a substitute for labeling k-nearest neighbor (kNN), a distances between input vector and the prototype vectors was used as a weighted decision. In order to perfectly characterized lipid and fibrous tissues. They have done some actual experiments for simulated data and IVUS image. In order to describe the luminal border, a method based on minimization of probabilistic cost function was presented in IVUS B-mode in [13]. Mendizabal-ruiz et al. [13] used concept of curve lumen contour due to minimizing a cost function. Support Vector Machine (SVM) classifier was applied to decide whether a pixel is blood or non-blood, lumen or non-lumen. Their proposed method can segment IVUS B-mode images in diverse type and frequencies. In order to shadow the detection in IVUS image, an adaptive threshold method was used by Basij et al. in [14]. They improved segmentation of the classified plaque by using active counters in order to separate the shadow region efficiently. This method segments the hard plaques tissues using intensity value in calcified region. The achieved outcomes for detecting shadow regions proved that this method has efficiency even in segmentation of challenging IVUS images. According to [1], a robust approach was recommended based on IVUS images for characterization of the plaques. Their algorithm involves shadow detection, feature extraction and classification. They used plaque area that is yield from VH-IVUS images for validation of their proposed method. Also, they applied texture analysis of both Ex-vivo and in-vivo data to extract Dense Calcium (DC), Necrotic Core (NC), Fibrotic Fibro-Fatty (FF) portion of plaque. Feature extraction method such as local binary pattern (LBP), modified Run-length (MRL) and neighbouring grey-level (NGL) were applied. In addition, the SVM and Error-Correcting Output Codes (ECOC) classifier are used for arrangement of each pixel into one of the three classes (calcium, necrotic core and fibro-fatty). Another method to overcome tissue characterization of coronary artery plaque challenges was proposed in [15]. Vachkov et al. introduced histogram-based and centre-of-gravity-based based on general computational strategy for moving window. The obtained similarity results of two real data sets - training and test data set are also demonstrated. Threshold is applied in hard decision and gradual changes were used for soft decision in the dissimilarity values by Athanasiou et al. [16]. They have proposed a method in order to various atherosclerotic plaque characterization using geometrical features in VH-IVUH Images. They applied intensity features, texture based features and two novel geometrical features along with Random Forests classification method.

3 CONCEPTS OF VULNERABLE PLAQUE

The phrase of vulnerable plaque descriptions which is related to concept of acute cardiovascular disease was introduced 20 years before [17]. In 2010, according to the study that has been done by PROSPECT study organization [9], prediction of vulnerable plaque is affected by three factors which are Virtual Histology-Thin Cap Fibro Atheroma (VH-TCFA), burden of plaque more than 70% along with an Minimal Lumen Area (MLA) a smaller amount than 4 mm [18]. According to experts, definition of the plaque rupture has different descriptions so it is not only related to the VP as shown in Fig. 2.

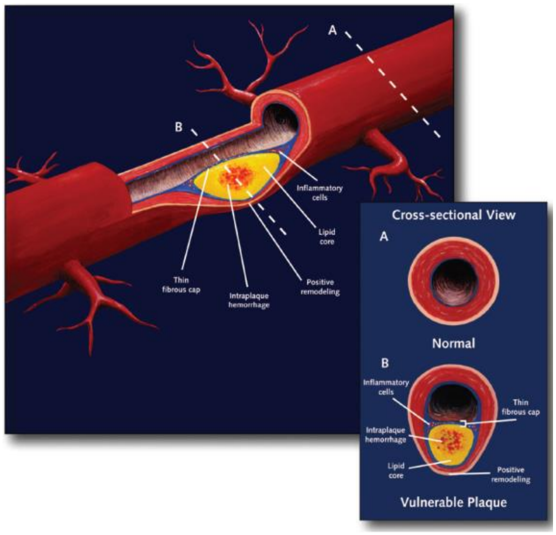


Figure 2. Normal arterial in segment (A) compared with a vulnerable plaque in section (B) along the length of the vessel and arterial wall cross-sectional image [17].

High likelihood of thrombotic complications and speedy progression are two major signs of atherosclerotic plaques indicate VP. Also, morphological parameters include thickness of cap, lipid core volume , plaque color, content of collagen, content of lipid ,solidity, calcification and stream pattern within coronary artery tree are indicators of VP [10]. There are two criteria in identifying plaque vulnerability. The major one is active inflammation cells (about 26% macrophage infiltration of fibrous cap); a thin fibrous cap less than <100μ along with a large lipid core less than 40% of the plaque's total area; endothelial denudation with superficial platelet aggregation; signs of recent rupture in fissured cap; or strict coronary stenosis and eccentric lumen [17].The minor case consists of calcified nodules; larger lipid core is indicated by yellow color; bleeding of intraplaque; dysfunction of endothelial and remodeling in positive side [17].

3.1. Virtual Histology-Intravascular ultrasound (VH-IVUS)

Ultrasound is a medical imaging technique based on catheter and is formed by ultrasound radiofrequency signal for constructing real time image inside of vessel. These waves produce electrical signals and send them to an external system in order to processing tasks such as amplification, filtering, and scan conversion [9]. IVUS strategy can provides a two-dimensional cross sectional views of image along with assessment burden of plaque [10].Three-vessel lumen demonstration and also characterization of plaque are provided by IVUS. Calcified areas and dense fibrous parts are homogeneous and have bright reflectance in gray-scale images of IVUS. On the contrary, “soft” or “mixed” lesions are represented by fewer echoes. The IVUS equipment has three main component include a catheter instrument along with a miniaturized transducer and a console for reconstructing and demonstrating the output IVUS image (Fig. 3).

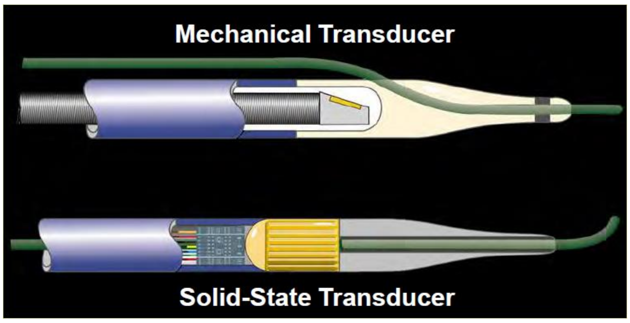


Figure 3: Catheter designs

To improve image resolution, ultrasound should be used in high frequencies. Since declining saturation of tissue is related to increasing the frequency of ultrasound, it's frequency should be limited to 45 MHz [8]. In order to generating cross-sectional image, electronic systems activated serially and use 64 transducer elements that is ordered in an annular array.Electronic catheters of IVUS can be used for characteristic lumen and wall of boundaries by demonstrating couloirs blood flow.Lately, virtual histology (IVUS-VH) as invasive clinical technique has been proposed based on the 20 MHz IVUS platform to assist analysis of image into special tissue components through identification, quantification, and characterization of VP [9]. Plaques components can be discriminated into three classes base of echogenicity consist of calcified tissue with extremely echo reflective range, fibrosis tissue that is characterized by hyper echoic areas and thrombus or lipid-rich tissue representing by hypo echoic regions. Whereas, tissue constituents are different in hardness, VH can classify soft and hard components along with the assessment’s properties of vessel wall [10]. This method classified plaque components into four types involves Dense Calcium (DC), Fibrotic Tissue (FT), Fibro-Fatty Tissue (FFT) and Necrotic Core (NC) that is shown in Fig. 4 [17].

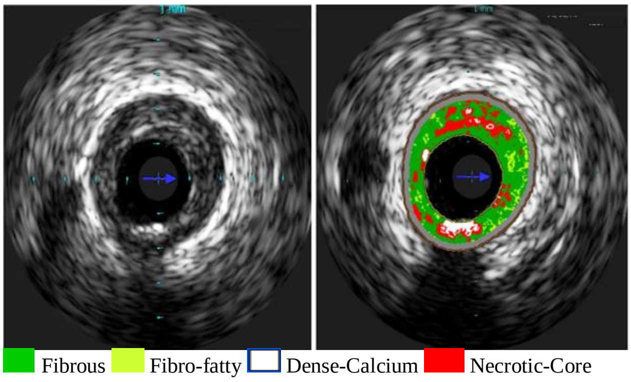


Figure 4: VH-IVUS color code

However,VH can distribute IVUS image into differentiate tissue components; but it is difficult to divide the plaque owing to heterogeneous nature along with overlapping components (Fig. 5).

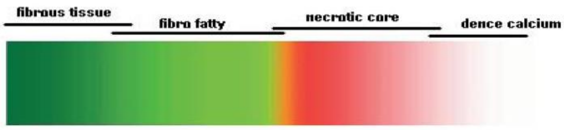


Figure 5. tissue spectrum of VH illustrating four plaque component include fibrous tissue, fibrofatty, necrotic core, and dense calcium by different color [19]

Fig. 5 demonstrates plaque components overlapping that causes inaccuracy difficulty in VH. Progression of disease between fibrous and fibro-fatty area have common characteristics in VH images as well as necrotic core and dense calcium. So, VH image analysis can detect at-risk lesions but non-critical. VH images along with interventional cardiologist can interpret plaque into four components and their related location in order to measurement the vulnerability of plaque [17].Several expert softwares have been implemented such as Volcano (Fig. 6).

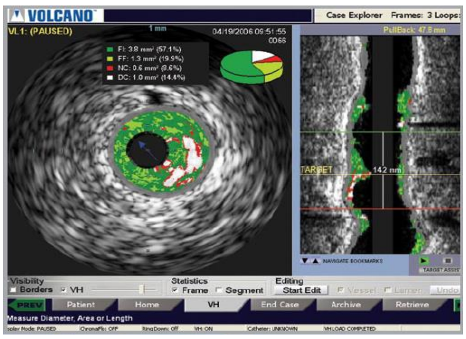


Figure 6.Volcano software

Volcano’s VH-IVUS technology classifies 4 tissue types of plaque by means of advanced, proprietary spectral analysis methods with 93-97% of accuracy. This equipment has been used for assessment of patient’s disease severity by image representation. Furthermore, Volcano software has advanced medical tools such as Lesion Risk Assessment, Transplant Patient Monitoring, Carotid Stenting, Longitudinal Stent Placement and Distal protection [20].

3.2 IVUS and VH limitation

IVUS technique has significant difficulties to distinguish and exact assessment plaque component especially in areas of low contrast include fibrous tissue, fibro-fatty tissue and thrombus part. Besides,VH technique has various limitations such as: (i) poor recognition of VH-TCFA as a result of fibrous tissue that causes false detection in thrombotic material; (ii) Misdiagnosis red halo around the white spots and (iii) The range of its resolution is more than limitation of VH-TCFA cap (65 μm) [18].

3.3 Virtual Histology–Thin Cap Fibroateroma (VH-TCFA)

According to recent studies , thrombus formation in lumen causes plaque rupture and led to more than 70% of acute coronary events. TCFA is recognized as the most frequent cause of thrombus formation. This type of plaque has fat necrotic core along with thin fibrous cap that is covered by macrophages particles, lymphocytes along with reduced smoothness of muscle.This circumstance causes clotting cascade triggering and thrombus formation that can provide adequate conditions for acute coronary syndrome (ACS). Rupturing lipid-rich atheroma causes coronary block due to VH-TCFA with thin fibrous layer along with necrotic core. It is reported to be the most common cardiac basis (60%) of myocardial infarction and death base on pathological trainings. VH-TCFA has been defined by two major principles include: (i) plaque burden more than 40%; (ii) necrotic core more than 10% that is located near the lumen directly with no signs of fibrous tissue. We can classify VH-TCFA into four categories which are VH-TCFA I, by calcium amount less than 5%; VH-TCFA II, by more than 5% calcium; VH-TCFA III, with various layers; VH-TCFA IV that is determined by necrotic core value more than 20% along with calcium more than 5%, remodeling criterion larger than 1.05, and plaque burden more than 50%. All four types of classification of VH-TCFA and measurement necrotic core are needed for identifying the degree of plaque vulnerability [18].

3.4 IVUS and OCT

OCT image is comparable to IVUS. IVUS-VH not able to assess the thin fibrous cap due to limitation in axial resolution (100–200 mm). However, OCT can identify the microstructure with high-resolution imaging technique (10–20 mm). The penetration of OCT signal is limited to 1–2 mm, so it cannot identify the lipid pools or amount of calcium behind thick fibrous caps and it produces inexact detection (Fig. 7). IVUS can not complete thrombus characterization precisely due to lack of resolution. In a test on patient's investigations it was revealed that OCT could observe all intracoronary thrombus's cases and angioscopy as a result of acute myocardial infarction (AMI) ; but, IVUS identified only 33% of them [9]. Recently, combination of OCT and IVUS as two imaging technique were used for vulnerable plaques detection.High resolution of OCT image can identify vulnerable plaque for in vivo image. Intravascular OCT has potential to visualisation of normal and atherosclerotic structure in coronary arteries. It seems that in order to achieve high precision of TCFA discovery, IVUS-VH image and OCT should be combined [9,10].

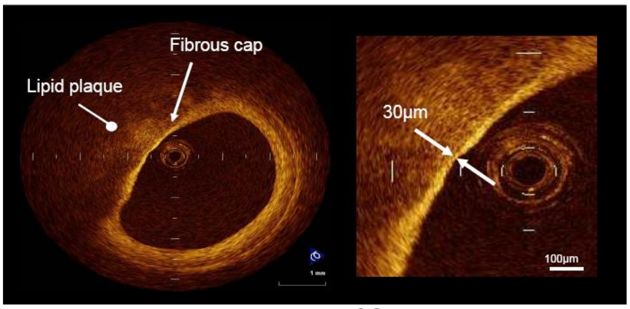


Figure 7 . OCT

4 SEGMENTATION

The first and essential step in plaque characterization is segmentation. Performing medical image segmentation manually is a difficult, individual, time-consuming procedure, and it depends on observer ability. Therefore, several methods of automatic image segmentation have been developed. These methods are categorized based on different parameters. Segmentation algorithms such as direct detection of the border(s) involves Edge-Tracking frameworks, Gradient-Based approaches, methods of Active Contour-Based, deformable models along with Statistical shape model, Probabilistic-Based Techniques, graph searching, gradient-driven algorithms and clustering methods.

4.1 Intuitionistic Fuzzy C-means Algorithm (IFCM)

Dividing distinctive homogeneous groups of image , including color and texture for classifying different patterns called as clustering. Clustering method is based on crisp or fuzzy. although, each pixel of a distinct clusters (i.e. 0 or 1) in non-fuzzy or crisp clustering is belonged to one group. Nevertheless, in fuzzy clustering method, membership value is associated with each pixel could be a member of several clusters. [21]. In IFCM algorithm, X is a nonempty fixed set and A in X is an object which defined as:

$$A := \{ < x, \mu_A(x), \nu_A(x) > \mid x \in X \} \qquad (5)$$

For each $x \in X$, $\mu_A : X \to [0,1]$ consider as the value of membership function .Also, the degree of non-membership group is $\nu_A : X \to [0,1]$. For every $x \in X$, μ_A and ν_A satisfy equation 6:

$$0 \le \mu_A(x) + \nu_A(x) \le 1 \quad (6)$$

Unlike traditional fuzzy, The summation of μ_A and ν_A is not necessarily had to be 1. IFCM function sums two equations: (i) conventional FCM objective function which is modified by IFS and (ii) IFE which shows entropy of intuitionistic fuzzy.
Minimizing objective function using IFCM is:

$$J_{IFCM} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{*m} d_{ik}^{2} + \sum_{i=1}^{c} \pi_i^{*} e^{1-\pi_i^{*}} \qquad (7)$$

$$u_{ik}^{*} = u_{ik} + \pi_{ik},$$

$$\pi_i^{*} = \frac{1}{N}\sum_{k=1}^{n} \pi_{ik}, \quad k \in [1, N]$$

Where the intuitionistic fuzzy membership is μ_{ik}^{*} ,the kth data in ith class. μ_{ik} is belong to conventional fuzzy membership and π_{ik} is Hesitation degree in this equation:

$$\pi_{ik} = 1 - u_{ik} - (1 - u_{ik}^{\alpha})^{1/\alpha}, \alpha > 0, \qquad (8)$$

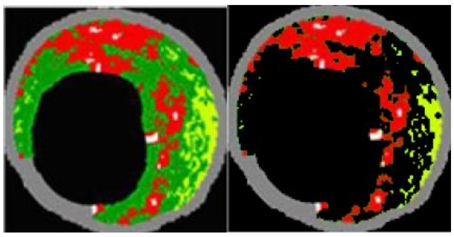
Yager's intuitionistic fuzzy complement used as fallow equation:

$$N(x) = (1 - x^{\alpha})^{1/\alpha}, \alpha > 0, \qquad (9)$$

Finally, intuitionistic fuzzy set equation is:

$$A_{\lambda}^{IFS} = \{x, \mu_A(x), (1 - \mu_A(x)^{\alpha})^{1/\alpha} | x \in X\} \qquad (10)$$

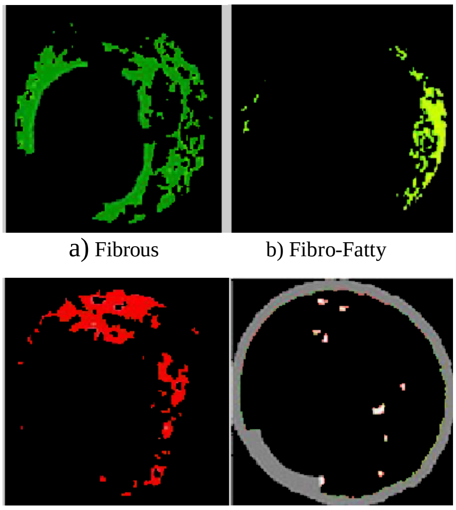
The algorithm is implemented using MATLAB software. The results with considering various numbers of clusters were different. Fig. 8 shows the result of clustering algorithm with three clusters. As seen in the figure, the algorithm is not able to separate the components properly, because the number of clusters is inappropriate.



a) VH-IVUS image b) segmented result

Figure 8: IFCM with 3 cluster

However, increasing the number of clusters can raise the accuracy of the IFCM algorithm. Fig. 8 shows the results by using seven clusters which could separate the components more accurately.



a) Fibrous b) Fibro-Fatty

c) Dense-Calcium d) Necrotic-Core

Figure 9: IFCM with 7 cluster

5 CONCLUSION

An arterial lumen can be threatened by plaque with lumen narrowing or stenosis and rupturing of plaque cap. Recent techniques, such as VH technology can identify atheromatous plaques. However, efforts should be made to improve its accuracy for identifying thrombotic material. Moreover, fibrous cap thickness in VH-TCFA must be measured carefully by increasing the resolution . In order to assess or have more accurate the coronary artery plaque, several imaging techniques have to be combined into a single catheter. Medical image segmentation as a essential step of identifying plaque, persist a challenge since lack of resolution, characteristics of imaging, motion, adjustable pathology and anatomy. Thus, suitable segmentation methods along with reliability, computation time reduction, higher accuracy, and full computerization will be desirable for treatment of clinical events. In VH-IVUS segmentation, unsupervised method such as IFCM algorithm can be applied for segmentation of input images in order to cluster pixels and visualize four components. Based on our experiments, performance of the clustering algorithm depends on the number of clusters that are chosen.

AKNOWLEDGEMENTS

The Universiti Teknologi Malaysia (UTM) under research grant 03H02 and Ministry of Science, Technology & Innovations Malaysia, under research grant 4S062 are hereby acknowledged for some of the facilities utilized during the course of this research work.

References

[1] K. Dulohery, A. Papavdi, M. Michalodimitrakis, and E. F. Kranioti, Evaluation of coronary stenosis with the aid of quantitative image analysis in histological cross sections., Journal of forensic and legal medicine 2012;vol. 19, no. 8, pp. 485–9, Nov.

[2] C. M. Nguyen and A. J. Levy, The mechanics of atherosclerotic plaque rupture by inclusion/matrix interfacial decohesion., Journal of biomechanics, vol. 43, no. 14, pp. 2702–8, Oct. 2010.

[3] A. Taki, A. Roodaki, S. K. Setarehdan, S. Avansari, G. Unal, and N. Navab, An IVUS image-based approach for improvement of coronary plaque characterization., Computers in biology and medicine May 2013; vol. 43, no. 4, pp. 268–80.

[4] T. S. Hatsukami, F. D. Kolodgie, J. Ohayon, R. Pettigrew, M. S. Sabatine, G. J. Tearney, S. Waxman, M. J. Domanski, P. R. Srinivas, and J. Narula, Detection of High-Risk Atherosclerotic Plaque, JCMG 2013; vol. 5, no. 9, pp. 941–955.

[5] A. Mauriello, F. Servadei, G. Biondi, E. Giacobbi, L. Anemona, E. Bonanno, and S. Casella, "Coronary calcification identifies the vulnerable patient rather than the vulnerable Plaque," Atherosclerosis 2013; pp. 6–11.

[6] W. M. Suh, A. H. Seto, R. J. P. Margey, I. Cruz-gonzalez, and I. Jang, Intravascular Detection of the Vulnerable Plaque, Circ Cardiovasc Imaging 2011;4;169-178. Available at: http://circimaging.ahajournals.org/content/4/2/169.full

[7] Reid, D.B., Watson, C. & Majumder, B.Intravascular Ultrasound: Plaque Characterization. 2012;10.1007/978-1-84882-688-5_32.

[8] A. Katouzian, E. D. Angelini, G. Carlier, J. S. Suri, N. Navab, A. F. Laine, and A. M. Background, A State-of-the-Art Review on Segmentation Algorithms in Intravascular Ultrasound, 2012; vol. 16, no. 5, pp. 823–834.

[9] H. M. Garcia-Garcia, M. a Costa, and P. W. Serruys, "Imaging of coronary atherosclerosis: intravascular ultrasound" European heart journal 2010;, vol. 31, no. 20, pp. 2456–69, Oct.

[10] M. Liang, A. Puri, and G. Devlin, The Vulnerable Plaque: the Real Villain in Acute Coronary Syndromes," 2011; pp. 123–129.

[11] X. Zhu, P. Zhang, J. Shao, Y. Cheng, Y. Zhang, and J. Bai, A snake-based method for segmentation of intravascular ultrasound images and its in vivo validation, Ultrasonics 2011;, vol. 51, no. 2, pp. 181–189.

[12] E. Uchino, K. Tokunaga, H. Tanaka, and N. Suetake, "IVUS-Based Coronary Plaque Tissue Characterization Using Weighted Multiple k -Nearest Neighbor," 2012; August.

[13] E. G. Mendizabal-ruiz, M. Rivera, and I. A. Kakadiaris, "Segmentation of the luminal border in intravascular ultrasound B-mode images using a probabilistic approach," Medical Image Analysis 2013.

[14] M. Basij, "Automatic Shadow Detection in Intra Vascular Ultrasound Images Using Adaptive Thresholding" 2012; pp. 2173–2177.

[15] G. Vachkov, E. Uchino, and S. Nakao, "Moving Window-Based Similarity Analysis and Its Application to Tissue Characterization of Coronary Arteries," 2012; vol. I.

[16] L. S. Athanasiou, P. S. Karvelis, V. D. Tsakanikas, K. A. Stefanou, K. K. Naka, L. K. Michalis, and D. I. Fotiadis, "Atherosclerotic Plaque Characterization Using Geometrical Features from Virtual Histology Intravascular Ultrasound Images," , 2010;pp. 1–4.

[17] A. A. Alsheikh-ali, G. D. Kitsios, E. M. Balk, J. Lau, and S. Ip, "Annals of Internal Medicine Review The Vulnerable Atherosclerotic Plaque: Scope of the Literature," 2010.

[18] J. D. Cascón-Pérez, J. M. de la Torre-Hernández, M. C. Ruiz-Abellón, M. Martínez-Pascual, R. Mármol-Lozano, J. López-Candel, P. Cano, C. Fernández, J. L. Ramos, M. Villegas, and F. Picó-Aracil, "Characteristics of culprit atheromatous plaques obtained in vivo by intravascular ultrasound radiofrequency analysis: results from the CULPLAC study" American heart journal Mar. 2013; vol. 165, no. 3, pp. 400–7.

[19] A. Nair, M. P. Margolis, B. D. Kuban, and D. G. Vince, "Automated coronary plaque characterisation with intravascular ultrasound backscatter: ex vivo validation,"2007; pp. 113–120.

[20] http://www.teachivus.com

[21] P. Kaur, "Novel Intuitionistic Fuzzy C-Means Clustering for Linearly and Nonlinearly Separable Data," 2012; vol. 11, no. 3, pp. 65–76.

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Cognitive approach to Computer Go programming

N. Tenys, I. Tanev, K. Shimohara
Doshisha University, Japan
(Tel: 81-080-9608-1313)
nerijus.tenys@gmail.com

Abstract: Go (Igo / Baduk / Weiqi) is a board game that is still one of the more difficult challenges for Artificial Intelligence research. While computers have defeated human champion in Chess 16 years ago, to this day best computer Go programs play only at a high amateur level. Humans play Go using experience, heuristic values, pattern recognition and whole board evaluations that computers are unable to match. Today's most successful Go programs use Monte Carlo search algorithm, but while this approach often finds good local moves on the board, due to its random nature it often fails to see the whole board position as well as human does. Our approach instead is implementation of human Go professional's "thinking" into a computer program - that is, creating a system, where computer uses algorithms and data constructs that mimic human's mind.

Keywords: Go programming, Artificial Intelligence, Game AI, Cognitive Programming

1. Introduction

Background

Go is an ancient board game invented in China ~4000 years ago, and while easy to learn, it is very deep and complicated game to play at a master level, even more so for computers.

One of the reasons why computer Go is so complicated is that it has a lot more variations than Chess. Chess has a game-tree complexity of about 10^{123} and Go ~10^{360}. For comparison of how large number that is, the number of atoms in the known Universe is about 10^{80}. Another difficulty in computer Go is the evaluation of board positions, combined with the fact that one stone placed in the beginning of a game can completely change the outcome in the end game. Humans play Go using experience, pattern recognition and whole board evaluation that computers are unable to match.

Motivation

Computer Go is an interesting subject in itself and has applications in the industry in the form of Computer Go engines, teaching programs and Go analysis tools, but even more interesting is the approach taken to solve Go problems. Outside of direct applications in Computer Go, this research is a test-bed for cognitive programming model approach that might have a lot wider application if proven successful.

Normally, game logic and similar deterministic systems are solved using specific computer algorithms and methodical approach to analyze each possible position and come to the best possible solution. When that is not possible, various heuristic methods are used to reduce the game tree complexity by removing nodes that are unlikely to lead to successful solution, like the approach widely used in Computer Chess programs - Alpha-beta pruning algorithm, which was used to reduce Chess game tree and allowed computers to beat human Chess champions, however in Go the game tree complexity is so large that, combined with the difficulty of evaluating game positions, the same method can not be applied successfully to Go programs.

Several other approaches were tried in Computer Go research, like Expert Knowledge system, which is a set of low level rules compiled by Go professionals and implemented into a system by programmers. This approach can be very successful when rules can be defined well enough to evaluate majority of possible situations for the system, however in Go, things are a lot more complicated. There are thousands of rules created throughout the centuries of Go play and analysis, but none of them are exhaustive, definite and unchangeable. Many rules are very situational and can be even contradictory in different games. That is one of the reasons why this approach wasn't very successful in Computer Go.

My approach is using Cognitive Psychology and Computer Science to create a distinct Software Development model. If tested and proven successful in Computer Go, this approach could be then applied to many other systems. While there aren't that many systems that are as complicated and open ended as Go, one particular application for this approach would be the implementation of human-like Artificial Intelligence - the merge of Cognitive Psychology and Computer Science is an ideal basis for building digital human brain model.

The long term goal of this research is understanding and implementation of human thought process into computer programs, allowing us to build digital brain models that eventually might lead to a true "Strong AI" envisioned by Computer Science pioneers more than 50 years ago.

Research Objective

Research goal is to write a Computer Go engine (software program) that is able to understand and use human-like concepts in the game of Go. Success of the research is difficult to define, however if the engine proves to be at least as strong at Go as mainstream programs using Monte Carlo approach, it could be considered a success, as it would create a unique approach to Go programming and would have a lot of unexplored potential for improvement in the future. One of the big improvements for this approach could be genetic programming, used to evolve constants and values used by the game engine.

The ultimate goal of the engine would be to one day beat the best human Go players.

2. Research Scheme

Approach

My approach is, in a way, a combination of deterministic algorithmic approach and Expert Knowledge system, but instead of defining many low level rules to determine all common game situations, I want to implement human Go professional's "thinking" into a computer program - that is, to create algorithmic Expert Knowledge system, where instead of a set of rules, computer uses algorithms and constructs that mimic human's mind. In programming terms, this is similar to adding another abstraction layer to create a program that manipulates on higher level concepts.

This kind of approach isn't used often, as, in many systems, deterministic approach is enough to define the system and its working rules, even if they sometimes are quite complex. Another reason for the lack of experimentation with similar approaches is that it requires a merging of two distinct science fields - Computer Science and Cognitive Psychology. To successfully create such system, good analysis of human mind is required as well as sufficient knowledge of Computer Science to translate human thought into digital computer program's structure and algorithms. In the past, 20-30 years ago, this approach probably wouldn't have been possible due to the lack of support from programming languages, but in recent years improvements in Computer Science, such as higher level programming languages, more advanced and intuitive Integrated Development Environments (IDEs), expansion of software Design Patterns and similar advances, have created possibilities for programs to work with more complicated and higher level concepts.

Concepts

There is a long list of human-like concepts, rules and gameplay traditions evolved through the centuries of playing Go and it would be impossible to gather and organize them all into computer program, however, I believe, just using main ideas from the plays of human professionals, combined with strengths of digital computation, can eventually achieve better results than most of contemporary Go programs, that mostly find good local moves, using Monte Carlo algorithm to search the game tree, but fail to recognize overall game strategy and often make critical long term mistakes.

Current programs keep getting better, partly because of increasing computation speeds, but also because of better heuristics being applied to evaluation of the game positions. In a way it is approaching a complex Expert Knowledge system that improves on computation based algorithm. While it is a valid approach, and someday improvements in computing speed and heuristic evaluation would reach the point where it could play as good as best human players, I believe that currently however, that possibility is quite far off.

My approach would take the Expert Knowledge system one step further and implement Go concepts like Influence, Stone frameworks, Connections, Safety, Eye potential, Sente / Gote, etc. into computer program that would use them to evaluate the board positions same way a human player does and then apply computational algorithms only to areas that are deemed important from this heuristic evaluation - similarly to how humans play.

As the evaluation and relative weights of these concepts are abstract and highly subjective, after initial implementation and balancing, system could be improved by applying genetic evolution to find best performing constants and values. It could also implement a scheme similar to Arthur Samuel's checkers program [1], which made groundbreaking work in early game AI implementations and Machine Learning, by playing games against itself (Reinforcement Learning).

Difficulties and Issues

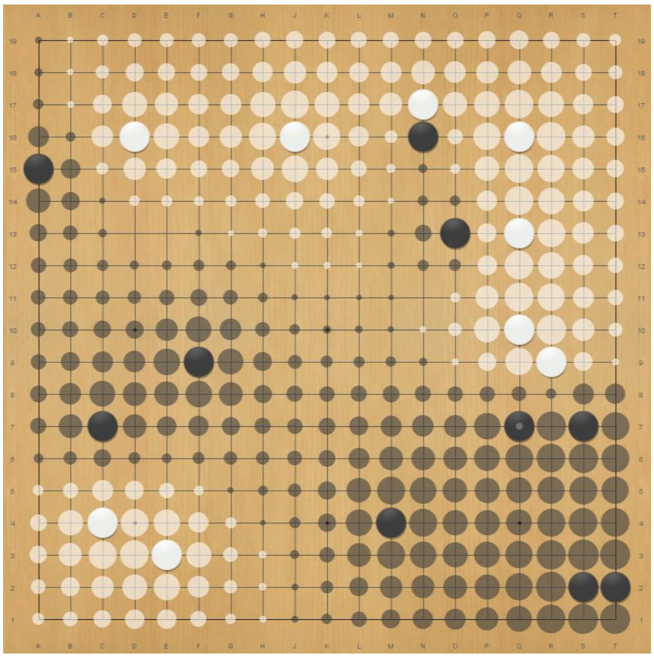
One of main difficulties of the research is difficulty to verify success of the approach until it is nearly fully completed, as Go engine can not function properly with only partially implemented features and play competitively until all the main concepts are implemented. As such, a lot of preparation and mundane work is needed for it to function properly, for example: implementing Go rules, capturing and calculating territories, eye recognition algorithms, integration with User Interface and Go Text Protocol for online play, etc.

3. Research Contents

Go Concepts and their implementation

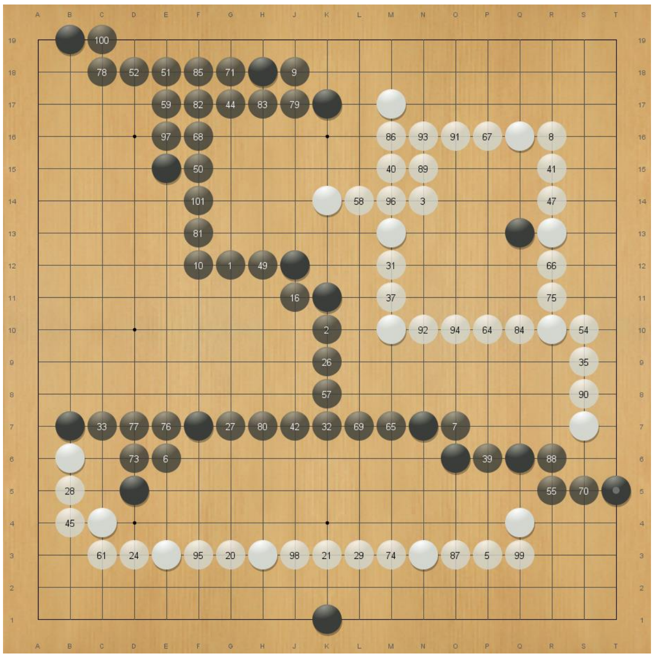
Main Go concepts and their currently planned implementations for the engine:

- *Influence*



Influence is implemented using Influence Maps, where each stone on the board "radiates" influence which degrades with distance, and each intersection on the board has certain influence value calculated from the influence of nearby stones.

- *Stone Connections*



Stone connection strength is calculated using influence values and algorithms to calculate possible connection paths and potential cutting stones.

- *Frameworks*

Frameworks are formed by connected stones that surround certain territory - it is one of key concepts in professional Go play. Frameworks have many possible values and attributes, like potential territory, safety, influence, connection strength, etc. Much of the success of play in Go depends on how successful player is in extending and protecting their frameworks and as such it should be one of most important concepts in Computer Go also.

- *Safety*

Framework and connection safety can depend on several factors, like nearby opponent's stones, supporting stones, inherent connection strength, eye making potential, etc. It is important attribute to successfully protect stone frameworks from being captured and to use opponent's safety weaknesses to threaten his frameworks.

- *Eye Estimation*

While at first glance a simple concept, it actually is very hard to implement in a computer program. To this day there does not exist an algorithm that can 100% successfully recognize and calculate the end result of a finished game between human players. Main reason for it is difficulty of eye recognition and estimation. Benson's Unconditional Life algorithm [2] can recognize unconditionally alive groups, however it can be applied only in very limited cases. My plan is to extend Benson's algorithm [2] to give a prediction of possible eyes in a framework, even if it is not 100% accurate. Like many other subjective concepts and values, it can be later evolved to find optimal point.

- *Sente / Gote*

Sente and Gote are simple concepts of move's priority or "forcing" factor of the move - Sente move is the one that requires immediate answer from opponent. Gote move is more passive, usually territorial or defensive move. Human professionals always keep these concepts in mind when making move and current Go programs are not good at recognizing that. My idea is to implement a formula that applies Sente value to potential moves and estimates possible opponent's replies.

- *Ko fights* and *Ko threats*

Sente move value ties directly into "Ko fights", where players search for a move that requires immediate answer to make use of the Ko rule. Ko evaluation is another area that is important in professional play and I feel is not given sufficient attention in computer Go.

Research Evaluation

The success of the research can be measured by having game engine play against humans or other game engines using Go Text Protocol through user interface or online server play. Most precise evaluation would be letting engine play on a live Go server, like KGS (gokgs.com), to establish a stable strength rank or participating in Computer Go competitions, like Computer Olympiad or UEC cup.

Partial measure of game engine's potential can also be estimated by a human player. Since game engine operates on abstract human-like concepts, if computer is able to come up with the moves that a good human player would make or is able to recreate joseki (standard Go opening) moves without having prior knowledge about them, it would prove that the approach was successful enough to reach human-like play ability.

4. Progress

Current Status

Current status of the project is a working version of computer Go engine that is integrated with Go UI and can be played against right now. The engine is able to recognize and manipulate concepts of Influence, stone Connections and stone Frameworks, though it is not making full use of them yet.

Next steps in the research are calculations for various framework attributes, like safety, territory, number of eyes. For that purpose work on the eye calculation is needed by using or extending Benson's Unconditional Life algorithm.

After that, the rest of the concepts can be added - like Sente / Gote and Ko fighting. Then, computational algorithms are needed to calculate specific local situations like capturing or "cutting" stones. For that purpose, even though Monte Carlo algorithm is successful in other programs, I would like to use Alpha-beta pruning, so successfully applied in Chess programs. The reason for that is that search space should be reduced dramatically by the previous heuristic evaluations to consider only areas that program deems important.

Lastly, test results can be improved using genetic evolution or Samuel's reinforcement learning [1] by setting up the Go engine to play against itself.

As mentioned previously, one of main difficulties of this research is that results won't be seen until all parts of the program are finished, but I believe there is value in successful completion of the project, even if the engine doesn't prove to be as strong as current programs, as this is completely new approach in Go programming and has a lot of potential to evolve in the future and be adapted in other areas.

5. Conclusion

It is hard to draw conclusions from current results, as the implementation of the Go engine isn't complete yet. The goal of the research is to explore new possibilities in Computer Science and Artificial Intelligence programming and it could be said, that, no matter the outcome and the strength of play of this Computer Go program, it will give insight into this distinct approach and maybe inspire new developments in similar direction in the future.

References

[1] Arthur Samuel (1959), Checkers Program, Machine Learning.
[2] David B. Benson (1976), Life in the Game of Go, Information Sciences 10, 17-29.

Circulative Narrative Generation Based on the Mutual Transformation between Narrative Conceptual Structures and Music in the Integrated Narrative Generation System

Taisuke Akimoto and Takashi Ogata

Iwate Prefectural University, 152-52 Sugo, Takizawa, Iwate, Japan
g236i001@s.iwate-pu.ac.jp

Abstract: We have proposed a new framework of narrative generation which generates narratives through the mutual transformation according to various pathways between narrative (conceptual representation) and music. This paper will continuously incorporate this framework into the integrated narrative generation system which is a comprehensive architecture for our narrative generation study and develop an experimental system combined with the integrated system. This new framework aims to apply various musical methods for composition and variation to narrative generation as a way to connect different types of media. An experiment of the mutual generation between music and narrative will show various transformation pathways or routes between music and narrative.

Keywords: integrated narrative generation system; narrative generation; music generation; structural transformation

I. INTRODUCTION

Issues on the similarity or correspondence relation between narrative and music have been discussed in musicology [1, 2, 3]. Overview, based on the studies, music itself cannot express concrete meaning as narratives. However, a narrative and a piece of music can commonly be composed of a sequence of events although a temporal sequence of events in narrative is different from a thematic sequence of events is music. We focus on the common structures.

The major purpose of our current research is developing an integrated narrative generation system to incorporate a variety of narrative generation mechanisms into an organic architecture [4]. The system consists of the following three large units: 1) Story generation mechanism generates the content of a narrative with a hierarchical conceptual structure using such knowledge units as conceptual dictionaries and story contents knowledge base. 2) Discourse generation mechanism transforms a story structure into the other conceptual structures considering how to narrate the story. 3) Expression mechanism generates representation by surface media including natural language, music and visual image.

One of the design concepts of the system is circularity or cyclicity. Although the one meaning is the ability to repeatedly generate narratives in one generation session, another meaning is the generation ability based on the mutual transformation relationship between narrative conceptual structures and the surface media expression. The transformation to a narrative conceptual structure from a surface medium, for example, will be also possible in the latter sense.

Music has been the first main object or medium to accomplish the concept. We have been studying narrative generation methods through mutual transformation according to various pathways between narrative and music based on the development of some experimental systems. The first system was presented by [5]. And the most recent system presented by [6] included eight routes of the transformation between narrative and music using automatic composition and musical variation mechanisms. However, the system was an independent experimental system in which only the relationship of narrative conceptual structures and music was simply defined.

The goal in this paper on the basis of the research is to incorporate the above experimental mutual transformation mechanism into the integrated narrative generation system. Music in this mechanism will not be just the accompaniment for a narrative. We aim to use musical knowledge & techniques and conceptual knowledge & techniques in fusion by the medium of music. Mechanisms for story & discourse and the musical expression mechanism are organically integrated together. This paper will introduce the current status of the development.

There are various systems that generate the background music for a narrative in the area of artificial intelligence (e.g., [7]). However, although narrative generation systems or computational narrative models are currently an independent study genre [8], the other researches do not exist to treat music from the view point of narrative generation or the structural correspondence between narrative and music except for a series of our study.

II. BASIC METHODS OF NARRATIVE AND MUSICAL PROCESSINGS

This section will present summaries of the integrated narrative generation system and the musical processing. Most of the foundational parts have been implemented with Common Lisp incrementally. The system is currently operating as a unified system.

1. Narrative Structures and the Processing

A story and a discourse generated in each of the mechanisms are respectively a same form of hierarchical tree structure. The basic units are an event or event concept and a relation. The former is a type of frame structure. And the latter connects two or more events, or two or more sub-structures with some events and one or more relations. An event concept is equivalent to a case structure with a verb concept and instances as values of some cases corresponded to the verb concept including agent, object, location, and so on. For example, "(event Eat (agent Boy) (object Apple) (location Park))" is the description of an event concept.

Both of structures of a story and a discourse are generated or transformed through the application of various types of techniques. We call the structural operational techniques associated with the story generation mechanism and the discourse mechanism "story techniques" and "discourse techniques" respectively. The name of "narrative techniques" is a collective term which contains story techniques, discourse techniques and moreover techniques for the generation by the surface expression media. Basically, for each of the story techniques and discourse techniques, the input (or target) is an event or a sub-structure with some events and the output is a newly expanded or transformed structure. The difference is the point that story techniques basically add a new element to the existing tree structure to expand it, while discourse ones transform a tree structure to another structure including the same event elements. A variety of types of narrative techniques can be integrated grounded on this common principle. The current version of the system has about thirteen types of story techniques and similarly ten types of discourse techniques.

2. Musical Structures and the Processing

The mutual relationship between a story structure and a discourse structure in the level of conceptual narrative are corresponded to the mutual relationship between a piece of "original music" and the "variation music" in the level of music. The basic method of the mechanism is to relate a story structure and a discourse structure to the structure of a piece of original music and a piece of variation music according to the concrete procedure explained in the following part. The mutual transformation is processed based on relational mappings.

The tree structure of a story is mapped to a musical tree structure with the similar form. Fig. 1 shows the structure of a piece of original music. Each of the internal nodes in the musical tree is equal to "primary-secondary" or "secondary-primary" relation between the child nodes. The primary side is more important musical part than the secondary side. Each of the leaf nodes, on the other hand, is corresponded to a musical event. A musical event is equal to a set of the following musical elements: one or more motifs, the harmony, and the duration defined in the number of quarter notes from 1 to 16. The musical events at the leaf nodes are sequentially played in original music. A motif is a basic musical unit consisting of a sequence of notes and has the same duration with the musical event. Each of the notes has a pitch and its duration. Each of the musical events is corresponded to an event in the story and each of the motifs is corresponded to one of the instances in the story.

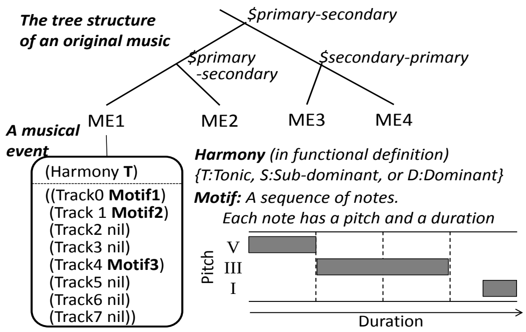


Fig. 1. The structure of a piece of original music

A piece of original music is generated through the operation of the structure and elements. The procedure of the original music generation from a story is as follows: 1) The mechanism generates motifs associated with each instance in the story. Each of the motifs is made with a sequence of notes which are selected at random. 2) It transforms each of the internal nodes, namely relations, in the story into the musical internal node using a set of rules for associating the two groups of internal nodes. 3) It decides a harmony in each musical event using a set of rules for deciding the harmony progression with a musical tree structure. 4) It transforms each event in the story into a musical event by substituting each instance to the corresponding motif.

On the other hand, the structure of a piece of variation music is also represented as a tree structure form which is transformed from the structure of a piece of original music using variation techniques. Each of the variation techniques is defined as a manipulation of the musical tree structure or the element. There are such techniques as the move and deletion of a sub-structure of a musical tree, the addition of a musical event created newly into a musical tree, and the modification of musical elements like tempo, pitch and volume.

III. A MUTUAL TRANSFORMATION MECHANISM BETWEEN NARRATIVE AND MUSIC

1. The System Configuration

We show in Fig. 2 the macro framework of the integrated narrative generation system including musical mechanisms. The musical mechanisms consist of an original music mechanism and a variation music mechanism. The former generates a piece of original music from a story, a piece of variation music, or a piece of original music. The latter, on the other hand, generates a piece of variation music according to a piece of original music, a piece of variation music, or a discourse structure. We have incorporated in addition to the above a mechanism for generating a discourse structure from a piece of variation music into the discourse mechanism. A mechanism to generate a story structure from a piece of original music is currently under development.

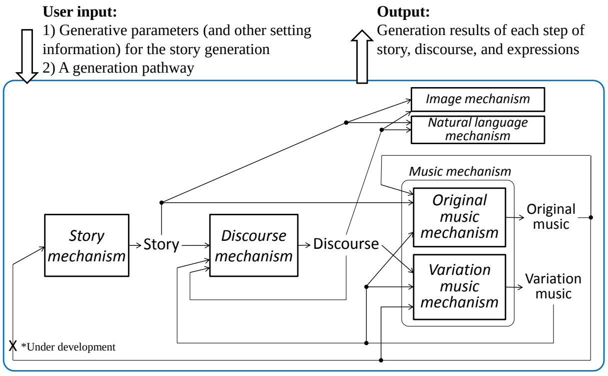


Fig.2. The expanded framework by musical mechanisms of the integrated narrative generation system

The user initially inputs generative parameters for directing the story generation. The following four parameters are prepared: "macro-structure", "length", "unreality" and "repetition". The user continuously sets the generation pathways from the next prescribed sequences: "S-D-NL", "S-D-NL-I", "S-D-VM" and "S-OM-VM-D-NL". Each of the S, D, NL, I, OM and VM mean Story, Discourse, Natural Language, Image, Original Music and Variation Music respectively.

The story mechanism is always executed firstly to generate a story by using adequate story techniques according to the generative parameters for story. Then the method sends the result by the previously executed mechanism to the next mechanism.

The user is necessary to input the following information in the middle of a continual generation process. When the discourse mechanism is executed, the user needs to set the generative parameters and the other settings for discourse generation. The user needs to select one or more variation techniques to be used for the musical variation processing. For generating a new piece of original music from an existing piece of original music, the user is required to select one or more operational techniques to change the original music structure.

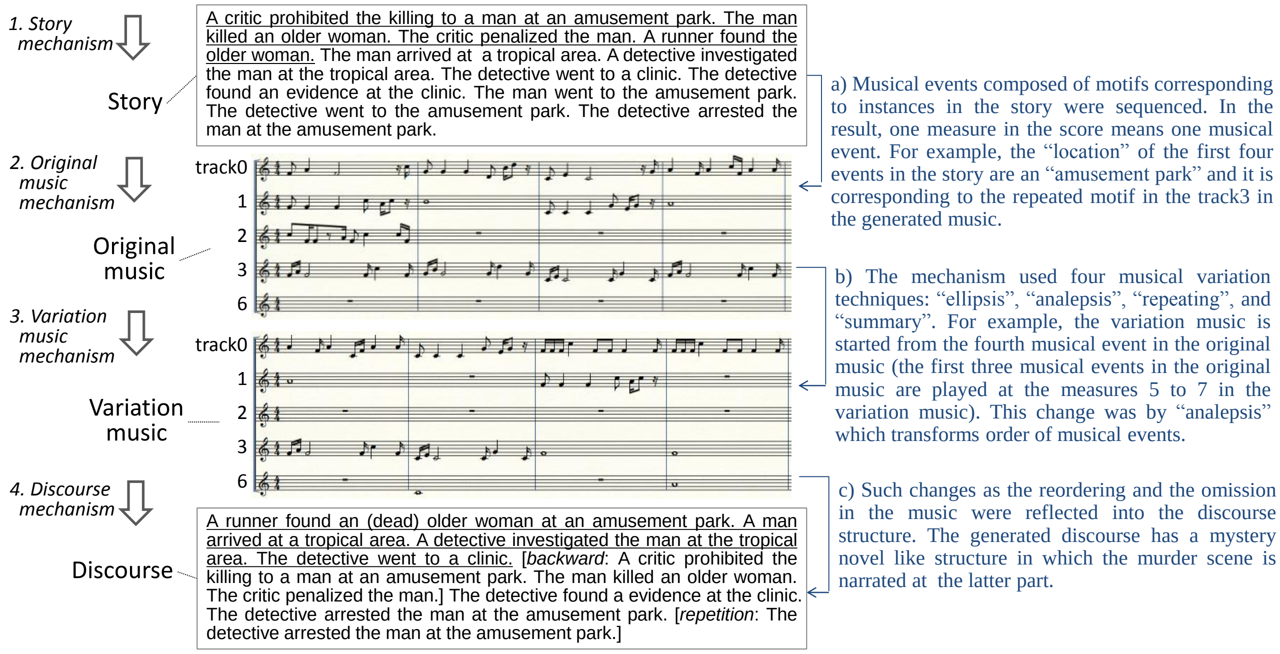
2. The Example of a Generation Flow

We will present the example of a narrative generation process including music generation by the method described above. The initial values inputted by the user are generative parameters such as "(macro-structure 2) (length 3) (unreality 1) (repetition 2)" and the generation pathway is set as "S-OM-VM-D-NL".

The result of each generation phase is shown in Fig. 3 in order. The surface representation by language of story and discourse in the figure is English translation by hand from Japanese language texts generated by a simple language generation mechanism in the integrated narrative generation system. These language texts are transformed from each of the conceptual representation forms. On the other hand, only first four musical events and the scores by hand are shown in the results of original music and variation music.

Conceptual processing and musical processing in a series of generation are mutually connected as shown in the example. To add the function of story generation from a piece of original music, which is not implemented in the current version, will contribute more free and flexible generation since we will be able to start the system's generation process from a piece of original music. It means the narrative generation from music in the extreme sense.

One of the issues to be solved in the current implementation is that the current discourse mechanism cannot transform musical structures into the corresponding discourse structure in several types of musical variation techniques, namely the modification techniques for such musical elements as tempo, pitch, and volume because of the theoretical problem on structure transformation considering the meaning in narrative and music.



*In the scores, track 4, 5, 7 are omitted because these tracks were blanks.

Fig.3. A generation flow example

V. CONCLUSION

This paper has presented the expanded framework of the integrated narrative generation system through the experimental system implementation. The important point of this framework is that musical generation and variation techniques are incorporated into the system to enable the mutual transformation between music and narrative. The previous version of the integrated narrative generation system basically generated a story structure, a discourse structure and expression representations including music in order, while introducing the mutual transformation mechanisms enabled the system to execute a variety of generation routes. Actually, we showed a generation example of the mutual transformation in which a piece of music is generated from a narrative structure and a musical variation of the piece causes a change of the narrative structure. We would like to verify the abilities in the several sides of the mutual transformation mechanisms between narrative and music after the completion.

The program of narrative structure processing has been implemented based on relatively sound theoretical foundations by a kind of knowledge-based architecture with narrative techniques and conceptual dictionaries. The background of the musical composition and variation, on the other hand, has been weak in both sides of theory and design. Thus, one of the next major topics is introducing musical theories and techniques into the part of the musical processing.

REFERENCES

[1] Walsh R (2011), The common basis of narrative and music, Storyworlds: A journal of narrative studies, 3(1): 49-71.
[2] Maus FE (1991), Music as narrative, Indiana theory review, 12: 1-34.
[3] Kramer L (1991), Musical narratology: A theoretical outline, Indiana theory review, 12: 141-162.
[4] Akimoto T & Ogata T (2012), Macro structure and basic methods in the integrated narrative generation system by introducing narratological knowledge, Proc. of 11th IEEE international conference on cognitive informatics & cognitive computing, 253-262.
[5] Kobayashi F & Ogata T (2004), Narrative and music as variation: Transformation of musical structure based on narrative discourse theory, Proc. of the 9th ISAROB, 1: 170-173.
[6] Akimoto T, Endo J, & Ogata T (2012), The expansion of paths in the mutual transformation mechanism of music and narrative, Proc. of the 11th IEEE international conference on cognitive informatics & cognitive computing, 230-239.
[7] Jewell MO, Nixon MX, & Prügel-Bennett A (2005), State-based sequencing: Directing the evolution of music, Proc. of the international computer music conference.
[8] Gervás P, Lönneker-Rodman B, Meister JC, & Peinado F (2006), Narrative models: Narratology meets artificial intelligence, Proc. of satellite workshop: Toward computational models of literary analysis, 5th international conference on language resources and evaluation, 44-51.

Dynamic Analysis of Tensegrity Systems Subject to Arbitrary Joint Constr aints

Youngsu Cho and Joono Cheong

Korea Univ., Sejong-City, S. KOREA
Tel : 82-44-860-1797; Fax : 82-44-860-1587
{youngsucho, jncheong}@korea.ac.kr

Abstract: This paper provides an approach for dynamic analysis of tensegrity systems subject to arbitrary joint constraints. Advantages of non-minimal dynamics including the simplicity of formulation and scalability are enjoyed. Due to the generality of the formulation, we can create a software module that can carry out dynamic simulation of numerous tensegrity systems. To illustrate the effectiveness of the proposed dynamics formulation, we show simulation results for a deployable tensegrity structure and a multi-stage tensegrity prism under various types of joint constraints such as pin, revolute, and planar types.

Keywords: Tensegrity, dynamics, joint constraints.

I. INTRODUCTION

A tensegrity system is one of truss structures that connect many rigid bodies via tension members such as wires and cables. Thus the tensions are the main forces that enable the structural stability and configuration [1-3]. A class k tensegrity system has as many as k bars connected at their ends by a ball joint, but the bar connections form a topological tree (no closed loops of bars)[4]. Examples of the static analysis include [4-9] among many others. On the contrary, dynamicists paid little attention to tensegrity structures, perhaps due to the apparent complexity of the structures and the equations. The papers [10,. 11] by Skelton show one of the simplest forms of the dynamics equations using the nonminimal coordinates.

Previous tensegrity dynamic analysis was performed only to those whose joint types were ball joints. However, there can be other types such as planar joints, revolute joints, or universal joints, although the ball joints are the most common types of joint pairs in the tensegrity systems. So, we will present a generic formulation for tensegrity dynamics with arbitrary joint types.

II. PRELIMINARY

NOMENCLARTURE

R = A matrix of bar's center coordinates
T = A matrix of tension vectors
W = A matrix of external forces
$\mathrm{C_b}, \mathrm{C_s}$ = Matrices of bar and cable connectivities
M = Mass matrix
K = Stiffness matrix
$\hat{\gamma}$ = tensile force density matrix
$\hat{\lambda}$ = compressive force density matrix
P = A matrix of node constraints
$\eta = [\eta_1, \eta_2]$ = A matrix of transformed coordinates

This section briefly introduces the dynamics of general tensegrity systems in a matrix differential equations using node coordinates. Reader refer to [7] for the details. Throughout this paper, the following assumptions on the bars and cables are made: (i) bars are uniform in their cross section, (ii) the moment of inertia about the bar axis is negligible (thin-bar assumption), and (iii) cable are uniform and massless.

The form of dynamics for tensegrity systems with B bars and a cables can be written by the following form

$$\ddot{N}M + NK = W + \Omega P^T$$

Where $\mathrm{N} \triangleq [\mathrm{n}_1\ n_2\ \cdots n_{n_n}] \in R^{3\times n_{n_n}}$ is the nodal coordinate matrix whose columns represent the end points of the bar; $\mathrm{W} \triangleq [w_1\ w_2\ \cdots w_{n_n}] \in R^{3\times n_{n_n}}$ is the external force matrix whose columns represent the forces at the corresponding nodes; M and K are defined as

$$\mathrm{M} = \frac{1}{12} C_B^T \hat{m} C_B + \frac{1}{4} C_R^T \hat{m} C_R$$

Where $\hat{m} = diag[m_1, m_2, \cdots, m_b]$ is the mass matrix where m_i denotes the mass of the i-th bar.

$$\mathrm{K} = \mathrm{C_S^T} \hat{\gamma} C_S - C_B^T \hat{\lambda} C_B$$

Where $\hat{\gamma} = \mathrm{diag}[\gamma_1, \gamma_2, \cdots, \gamma_a]$ is a force density matrix for cables whose components are

$$\gamma_i = \begin{cases} k_i \dfrac{\|S_i\| - \|S_i^0\|}{\|S_i\|} & if \ \|S_i\| > \|S_i^0\| \\ 0 & otherwise \end{cases}$$

Where $\|S_i\|$ is the length of the i-th cable vector, $\|S_i^0\|$ is its rest length, and k_i is the material stiffness of the i-th cable; $\hat{\lambda} = \mathrm{diag}[\lambda_1, \lambda_2, \cdots, \lambda_b]$ is the force density in the bars defined as

$$\hat{\lambda} = \frac{1}{2}\hat{l}^{-2}B^T(S\hat{\gamma}C_S - W - \Omega P^T)C_B^T - \frac{1}{12}\hat{l}^{-2}\hat{m}\dot{B}^T\dot{B}$$

III. Vector Form of Dynamics

1. Vector Notation

Dynamics using matrix form has limitation in handling constraints. Matrix form can only define constraint like pin joint type connection. When we apply tensegrity system to robots and mechanisms however, it is good to consider other joint constraints including revolute and prismatic types. So, we present a vector form of dynamics in this paper. Vector notation and dynamics can be defined as

$$\mathrm{n} = [\mathrm{n}_1^T, n_2^T, \cdots, n_n^T]^T$$
$$\mathrm{C_b} = \mathrm{kron}(\mathrm{C_B}, I_3)$$
$$\mathrm{C_s} = kron(C_S, I_3)$$
$$\mathrm{C_b} n = b$$
$$\mathrm{C_s} n = s$$
$$m\ddot{n} + kn = w + p^T\Omega$$
$$\lambda = \frac{1}{2\sum_{i=1}^{n_b} l_i^{-2}}\left(b^T C_b(C_s^T\hat{\gamma}C_s - w - p^T\Omega) - \frac{1}{6}\dot{b}^T\dot{b}\right)$$

This form is derived from the matrix form of equation in Section II.

2. Arbitrary Constraints

A. Pin-joint

If the i-th node is connected on the ground by pin-joint type, constraints condition can be written as

$$\mathrm{P_{i,x}n} = \mathrm{n_{i,x}}$$
$$\mathrm{P_{i,y}n} = \mathrm{n_{i,y}}$$
$$\mathrm{P_{i,z}n} = \mathrm{n_{i,z}}$$

If a tensegrity model is class-k type, assume there are k-1 virtual nodes and there are 3(k-1) conditions.

$$\mathrm{n_{i,1}} - n_{i,2} = 0$$
$$\mathrm{n_{i,1}} - n_{i,3} = 0$$
$$\vdots$$
$$\mathrm{n_{i,1}} - n_{i,k} = 0$$

B. Planar, prismatic joint

If the i-th node can move on the ground,

$$\mathrm{P_{i,x}n} = \mathrm{n_{i,x}}$$
$$\mathrm{P_{i,y}n} = \mathrm{n_{i,y}}$$

IV. Simulation Result

In this section, we present some simulation result to demonstrate the validity of the proposed analysis. We used a tensegrity prism shown in Fig.1. We assumed that the masses of bars are identically to be m =1[kg] and that the stiffness coefficient in the cable is uniform as k= 500[N/m]. The height of stage was set to be $1/2 + \sqrt{3}/6$ [m], and each three nodes in horizontal faces formed an equilateral triangle with the unity side length. This corresponded to the length of each bar equal to $(1+\sqrt{3})/2$[m].

Using the proposed dynamics formulation with physical parameters, we obtained the simulated response of the tensegrity prism under the external forces $[0,0,-1]^T$, applied on each node at the top face. The proposed numerical correction algorithm was applied. Figs. 2-4 show the responses of node n_3 and n_6 and tension s_4 and s_9. The correctness of the simulation results was verified by comparing with the results from a commercial package, ADAMS.

Since the formulation for the dynamic analysis was general, we could apply for any shape and scale of tensegrity systems.

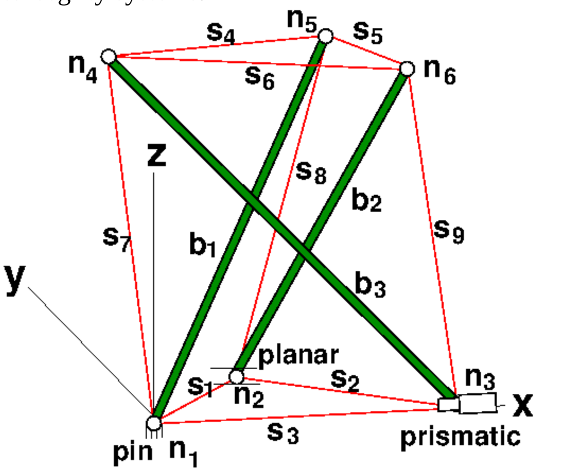


Fig.1. Tensegrity prism

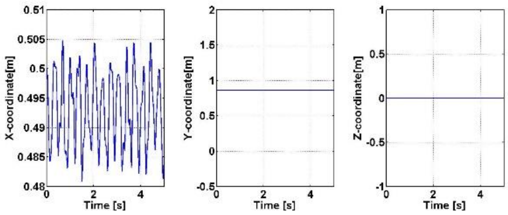


Fig.2. N3

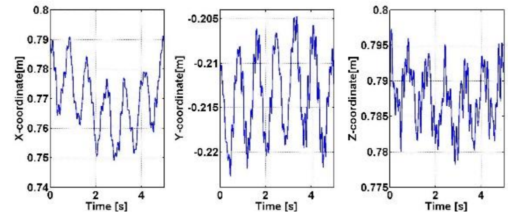
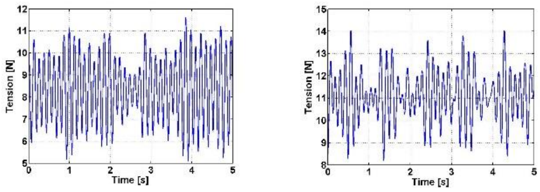


Fig.3. N6



(a) t_4 (b) t_9

Fig.4. Tension

VI. CONCLUSION

In this study, we presented a vector form of dynamics of tensegrity systems using node coordinates. The formulation in particular allowed arbitrary types of joint connections between nodes, which was impossible in the previous analysis. Due to the generality of the formulation, we can apply for any shape or scale of tensegrity systems once the necessary topology matrices are defined.

Acknowledgement

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (2012R1A1A2042823). And the Human Resources Development Program for Convergence Robot Specialists supervised by the NIPA(National IT Industry Promotion Agency, NIPA-2012-H1502-12-1002)

REFERENCES

[1] Pugh A (1976), An introduction to tensegrity. University of California Press

[2] Ingber DE (1998), The architecture of life, Scientific American 48-57

[3] Connelly R (1999) Tensegrity structures: why are they stable? (in

[4] Pellegrino S (1993) Structural computations with the singular value decomposition of the equilibrium matrix, Int. Journal of Solids and Structures, 30:3025-3035

[5] Connelly R and Terrell M (1996) Globally rigid symmetric tensegrities, Structural Topology 21:59-78

[6] Guest S (2006), The stiffness of prestressed frameworks: A unifying approach, Int. J. of Solids and Structures 43:842-854

[7] Micheletti A and Williams WO (2007), A marching procedure for form-finding for tensegrity structures, Journal of Mechanics of Materials and Structures 2:857-882

[8] Zhang JY and Ohsaki M (2007), Stability condition for tensegrity structures, Int. J. of Solids and Structures 44:3875-3886

[9] Kmet S and Kokorudova Z (2009), Non-linear closed-form computational model of cable trusses, Int. J. of Non-Linear Mechanics 44:735-744

[10] Skelton RE (2005) Dynamics and control of tensegrity systems, in: Proc. Of IUTAM symposium on Vibration Control of Nonlinear Mechanisms and Structures 309-318

[11] Skelton RE (2010), Efficient models of multi-body dynamics, in: R.Blockley and W.Shyy(ED.), Encyclopedia of Aerospace Engineering

Road map generation from Smartphone data

Masao KUBO, Chau viet DAN, Hiroshi SATO, Akira NAMATAME

National Defense Academy of Japan, Yokosuka, Japan
Tel : 81-46-841-3810; Fax : 81-46-844-5911
masaok@nda.ac.jp

Abstract: Recently, methodologies to retrieve information for achieve a safety mobility environment from driving data have been developed actively. In this paper, we propose a method to make a safety map which reveals incidents that a driver takes care. Many researches discuss the relation between a driver and its control but the driver's response against incidents is not discussed enough. The proposed method uses GPS data so our method is applicable to any smartphones. We verify our method by experiments in our commuter environment.

Keywords: smartphone, Models of Driving Behavior, safety map, collective intelligence

I. INTRODUCTION

In this paper, a method to make a safety map from GPS path recorded by smartphone which shows incidents to take care.

Traditionally, safety map used to describe spots which an accident happens. Therefore, this map is made by mainly government and public administration, for example, Kanagawa pref. [1], Aichi pref. [2] and etc..

Recently, people try to share their own driving experiences to achieve a more safety mobility environment. By analyzing their driving data collected by social network services, it is possible for the community members to share their each aspects about their mobility environment, for example, Nakajima [3], Hilton [4], Honda[5].

In this paper, we propose a method to make a map which reveals incidents that a driver should take care from GPS data recorded by smartphones.

II. APPROACH

In this chapter, we propose a method to make an occupancy grid map of incidents. In this paper, we call a geographical factor which forces a driver pays more attention an incident. GPS path is mainly used for because the pose of smartphone is usually unknown beforehand. Each episode of the data is not accurate by noise for example, included GPS signal. We expect the proposed method can extract a plausible map from a large collection of the episodes.

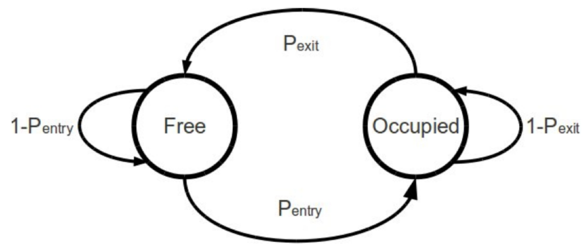


Fig.1. state transition model of a grid of IMAC.

Independent Markov Chain Occupancy Grid Map(IMAC) by Saarinen[6] is adopted for our purpose. IMAC is a kind of occupancy grid map (Thrun[7]) . IMAC consists of grids which record information around. Basically this map is adequate for representing existence of static obstacles. However, normally there is no static obstacles on a road while cars and pedestrians should be dynamic obstacles, and traffic signals and stop lines should be consider as semi-static obstacles. We expect that IMAC is very adequate for representation of semi-dynamic obstacles because state of each grid of IMAC is switched as a 2 states Markov model.

In this section, we explain IMAC first. Next, the algorithm to make a map of incident is proposed.

1. Independent Markov Chain Occupancy Grid Map(IMAC)

Independent Markov Chain Occupancy Grid Map(IMAC) is an occupancy grid map based on Luber [8]. Each grid has one of the states, *free* or *occupied*. The state transition is described as Fig.1. Each transition is follows as Poisson process.

$$P = \begin{pmatrix} 1-P_{entry} & P_{entry} \\ P_{exit} & 1-P_{exit} \end{pmatrix} = \begin{pmatrix} 1-\lambda_{entry} & \lambda_{entry} \\ \lambda_{exit} & 1-\lambda_{exit} \end{pmatrix}$$

Let suppose m is a grid and the average transition probability from occupied to free and the average transition probability from free to occupied are $\lambda_{m,exit}$, $\lambda_{m,entry}$ respectively.

$$\hat{\lambda}_{m,exit} = \Pr ob(m = free \mid m = occupied)$$

$$\hat{\lambda}_{m,entry} = \Pr ob(m = occuppied \mid m = free)$$

Saarinen[6] proposes the following estimation method for these transition probabilities. After m is observed in a sufficient number of times, these probabilities are estimated as follows.

$$\hat{\lambda}_{m,exit} = \frac{\#events : occ.to\ free + 1}{\#observations\ when\ occ. + 1}$$

$$\hat{\lambda}_{m,entry} = \frac{\#events : free\ \ to\ occ. + 1}{\#observations\ when\ \ free + 1}$$

where #observations when occ. means the number of time when m is occupied. #events:free to occ. indicates the number of times when m is switched from free to occupied.

Saarinen[6] showed that this rule works well by experiment in warehouse.

2. The proposed rules for making incident map on road environment.

In this section, we explain the method to make a map of IMAC from GPS path by estimation of status of neighbor grids.

We suppose that a driver of vehicle takes an action to reduce the risk to make an accident (for example, Masuda [9]). We propose 3 rules divided into 2 categories which estimate status of grids around its vehicle.

Fig.2 shows the first rule to estimate grid's status which is used for when a vehicle is moving forward. Its driving environment is discretized as grids. First vehicle speed at time t V(t) is calculated from a pair of positions of GPS at time t P(t) and P(t-1). If the V(t) and Azimuth(t) are stable, (Rule1) the number of times of free of all grids inside Sector(t) is added 1.

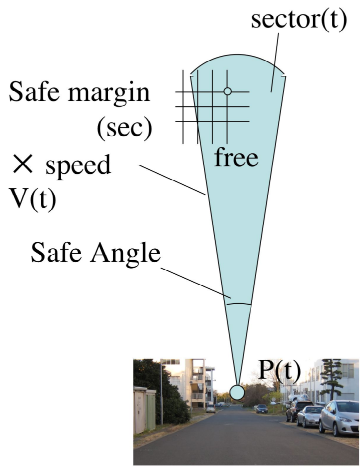


Fig.2. the update rule when a vehicle goes forward.

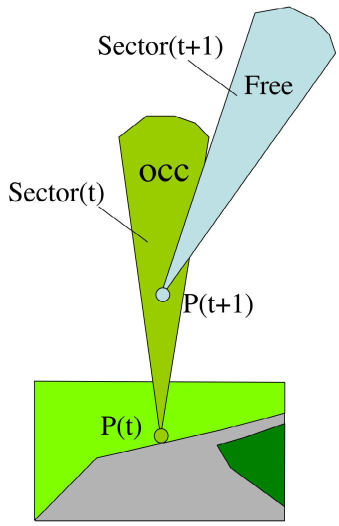


Fig.3. the update rule when a vehicle turns.

Fig.3 shows the second rule. When a vehicle is turning, we suppose that its driver tries to avoid obstacles in Sector(t). Therefore, the number of times when each grid is in Occupied added 1.

(Rule2) $sector(t) \cap \overline{sector(t+1)}$

Also, we suppose that there should be obstacles when a vehicle is moving slower.

(Rule3) $V(t-1) > V(t)$

III. Experiments

1. Test in "sandbox"

A. Test environment

We verify the basic performance of our method by a well organized experiment in the closed area. Vehicle's behavior includes GPS, accelerations, and azimuth, etc is recorded in 100 Hz by iPhone by Apple which is one of popular smartphones. We drive about 90 minutes around one of our academy's buildings clockwise and about 60 episodes (=60 laps) and 527640 records are obtained. During this experiment, we met only 1 pedestrian and no cars. Fig.4 shows the test environment and the episodes over the Google map. In the center of this figure, there is the building and the lines around it show the GPS path. The road around the building has a single lane road on each side. The driver attempted to drive through a same location while the paths recorded are seriously scattered. Therefore this data of GPS path contains a lot of noise.

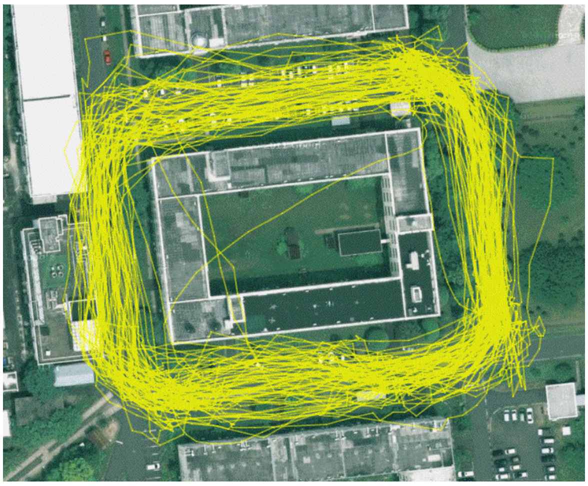


Fig.4. The experiment in the closed installation

B. Result of the closed area.

Fig.5 shows the map obtained. The diameter of green circle at each grid indicates $\lambda_{m,entry}$ that of the red circle indicates $\lambda_{m,exit}$.. The area around the upper right corner shows the performance of our method. During turning the corner green grids are impressive. There are larger red circles at around where the vehicle starts to accelerate at the end of the curve. We suppose that these characteristics of the obtained map well consistent with our expectations. We can say our method can make incident map under noisy environment if there are no other vehicles.

2.Road test

Next, our method is evaluated by public road environment. We drove about 20 times during from June to July 2013 around our academy (mabori, yokosuka, kanagara prefecture, Japan). Fig.6 shows the path and the incident map obtained. The driving route is as follows. We start our academy (the right of A)at about 83 m height above sea level and go down to D at about 1 m height above sea level . Then we turn right at D and go to the end of the right. We go to the end of the left hand side and return to the academy. We seldom meet a traffic congestion while we frequently stop at a red light of traffic signals.

For simplicity we show only the result of $\lambda_{m,exit}$ in Fig.6. A filled circle at each grid obtained indicates the value of $\lambda_{m,exit}$. The green shows the high, the red is about middle and the blue represents a value of nearly zero.

Simply speaking, markers where we release the brake pedal after a large and/or long speed down tend to be green. For example, this map shows the 6 grids labeled by alphabets are green. There is a zebra crossing without traffic signal at A and B. C and D are grids just beside of traffic signal on a national road. E is the spot where 2 lanes merge 1 lane. At these all spot, we have to stop or take the corresponding careful driving.

Also, grids where we drive slowly become to be blue. The area of the end of the right is a residential area and we have to drive very carefully. Also the rectangle shape path of the left side of "F" indicates roads in a residential district.

By this experiment, we can say our method can make an incident map at least if there is no congestion.

VI. CONCLUSION

In this paper, we propose a method to make a safety map that reveals incidents which a driver takes care from GPS path recorded by smartphone. By the 2 simple but well conducted experiments, we showed our method can make the map adequately.

REFERENCES

[1]Kanagawa prefecture, http://www.police. pref. kanagawa. jp/mes/mesf0152.htm

[2]Aichi prefecture, http://www. pref.aichi. jp/cmsfiles/ contents/0000041/41119/sitara-jikomap.pdf

[3]Nakajima Y,Makimura K,Masuko T, Applications for Road and Urban Transport planning administration using data on driver reaction in unforeseen circumstances by probe vehicle,IBS Annual Report2005,pp81-86, 2005

[4] Hilton B.N., Horan T.A., Burkhard R., Schooley B.,SafeRoadMaps: Communication of Location and Density of Traffic Fatalities through Spatial Visualization and Heat Map Analysis Information Visualization.

[5] Honda, http://safetymap.jp/

[6]Saarinen1 J., Andreasson H.,. Lilienthal A.J., Independent Markov chain occupancy grid maps for representation of dynamic environment, 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.3489 - 3495,2012

[7] Thrun et.al., Probabilistic robotics, MIT press, 2005
[8] Luber M, Diego G., Tipaldi, and Arras K.O.. Place-dependent people tracking. The International Journal of Robotics Research, 30(3):280, 2011.
[9]Masuda, Haga, Models of Driving Behavior and Accident Prevention, 2009, Reliability Engineering Association of Japan(REAJ), Vol.31,No.3,pp223-228

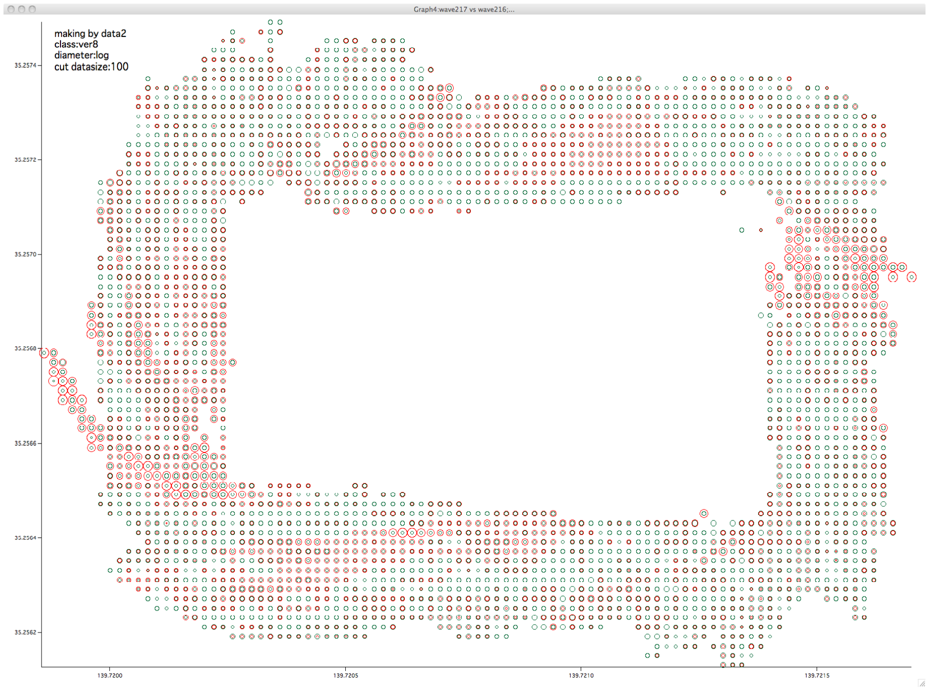


Fig.5. The incident map for the closed area.

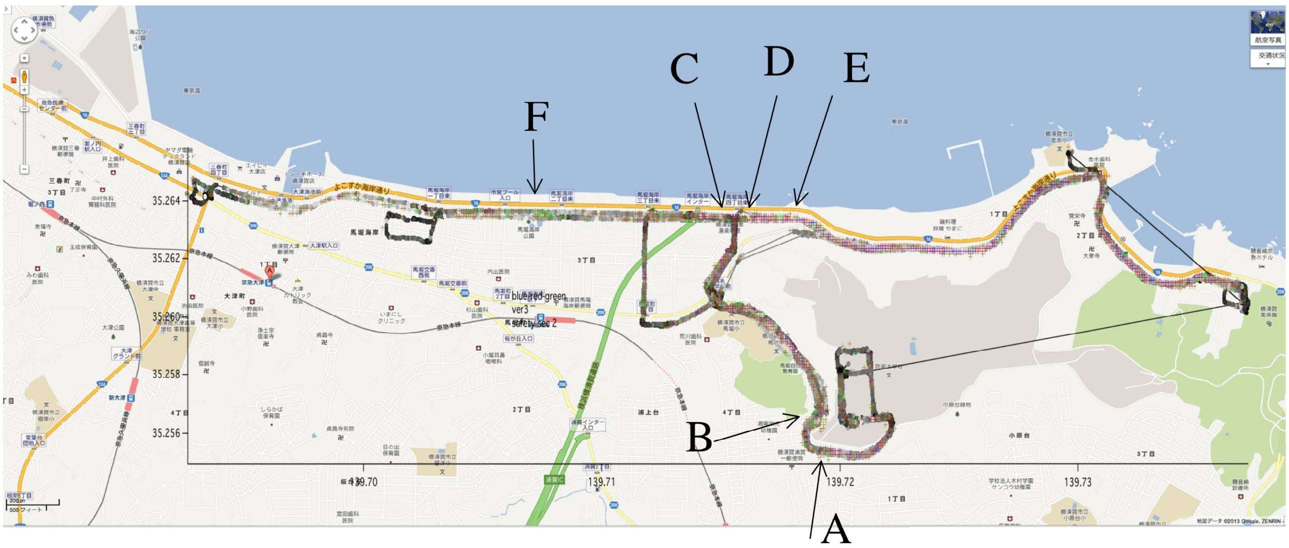


Fig.6. The incident map for the commuter area of our academy.

Zigzag-Perceptually Important Points for Financial Time Series Indexing

Chawalsak Phetchanchai[1], Ali Selamat[2] , Md Hafiz Selamat[3]

K-Economy Research Alliance & Faculty of Computing,
University Teknologi Malaysia
81300 Skudai, Johor
[1]chawalsak_phe@dusit.ac.th, [2]aselamat@utm.my, [3]Md Hafiz Selamat

Abstract: In this paper, we propose a zigzag based m-ary tree (ZM-Tree) to index financial time series. The index is done by mean of zigzag-perceptually important points (ZIPs). The tree represents the zigzag feature of financial time series which benefits in reversal pattern searching, and wave counting. We also propose a method of identifying ZIPs in multi-resolution. The identification process uses the vertical distance as a measurement to evaluate the point importance. The experiments demonstrate the performance of ZM-Tree in tree building, tree retrieval, tree pruning, and dimensionality reduction.

Keywords: Financial time series analysis, time series indexing, zigzag-perceptually important points.

1 INTRODUCTION

A time series is a sequence of observed data which are generated chronologically. A financial time series is a time series that generated in financial activities; including stock prices, exchange rates, fund prices, and dollar indexes etc. Recently, there are varieties of financial time series research areas, for examples, pattern searching [1, 8], segmentation [14], dimensionality reduction [13], clustering [12], classification [25], and prediction [15]. Due to time series are usually large of size, unstructured and high dimensionality; the way of locating time series of interest is a non-trivial problem. Thus, indexing those time series is one of the possible solutions for increasing processing performance. A financial time series illustrates its own specific characteristics comparing to other time series data such as crime time series data [26]. The illustration of a financial time series shape is typically characterized by a few important points within multi-resolution view. These important points are always formed in a zigzag direction. A zigzag direction movement can be considered as a zigzag feature which is collected by filtering out random noise [16]. A zigzag feature demonstrates past performance trends and only the most important changes. Most of using a zigzag feature in technical analysis is related to reversal patterns identification, price retracement measures, and corrective waves counting [27]. In this paper, we propose a financial time series indexing approach. Indeed, we also propose a method of identifying zigzag important points in multi-resolution so-called Zigzag-Perceptually Important Points (ZIP) identification method. Further, these zigzag important points are indexed by using our proposed Zigzag based M-ary Tree (ZM-Tree). The rest of this paper is organized as follows. Section 2 discusses in the previous related works. Section 3 describes the proposed ZIP identification method and the indexing method. Section 4 represents the experiments. The last section, section 5 discusses the conclusion and recommendation for future works.

2 RELATED WORKS

2.1 Time Series Indexing

Indexing a time series is necessary in several researches of similarity searching. Faloutsos et al. [2] proposed a fast subsequence matching technique that allows the similarity search in between different size-time series. They extracted a time series feature and divided it into several subsequences by using sliding window. Each subsequence is represented by a Minimum Bounding Rectangle (MBR) which is further indexed and stored using ST-index. Lin et al. [17] proposed a new method based on Piecewise Aggregate Approximation (PAA) [18, 19] so-called Symbolic Aggregate Approximation (SAX). SAX allows a time series of arbitrary length s to be reduced to equi-length segmented time series p by using PAA, where s >> p, and then symbolize the PAA representation into a discrete string. SAX was later improved as indexable SAX or iSAX and indexed them for mining terabyte sized time series [20]. Roh and Yi [21] introduced a new class of time series sequences where each observation is represented by an interval rather than a number. They then efficiently indexed these intervals for fast retrieval of similar interval time sequences from large databases. However, above approaches of indexing focus on dividing a time series into several subsequences without interesting in important points which are represented in financial time series.

Important points are very important and non-trivial in financial time series analysis. Chung et al.[6] introduced the perceptually important points (PIPs) identification method for using in patterns matching of financial applications. Later, Fu et al. [7] and Fu et al. [9] indexed the identified PIPs individually. Fu et al. [7] introduced the indexing method of financial time series based on low resolution clustering. PIPs are identified in multi-resolution and they are grouped together by mean of its shape by using a clustering method such as K-means clustering. The index is produced during the clustering process. More structurally, Fu et al.[9] proposed the method to index PIPs by means of the tree data structure so-called specialized binary tree (SB-Tree) with the benefits of efficient computation of cumulative new data points, retrievals, and access. Unfortunately, the representation of important points in a time series is not in the zigzag form.

Perng et al.[22] proposed a Landmark Model which relies on human spatial memory. Landmark Model produces a zigzag compressed time series. The model collects the greatest important points in a time series with a specific minimal distance/percentage principle (MD/PP). MD/PP has its own intuitive meaning. For example, if a stock trader trades once for a week (5 business days) and regards a 10% gain or loss as significant, then MD is set to 5, and PP is set to 10. Landmark Model represents a time series which is filtered out the noise and retained only the significant important points underlying the specific MD/PP. For querying the Landmark sequence, they recommended that the Landmark sequence is more similar to a string sequence rather than a multidimensional sequence, thus the string indexing is more suitable than using R-tree. The similar concepts of the important points identification can be found in studies done by Pratt and Fink [24] and Fink and Pratt [23]. They compressed a time series by selecting some of minima and maxima or important points of peaks and troughs, and dropping the other points. The intuitive idea is to discard minor fluctuations and keep major minima and maxima. The compression rate is controlled with a parameter R, which is always greater than one; an increase of R leads to selection of fewer points. They gave a technique for indexing of compressed series, which supports retrieval of series similar to a given pattern by considering their major inclines, which are upward and downward segments of the series. However these techniques do not illustrate data structure to well-organize and store the identified important points.

2.2 Perceptually Important Points Identification

The concept of PIP identification [6] is based on the data point importance. The importance is defined by the domination of a data point on the shape of the time series. A data point that has a greater domination to overall series shape is thought as more importance. For a given time series T, all the data points, $t_1, t_2, \ldots, t_m$ in T will go through the PIP identification process. Initially, the first two PIPs are collected from the first and the last points of T. The next PIP will be the point in T with the greatest distance to the first two PIPs. The fourth PIP will then be the point in T with the greatest distance to its two adjacent PIPs, either between the first and second PIPs or between the second and the last PIPs. The process of locating the PIPs continues until all the points in T are attached to a list. To calculate the distance to the two adjacent PIPs, three data point importance evaluation methods are proposed to measure this distance. Fu et al. [9] introduced three methods for evaluating the importance of the PIPs in a time series, they are: Euclidean distance (PIP-ED), perpendicular distance (PIP-PD) and vertical distance (PIP-VD) as shown in fig. 1.

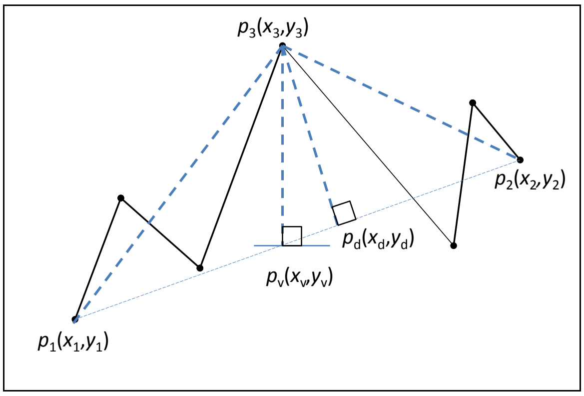


Fig. 1. Measurement of PIP-ED, PIP-PD, and PIP-VD.

PIP-ED calculates the sum of the Euclidean distance of the test point to its adjacent important points. PIP-PD calculates the perpendicular distance between the test point and the line connecting the two adjacent PIPs. Finally, PIP-VD calculates the vertical distance between the test point and the line connecting the two adjacent PIPs.

3 THE PROPOSED METHOD

In this section, we propose the method of zigzag-perceptually important point identification method, the algorithm of constructing the ZM-Tree from ZIPs and also the algorithm of accessing, updating, and retrieving ZIPs from ZM-Tree.

3.1 Zigzag-Perceptually Important Point Identification Method

For a time series $T = \{t_1, t_2, \ldots, t_m\}$; where m is a number of time instances, the main idea of ZIP algorithm is to aim to collect the important points that influence the zigzag shape of the time series. These important points are collected from a time series by evaluating the distance of the point to its adjacent important points. The selected data point importance evaluation is based on the vertical distance (VD) [9]. However, in our case, we make some modification of VD measurement for utilizing evaluating their zigzag states by removing the absolute function according to retain the sign (+ or −) of the VD for later evaluating their zigzag turning signals (ZTS).

The modified function depicts as follow:

$$VD(p_3, p_v) = y_v - y_3 = \left(y_1 + (y_2 - y_1)\frac{x_v - x_1}{x_2 - x_1}\right) - y_3 \tag{1}$$

where (x_1,y_1) and (x_2,y_2) are coordinates of points at the start and end of a segment or p_1 and p_2 respectively, $p_3(x_3, y_3)$ is a point which we try to find its VD comparing to p_1 and p_2 . Finally, $p_v(x_v,y_v)$, a projection of p_3 on the line connecting between p_1 and p_2. The algorithm of identifying ZIPs is illustrated as follows.

***Algorithm** ZIP_Identification*(T)

```
 1.    Choose the first and the last point of the time series T; p1, pm as the first two ZIPs, add to S;
 2.    Z ← Find the next ZIP
 3.   IF Z is found THEN
 4.         Determine ZTS of the Z;
 5.         Add the ZIP and its Z to S;
 6.         Loop_cont=TRUE
 7.         WHILE Loop_cont
 8.                  C ← Collect all available segments in T
 9.                  IF C is not empty THEN
10.                       sm ← Count a number of segments in C
11.                       FOR i = 1:sm:i=i+1
12.                         seg ← Read seg i
13.               Remove seg i from C
14.                           IF seg is not marked as 'N' THEN
15.                    Z1 ← Find the ZIP of seg
16.                    IF Z1 is found THEN
17.                           Determine ZTS of Z1,
18.   IF ZTS of Z1 equals to one of its adjacent ZIP THEN
19.                                  Z2 ← Find the ZIP of this sub segment of seg
20.                                           IF Z2 is found THEN
21.                                  Determine ZTS of Z2
22.                                                IF ZTS of Z2 equals to one of its adjacent ZIP THEN
23.                                         Delete Z2, Z1
24.                                         Mark seg as 'N'
25.                    ELSE
26.                    Add Z1, Z2 and their ZTS to S sequentially
27.                                    END IF
28.                              ELSE
29.                                              Delete Z1
30.                                    Mark seg as 'N'
31.                              END IF
32.                  ELSE
33.                                    Add  Z1 to S
34.                              Mark seg as 'N'
35.                             END IF
36.                     END IF
37.            END FOR
38.       ELSE
39.            Loop_cont  = FALSE
40.       END IF
41.  END WHILE
42.   END IF
END_ZIP_Identification
```

Algorithm 1. ZIP identification process

3.2 An example of ZIP identification

For representing of the ZIP identification process, we illustrate step-by-step of the method graphically. The empirical example is a synthetic data sequence with 17 data points. Algorithm 1 depicts the steps of ZIP identification from the first step until the last step. Table 1 shows the ZIP list ordered by their importance level collected from the ZIP identification process of algorithm 1.

4 TIME SERIES INDEXING

In this section, we describe the indexing technique of financial time series based on ZIP identification. The ZIPs which collected from the ZIP identification method are used to construct the tree data structure. The constructed tree is a variety of m-ary tree; so-called Zigzag based M-ary Tree (ZM-Tree). ZM-Tree follows the M-ary Tree constraints with some additional constraints. Section 4.1, we describe the structure of ZM-Tree. Section 4.2, we introduce the ZM-Tree building algorithm. Finally, section 4.3, we propose some basic tree operations.

4.1 ZM-Tree data structure

The proposed ZM-tree is a variety of the incomplete m-ary tree with m = 3 where m is a number of available children for each node. ZM-tree is an ordered tree in which:

(1) The nodes hold between 1 to 2 distinct keys.
(2) The keys in each node are sorted.
(3) A node with k elements has k+1 sub trees, where the sub trees may be empty.
(4) The ith sub tree of a node $[v_1, ..., v_k]$, $0 < i < k$, may hold only values v in the range $v_i < v < v_{i+1}$ (v_0 is assumed to equal - infinity, and v_{k+1} is assumed to equal + infinity).

The additional constraints and information to the m-ary tree are added to ZM-Tree. The root contains exactly two keys of the ZIPs from level zero important points. And, each other node has one or two key(s) depending on a number of retrieved ZIPs on that iteration level. Information storing in each key comprise of ZIP, and ZTS (placed over the key block).

4.2. A node structure and contents

There are three available types of node in ZM-Tree;, a single key node, a double keys node, and a root node. Firstly, a single key node is a node that contains only one key while holds available two children at left and right sub-tree. A double keys node is a node that contains two ordered keys. It holds available three children at left, middle and right sub-tree. Lastly, a root node, it is a special kind of a double key node but it holds only available one child at the middle key while the left and the right keys are set to null. For considering which node is a single key node or a double keys node, this depends on the collected ZIPs of a sub-segment at the specific level.

Each node contains a key or keys depending on the kind of that node. However, each key also attaches some necessary information, for examples, key's ZTS, and its VD. ZTS is placed on the top of the correspondent key of each node except the root node which there is no ZTS. VD is stored in a key structure moreover a key value. VD is necessary for tree pruning or tree retrieval activities.

4.3. Tree Building

To build a ZM-Tree, firstly all ZIPs are needed to collect and kept in a ZIPs list sequentially. The root of the tree is created by reading the first two ZIPs from the ZIPs list and storing them in the root node. Further, all other ZIPs are read and put to the appropriated nodes and positions. The algorithm of creating a ZM-Tree is illustrated as follows:

```
Algorithm Insert( root, newKey)
cur ← root of the tree ;
prev ← null;
IF (root   is null)   THEN
      root   ← create a new TreeNode(newKey);
ELSE
      pos ← seek_position(root, newKey);
IF( pos is the tree root) THEN
                  IF (newKey.zip < pos.key(1).zip) THEN
            pos.key(1) ← pos.key(1);
            pos.key(2) ← newKey;
                  ELSE
            pos.key(2) ← newKey;
      END IF
   ELSE
      IF (pos.key(2) is null) THEN
            IF  (pos.key(2).level  equals  to  newKey.level) THEN
                  IF (newKey.zip < pos.key(1).zip) THEN
                        pos.key(2) ← pos.key(1);
                        pos.key(1) ← newKey;
                  ELSE
                        pos.key(2) ← newKey;
                  END IF
            ELSE
                  IF (newKey.zip < pos.key(1).zip) THEN
                   pos.left ← new TreeNode(newKey);
                  ELSE
                   pos.right ← new TreeNode(newKey);
                  END IF
               END IF
      ELSE
               IF (newKey.zip < pos.key(1).zip) THEN
                  pos.left ← new TreeNode(newKey);
               ELSE  IF  (newKey.zip  <  pos.key(2).zip) THEN
                  pos.middle ← new TreeNode(newKey);
               ELSE
                  pos.right ← new TreeNode(newKey);
               END IF
      END IF
    END IF
END IF
RETURN root;
End Insert
```

Algorithm 2. Insert a new key to the ZM-Tree

```
Algorithm Seek_position( root, newKey)
cur ← root ;
prev ← null;
REPEAT WHILE cur < > null
BEGIN
      prev ← cur;
    IF (cur is root) THEN
          IF (cur.key(1) is null) THEN
                      cur ← cur.key(1);
          ELSE IF (cur.key(2) is null) THEN
             cur ← cur.key(2);
          ELSE
             cur ← cur.middle;
          END IF
    ELSE
          IF (cur.key(2) is null) THEN
     IF (cur.key(2).level equals to newKey.level) THEN
                cur ← cur.key(2);
              ELSE
                IF (newKey.zip < cur.key(1).zip) THEN
                   cur ← cur.left;
                 ELSE
                   cur ← cur.right;
                 END IF
             END IF
          ELSE
             IF   (newKey.zip   <   cur.key(1).zip)   THEN
                cur ← cur.left;
             ELSE IF (newKey.zip < cur.key(2).zip) THEN
                cur ← cur.middle;
             ELSE
                cur ← cur.right;
               END IF
          END IF
    END IF
END
RETURN prev;
End Seek_position
```

Algorithm 3. Seek a position in a ZM-Tree for inserting a new key

In order to create a ZM-tree, we start with creating the root of the tree. The root node contains two elements. Theoretically, in the case, the root must hold three sub trees, but however, the left and the right sub trees are set to be null due to the node containing the elements which their values larger than the last point in time series and the node containing the elements which their values smaller than the first point in the time series are not available. The remaining sub tree is the middle sub tree between two elements of the root. Children nodes may comprise of one or two elements ascending ordered from the left to the right. The creation of a tree can be done recursively. The next importance level is considered as the children of the current importance level. The conclusion of the recursion of creating the ZM-tree can be illustrated in algorithm 3, and an example of the ZM-Tree from ZIPs of table 1 is demonstrated in fig. 2.

Table 1. ZIP list from ZIP identification process.

level	ZIP	y	VD	ZTS
0	1	0.50	n/a	n/a
0	17	0.00	n/a	n/a
1	9	1.00	0.75	–
2	6	0.21	-0.60	+
2	15	0.53	0.28	–
2	13	0.41	-0.28	+
3	2	0.18	0.26	+
3	3	0.46	-0.27	–
3	7	0.80	0.33	–
3	8	0.63	-0.27	+
3	10	0.86	0.16	–
3	11	0.71	-0.22	+
4	4	0.25	-0.13	+
4	5	0.39	0.16	–

4.4 Basic Tree operations

a) ***Retrieval***

To retrieve a time series from the ZM-tree, data points are retrieved according to their importance. The retrieval process starts from the root to their leaves recursively. The steps of retrieval process depicts as follow:

- Firstly, the first two important points which stored in the root are retrieved and put in the sorted heap.
- The children nodes are determined their important levels and VDs. Children nodes are retrieved ordered from left to right. Hence, nodes comprise of one or two keys, in this case, all keys in a node must be retrieved at the same time.
- The retrieval can be done recursively until the maximum number of required ZIPs is reached or until the last node is retrieved.

b) *Pruning*

In this study, we propose a method of financial time series dimensionality reduction based on ZIPs retrieval from ZM-Tree. Our method is based on the idea that the time series always represent their shape in a zigzag movement. All points between the two important points which are closely laid on the strait line between these two points are removed according to their less importance. Additionally, the pruning technique proposed in [6] is adapted for this study. Firstly, we attempt to remove the less significant points of the time series by considering their VD. The nodes with at least one element have smaller absolute VD values than the threshold γ are removed. The meaning of this condition is that the corresponding nodes and their ascendants have less fluctuation. However, the threshold γ can be appropriated assigned by users.

5 EXPERIMENTAL RESULTS

In this section, we evaluate our methods of ZIP identification, indexing, and time series dimensionality reduction. The evaluation experiments are implemented with Java programming language performing on IBM PC compatible (Windows XP Professional, CPU Intel Pentium IV 1.7 GHz, RAM 512 MB). Our dataset is the daily close price stock index time series of Stock Exchange of Thailand (SET) collected during Jan 4, 1999 and July 31, 2006. Measuring of the indexing performance, we experiment on the dataset by measuring the CPU cost against the number of ZIPs. The varieties of number of ZIPs to build the tree tend to show that the increasing of the number of ZIPs causes the increasing of time to build the tree linearly, as can be seen in figure 3.

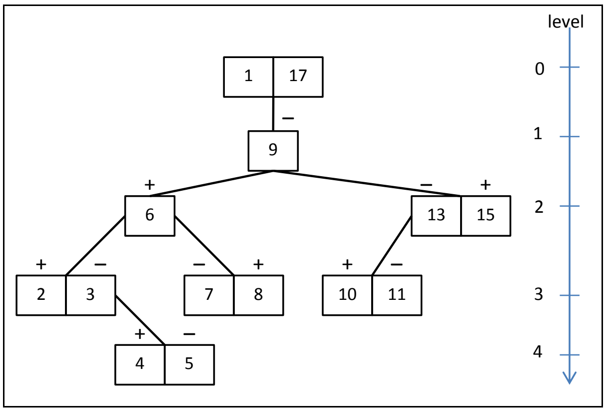


Fig. 2. ZM-tree data structure of retrieved important points in Table 1 .

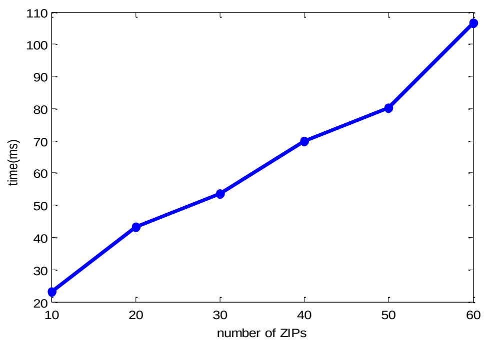


Fig. 3. CPU cost for building the index against number of ZIPs.

For the tree accessing, we measure two factors; one is the time usage against the number of ZIPs retrievals, and another one is the time usage against the tree-height level retrievals. Figure 4 demonstrates the CPU cost (time) against the number of retrieved ZIPs. The increasing of number of retrieved ZIPs causes the increasing of CPU cost linearly. While the figure 5 shows the CPU cost against the tree-height level retrievals in the same way.

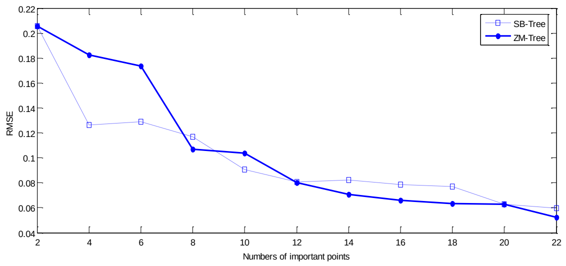


Fig. 4. Comparison of dimensionality reduction errors between ZM-Tree and SB-Tree.

Finally, we experiment our study on the task of dimensionality reduction using the tree pruning approach. The comparison of errors of the dimensionality reduction found that the cost of dimensionality reduction in ZM-Tree and SB-Tree approaches are very high when using a very few number of important points and the errors are dropped quickly when the number of important points are increased. The graph shows slightly decline when the numbers of important points are more than twelve. This can be seen in fig. 7, and the examples of reconstructed time series of the retrieved important points comparing to their original time series are illustrated in figure 8a for ZM-Tree, and figure 8b for SB-Tree.

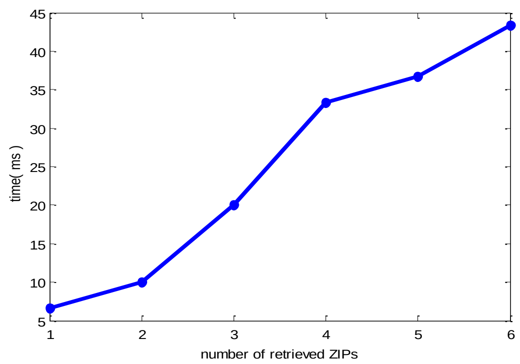


Fig. 4. CPU cost of tree retrieval against numbers of retrieved ZIPs.

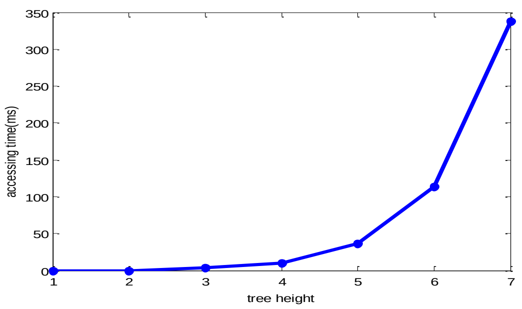


Fig. 5. CPU cost of the tree retrievals against the tree-height level.

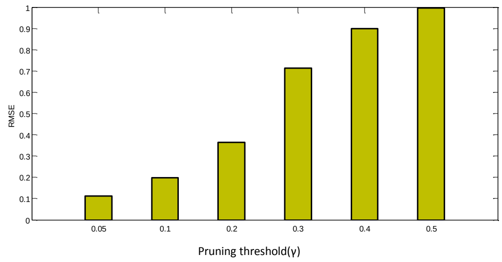


Fig. 6. Tree pruning thresholds and their errors.

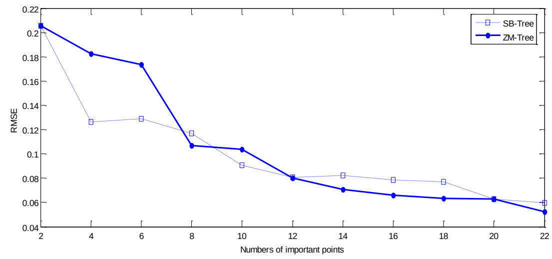


Fig. 7. Comparison of dimensionality reduction errors between ZM-Tree and SB-Tree.

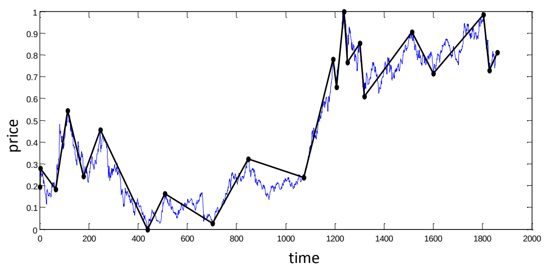


Fig. 8a - ZM-Tree

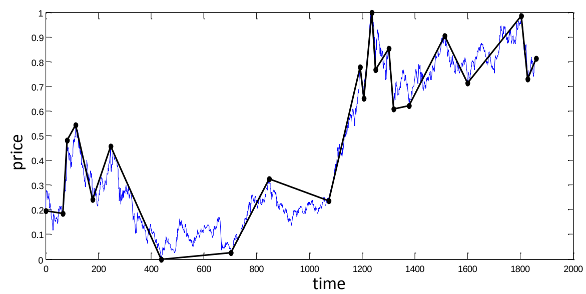


Fig. 8b - SB-Tree

Fig. 8. Comparison of dimensionality reduction errors for (a) ZM-Tree and (b) SB-Tree.

6 CONCLUSIONS

This study, we propose the methods of financial time series indexing by considering their zigzag shape movement. The methods consist of two major algorithms; firstly, the identification of important points so-called Zigzag-perceptually Important Points(ZIPs) identification method, secondly, the indexing method by applying the variety of Multi-way search tree so-called ZIP-based M-ary tree (ZM-Tree). This tree data structure illustrates significantly their benefits in dimensionality reduction comparing to SB-Tree. Further, the experiments should conduct to the several sizes of the time series and several datasets. This to comparing its clearly performance in various conditions.

AKNOWLEDGEMENTS

The Universiti Teknologi Malaysia (UTM) under research grant 03H02 and Ministry of Science, Technology & Innovations Malaysia, under research grant 4S062 are hereby acknowledged for some of the facilities utilized during the course of this research work.

REFERENCES

1. C.D. Kirkpatrick, J.R. Dahlquist, Technical Analysis: The Complete Resource for Financial Market Technicians, Financial times press, 2007.
2. R. Agrawal, C. Faloutsos, A. Swami, Efficient similarity search in sequence databases, In: 4th Conference on Foundations of Data Organization and Algorithms (1993) 69-84.
3. K.-P. Chan, and A.-C. Fu, Efficient time series matching by wavelets. In: Proc. of DE., 1999, pp. 126.
4. E., Keogh, K. Chakrabarti, and M. Pazzani, Dimensionality reduction for fast similarity search in large time series databases. Journal of Knowledge and Information Systems 3(3) (2001) 263-286.
5. E. Keogh, and P. Smyth, A probabilistic approach to fast pattern matching in time series databases, In: Proc. of KDD, 1997, pp. 24–30.
6. F.L. Chung, et al. Flexible time series pattern matching based on perceptually important points. In: Proc. of IJCAI-01, 2001, pp. 1–7.
7. C. Fu-Lai, et al., An evolutionary approach to pattern-based time series segmentation. Evolutionary Computation, IEEE Transactions on 8(5) (2004) 471-489.
8. T.-C Fu, F.-L. Chung, R. Luk, C.-M. Ng, Stock market time series pattern matching: Template-based vs. rule-based approaches. Engineering Applications of Artificial Intelligence 20 (2007) 347-364.
9. T.-C Fu et al., Representing financial time series based on data point importance. Engineering Applications of Artificial Intelligence 21(2) (2008) 277-300.
10. D. Bao. A generalized model for financial time series representation and prediction, Applied Intelligence 29(1) (2008) 1-11.
11. D. Bao, Y. Zehong, Intelligent stock trading system by turning point confirming and probabilistic reasoning. Expert Systems with Applications 34(1) (2008) 620-627.
12. W. Ohta, An analysis of intraday patterns in price clustering on the Tokyo Stock Exchange., Journal of Banking & Finance 30(3) (2006) 1023-1039.
13. A. Lendasse et al.,Dimensionality reduction of technical indicators for the prediction of financial time series application to the BEL20 Market Index., European Journal of Economic and Social Systems 15(2) (2001) 31-48.
14. T.-C. Fu, F.-L. Chung, C.-M. Ng, Financial time series segmentation based on Specialized Binary Tree representation. In Proc of DMIN, 2006, pp. 3-9.
15. W. Huang, Yoshiteru, Nakamori, S.-Y. Wang, Forecasting stock market movement direction with support vector machine. Computer and Operation Research 32(10)(2005) 2513-2522.
16. Zigzag. StockCharts.com, accessed Feb. 20, 2010.
17. J. Lin, E. Keogh, S. Lonardi, and B. Chiu, “A Symbolic Representation of Time Series, with Implications for Streaming Algorithms”, In: Proc. of ACM SIGMOD, 2003, pp.2-11.
18. B. Yi, C. Faloutos, Fast time series indexing for arbitrary Lp norms. In: Proc. of VLDB, 2000, pp. 385-394.
19. E. Keogh et al., Dimensionality reduction for fast similarity search in large time series databases. Journal of Knowledge and Information System 3(3)(2000), 263-286.
20. J. Shieh, E. Keogh. iSAX: indexing and mining terabyte sized time series. In: Proc. of ACM SIGKDD, 2008, pp. 623-631.
21. J.-W. Roh, B.-K. Yi, Efficient indexing of interval time series sequences. Information Processing Letters 109 (2008) 1-12.
22. C.-S. Perng, H. Wang, S.-R. Zhang, D.-S. Park, Landmarks: a new model for similarity based pattern querying in time series database. In: Proc. of ICDE, 2000.
23. E. Fink, B. Pratt, Indexing of compressed time series, Data Mining in Time Series Databases (2003) 51–78.
24. B. Pratt, E. Fink, Search for patterns in compressed time series, Image and Graphics 2 (1)(2002) 89–106.
25. Z. Zeng, H. Yan, Supervised classification of share price trends. Information Sciences 178 (2008) 3943-3956.
26. W. Gorr, A. Olligschlaeger, and Y. Thompson, Short-term forecasting of crime. International Journal of Forecasting 19 (4) (2003) 579-594.
27. R. D. Edwords, J. Magee, W.H.C. Bassetti. Technical Analysis of Stock Trends 9th Edition. CRC Press, 2007

Recognizing a Road Environment Using Multiple-window Bag of Features

S. Morita, J.K. Tan, H. Kim and S. Ishikawa

Kyushu Institute of Technology
Tobata Kitakyushu 804-8550, Japan
Tel : 81-93-884-3183
e-mail: morita@ss10.cntl.kyutech.ac.jp, {etheltan, kim, ishikawa}@cntl.kyutech.ac.jp

Abstract: The idea of Bag of Features (BoF) is recently often employed for general object recognition. But, as it does not take positional relations of detected features into account, the recognition rate is still not very high for practical use. This paper proposes a method of describing the feature of an object by the BoF representation which considers positional information of the features. Although the original BoF representation is applied to an entire image, the proposed method employs multiple windows on an image. The BoF representation is applied to each of the windows to represent an object in the image interested for recognition. The performance of the proposed method is shown experimentally.

Keywords: Object recognition, bag of features, multiple-windows, VLAD, computer vision.

I. INTRODUCTION

The future of a mankind will be more and more complicated and will definitely need the help of an intelligent robot. Then the robot must be equipped with a strong ability of object recognition. On the other hand, a hand-held camera and a wearable computer system which can recognize every object around a blind person may help him/her a lot in living a daily life safely as well as conveniently. Such a system again needs to have a strong ability of object recognition. Various techniques of object recognition have been developed to date. But such techniques normally employ the features depending solely on the objects interested. The features common to every object should be considered to develop a general objects recognition method.

General objects recognition has been paid much attention among computer vision researchers recently. A well-known general object recognition technique is the idea of Bag of Features (BoF) [1]. It is a point-based feature description method and describes every object using a visual word dictionary. But it describes an object as a set of feature points without taking positional information into account. The positional information, or to know how feature points distribute on an object, is actually important information for its recognition. An idea of spatial pyramid matching (SPM) [2] is proposed in order to take positional information of feature points into account. But it is not very effective, since the method segments an image into 2^n by 2^n regions with no overlap some of which may contain only the background of the image.

The present paper proposes a method of describing the feature of an object by BoF representation which considers positional information of the features. Although the original BoF representation is applied to an entire image, the proposed method employs multiple overlapping windows on an image. The BoF representation is applied to each of the windows to represent an object in an image. In this way, the positional information among obtained BoFs is employed for recognizing an object interested.

II. BOF AND VLAD

The idea of BoF is overviewed in the first place followed by giving the concept of VLAD (Vector of Locally Aggregated Descriptors) [3] proposed as another representation of BoF.

Given an object image, the SIFT operator [4] is applied to the image to derive a number of feature points on the object. The point is described by a 128-dimensional vector. It is then projected into a 128-dimensional feature space. A number of object images are respectively transformed into the feature space each as a set of feature points. The feature space then contains a large number of the feature points, to which clustering is applied to make some hundreds or thousands of prominent classes. Let the number of the class in the feature space by M. A class C_i is represented by a feature vector $\boldsymbol{v}_i$ (i=1,2,...,M). The feature space is then defined as a visual word dictionary (VWD) by the set $V=\{\boldsymbol{v}_i| \ i= 1,2,...,M\}$: Vector $\boldsymbol{v}_i$ is referred to a visual word (VW) within the dictionary.

An object image is then described using the VWD. Given an object image, the SIFT operator is applied to the image and the feature points are extracted from the object. Once they are projected into the VWD, they distribute around the VWs which represent the object. Let the number of the feature points distributing around a VW $\boldsymbol{v}_i$ be denoted by f_i. This is actually the frequency of a histogram of the chosen VWs. The object is then characterized by a M-dimensional feature vector

$$\boldsymbol{w}=(f_1, f_2, ..., f_M). \quad (1)$$

This is called Bag of Features (BoF). An object is finally identified by the BoF $\boldsymbol{w}$.

Instead of using the frequency of the VWs, another description of an object [3] is proposed using a VWD (Vector of Locally Aggregated Descriptors). If a feature point extracted from an object image by SIFT is denoted by $\boldsymbol{x}$, the VLAD feature vector is defined by

$$\boldsymbol{w}_i = \sum_{v(\boldsymbol{x})=i} (\boldsymbol{x} - \boldsymbol{v}_i). \quad (2)$$

After all, the VLAD expression provides a 128M-dimensional feature vector of the form

$$\boldsymbol{w}=(\boldsymbol{w}_1, \boldsymbol{w}_2, ...,\boldsymbol{w}_M). \quad (3)$$

The magnitude of the component $\boldsymbol{w}_i$ depends largely on the feature points distributing around VW $\boldsymbol{v}_i$.

III. MULTIPLE-WINDOW BOF

The proposed method puts K mutually overlapping windows (W_1, W_2, ...,W_K) on an image in order to consider positional relation among extracted feature points, which is a strategy different from the original BoF [1]. It also differs from SPM [2] in the overlap of the windows. The idea of multiple windows is shown in **Fig. 1**. The proposed method also introduces VLAD for describing BoF. This means to put more emphasis on the VWs which have many feature points around themselves than the frequency description.

Locating windows on an image has three variations:
(i) Random location; Windows are randomly located on an image;
(ii) Considering feature points distribution: Arranging windows more at the spots where many feature points distribute;

Fig. 1. Multiple-windows set on an object image.

(iii) Combining (i) and (ii): Arranging windows randomly under the condition that the spots where there are many feature points have priority in the arrangement.

Among the above three strategies, (ii) is natural and reasonable, since the present object recognition is feature-points-based recognition. It is, however, important to take some randomness into account to escape from over-learning against training images. This is the reason why (iii) is considered. (i) is conducted for the comparison with (ii) and (iii).

The size of the window and the randomness in the windows placement is determined experimentally.

Let the number of the visual words in a window W_k (k=1,2,...,K) be denoted by M_k. In the original idea of BoF, frequency in the BoF histogram is employed for the components of feature vector $\boldsymbol{w}_k$. Instead of using the histogram, the present method introduces VLAD for describing a BoF. The magnitude of the VLAD $\boldsymbol{w}_{km}$ at visual word $\boldsymbol{v}_{km}$ (m=1,2,...,M_k) in window W_k becomes large, if many feature points distribute close to $\boldsymbol{v}_{km}$ in a biased way. After all, the overall feature vector $\boldsymbol{w}$ is defined by

$$\boldsymbol{w}=(\boldsymbol{w}_1,\boldsymbol{w}_2, ..., \boldsymbol{w}_k, ...,\boldsymbol{w}_K), \quad (4a)$$

$$\boldsymbol{w}_k=(\boldsymbol{w}_{k1},\boldsymbol{w}_{k2}, ..., \boldsymbol{w}_{Mk}). \quad (4b)$$

The dimension of the feature vector $\boldsymbol{w}$ is therefore 128M (M=M_1+M_2+...+M_K).

The recognition strategy employs a nonlinear SVM based on one-versus-rest classification.

IV. EXPERIMENTAL RESULTS

An experiment was conducted using the images in an road environment. The employed objects for recognition are a pedestrian, a traffic signal, a car and a bicycle (See **Fig. 2**). They are all principal objects in the road environment. The number of images used for the training of a SVM is 800; 100 with each object and 400 negative images. On the other hand, the number of images used for test is 500; 100 with each object and 100 negative images. Used PC has a 3.40 GHz CPU with 8 GB memories.

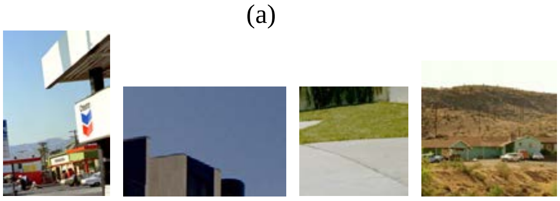
The first experiment, **Exp_1**, was done to examine the performance of VLAD. The original method and SPM employing a frequency histogram for BoF representation were compared to those employing VLAD for BoF representation. The result is given in **Table 1**.

In the second experiment **Exp_2**, the proposed method employing multiple windows is examined its performance with respect to the three cases of windows placement explained in the former section; (i) placement at random, **P_R**, (ii) placement considering feature points distribution, **P_FPD**, and (iii) placement considering feature points distribution and randomness, **P_FPD&R**.

The experimental result is shown in **Table 2**. In Exp_2, the number of VWs is parameterized and it varies from 50 to 500 per window.

As seen in Table 2, the recognition rate is the maximum when P_FPD&R is adopted for windows placement and 200 VWs are employed with every window. The third experiment, **Exp_3**, was conducted under the employment of multiple windows with P_FPD&R, 200 VWs with each window, and VLAD expression for BoF. The result is shown in **Table 3**.

(a)



(b)

Fig. 2. Objects for recognition: (a) Positive samples; a pedestrian, a traffic signal, a car and a bicycle, (b) negative samples.

Table 1. Result of Exp_1: Original BoF & SPM without/with VLAD.

Methods	O-BoF*	O-BoF +VLAD	SPM	SPM +VLAD
Rec rate [%]**	63.0	67.4	69.8	72.8

* Original BoF

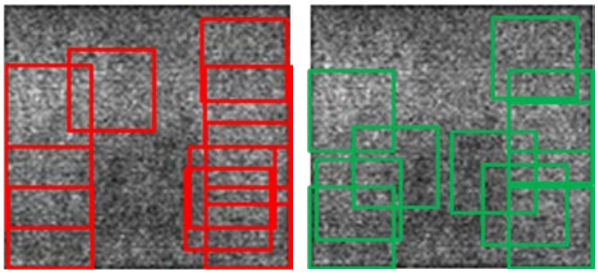
** Recognition rate

Table 2. Result of Exp_2: Multiple window BoF.

	P_R			
No. VWs / window	50	100	200	500
Recogn' rate [%]	68.6	72.8	69.4	68.4
	P_FPD			
No. VWs / window	50	100	200	500
Recogn' rate [%]	71.0	69.4	69.4	68.8
	P_FPD&R			
No. VWs / window	50	100	200	500
Recogn' rate [%]	71.8	73.0	73.2	67.6

Table 3. Result of Exp_3: Multiple window BoF with VLAD.

	Proposed method
Recognition rate [%]	74.8



(a) (b)

Fig. 3. The windows employed in Exp_2 and Exp_3: (a) Case (ii), (b) case (iii).

The number of used windows in Exp_2 and Exp_3 is 10 whose size is approximately 1/3 of the entire image. They are arranged as shown in **Fig. 3**, in which windows are placed where many feature points exist in (a), whereas randomness is considered in addition to the feature points distribution in (b). These windows placements are kept unchanged through Exp_2 and Exp_3.

Table 4. Average recognition rate with respect to three strategies of windows placement.

	P_R	P_FPD	P_FPD&R
Ave. rec. rate*	69.8	69.7	71.4

*Average recognition rate [%]

V. DISCUSSION

In Exp_1, the positive performance of the VLAD was recognized as shown in Table 1. The idea of VLAD is to put emphasis on the VW which is characteristic to a particular object more than frequency, and it worked affirmatively in the recognition of the 10 objects employed in the experiment.

On the other hand, in Exp_2, the maximum recognition rate was 73.2% when strategy (iii) in section III was employed as windows location and 200 visual words were used with each of the 10 windows. When a single window, an image itself, is employed, which is the original way of using the BoF, the recognition rate is 63.0% as seen in Table 1. Although SPM is a multiple window method, the recognition rate is worse, 69.8%, than the proposed method which is more flexible in windows placement than SPM. This fact indicates the effectiveness of the proposed use of multiple windows in the BoF-based object recognition.

Finally, the proposed method, employing multiple windows and VLAD expression, achieved 74.8% of the recognition rate as given in Table 3. This is the best result at the moment.

As for the three strategies of windows placement, strategy (iii) seems to act better than the other two, which is seen in **Table 4**. It shows average recognition rates with respect to each windows placement in Table 2. It may, however, be necessary to perform more experiments to make the superiority certain, since the difference is not very large.

In the employment of BoF, various weights could be considered including frequency of the SIFT feature points [1], VLAD [3], TF-IDF [5] and weighted BoF [6]. But they still give not very high recognition rates to general objects. One may need to improve this in some way.

VI. CONCLUSION

In this paper, multiple window bag of features was proposed which took positional relation of the feature points on an object into account. For BoF representation, the vector of locally aggregated descriptors, VLAD, was also employed for recognizing ten familiar objects in a traffic environment. By effective placement of the multiple windows on an image, 74.8% of recognition rate was achieved, which is satisfactory for general object recognition. However, the research should be continued to raise the recognition rate more in order to put the method into practical use.

REFERENCES

[1] Csurka G, Dance CR, Fan L, Willamowski J, Bray C (2004), Visual categorization with bags of keypoints, Proc. ECCV: Learning in Computer Vision, 1-22.
[2] Lazebnik S (2006), Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories, Proc. IEEE Computer Vision and Pattern Recognition, 2169-2178.
[3] Jégou H, et al. (2010), Aggregating local descriptors into a compact image representation", Proc. CVPR.
[4] Lowe DG (2004), Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, 60(2):91-110.
[5] Frakes WB, Baeza-Yates R (1992). Information Retrieval: Data Structures & Algorithms. Prentice-Hall, Engelwood Clifs.
[6] Manabe T, Tan JK, Kim H, Ishikawa S, Recognizing indoor objects by weighted bag of features, 2014 IEEE Tensymp. (submitted)

Recognized Face Tracking for CONBE Robot

Sakmongkon Chumkamon and Eiji Hayashi

Department of Mechanical Information Science and Technology, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka, Japan, 820-5202
Tel: +819 0 9499 2408
E-mail: m-san@mmcs.mse.kyutech.ac.jp, haya@mse.kyutech.ac.jp

Abstract: In our research, we develop the robot to combine with the consciousness and behavior which is the conscious and behavior robot (CONBE). In this paper we present the basis system in the CONBE robot that consist of the two main sections as the face recognition and robot gaze tracking. This system uses the face recognition using Fisherfaces that also refer to linear discriminant analysis (LDA). For the robot gaze tracking, we implement with the head of the CONBE robot that has a camera, two displays used for robot eyes and the actuators of 2 degrees of freedom. The actuators use to control the robot gaze for tracking the human face which the robot can recognize. In this paper, we experiment and present the results of the face recognition system and the tracking control system such as the tracking error and time.

Keywords: Consciousness-Based Architecture, Fisherfaces, Face recognition, Cognitive robot, face tracking**.**

I. INTRODUCTION

An autonomous robot requires efficient operation of their action. Decades, many good approaches have been proposed for controlling the autonomous robot under their condition such the development of the self-aware robot. Due to a robot with self-awareness has a realistic action to operate under the diversity situation more functional than a system without self-awareness. In this paper, we develop a conscious behavior robot (CONBE-I) that imitates a human head. The mechanism of a head CONBE robot has 2 degrees of freedom that are orthogonal to each other in Cartesian coordinate. In our laboratory, we develop the robot to have the consciousness and behavior as the mention and propose the consciousness-based Architecture (CBA) system as a hierarchical structure model base on the mechanistic expressed model of animal consciousness and behavior [1]. The motivational action of animal is derived from recognition and comprehension respectively. Therefore, in this research, we focus on developing the recognition of the robot to be a basis of CBA system. We develop the face recognition with the CONBE robot to tracking face which is recognized. In face recognition, we use Fisher-faces algorithm that also refers to Linear Discriminant Analysis (LDA), use a video camera as an input instant image [2]. For hardware of the head robot, we use Dynamixel actuators that are controlled via RS485 bus to control the rotation of yaw and pitch axis of the head. For controlling, in the instant image, we determine the distances between the center of the tracking object and the center of camera image in vertical and horizontal axis. We then calculate the distances of the angular errors for each rotation in yaw axis and pitch axis and command the target angle to the actuators to tracking object. As previously mentioned, this research studies the face recognition that is one of recognition models in our CBA, and studies the part of the robot behavior that is an action of following gaze. This system imitates the animal character in the case of the animal seeing some people. It could recognize the persons and interested in them, it then acts some behavior such as a gaze following in our case. Our system imitating the animal system, due to our attempt is to increase the function of the robot.

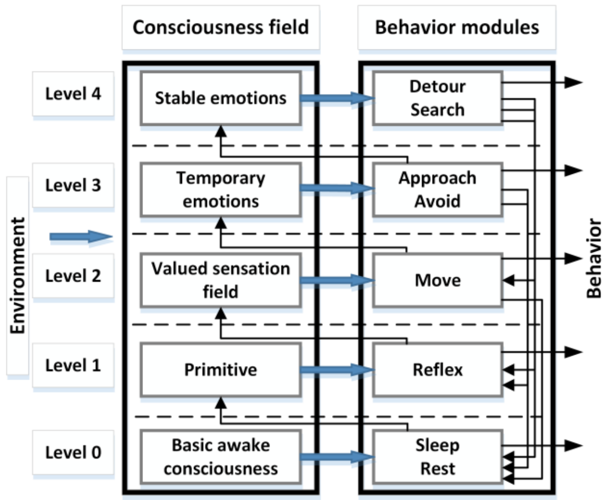


Fig.1. Consciousness-based architecture

The robot would operate more than following the command from human such as when the robot combines CBA system; it can introspect to act some suitable behavior when the robot can recognize the user being sad, happy and so forth. Moreover, this robot would be more attractive for human or user with this proposed.

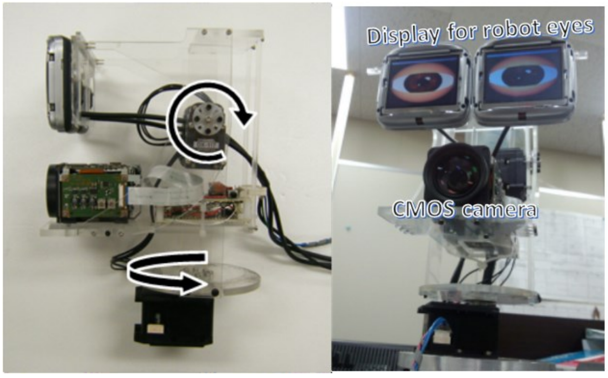


Fig.2. Head robot.

II. ROBOT SYSTEM OVERVIEW

The system based on CBA system combining with robot mechanism and software that we develop to archive face recognition and tracking face which is recognized. This system consist of two main such as hardware and software system.

1. Hardware

For head robot, we designed and constructed the structure with the acrylic plastic. The head robot also assembles with two actuators to rotate in Cartesian coordinate which the first actuator operates as a robot neck which base Dynamixel RX-64 and the second actuator for up down rotating which base Dynamixel DX-117. These actuators are DC gear motor with plastic body which are controlled by packet command containing with an identity number of the actuator via serial RS485 multi drop and yield the feedback such as position of rotation degree, temperature, load, Input voltage, speed and so forth. For obtaining the image to process, we use embedded CMOS camera based on SONY FCB-H11 which support full HD (1080i high definition) video format with; NTSL and PAL. The camera model is equipped with ×10 optical auto focus zoom lens and the ×12 digital zoom. The camera module can control the features of the camera with VISCA protocol via serial port. In this paper, we use this camera in PAL mode connecting with composite video channel and convert composite video channel to USB by USB video capture device. All the device in the robot system is controlled by a laptop which base CPU i7-2630QM, RAM 4GB.

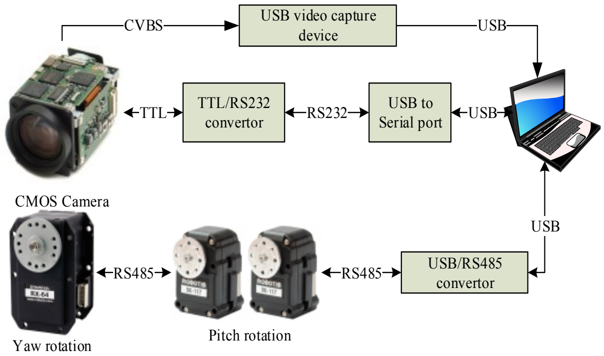


Fig.3. Connection diagram of the system.

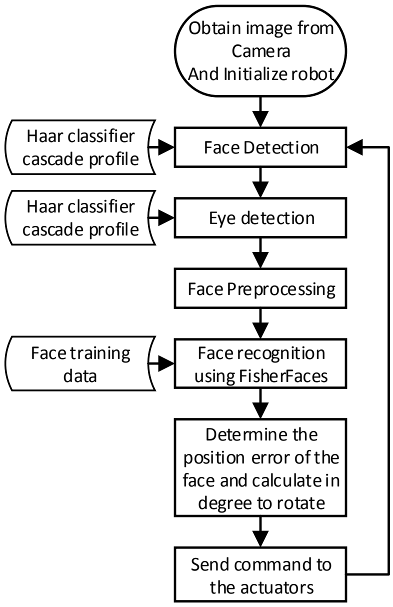


Fig.4. Software flow diagram.

2. Software

The system software has two major such face recognition and control which are implemented by C++ language in a laptop. For the image processing library, this software use OpenCV library. At the beginning software from the face recognition, the system would capture the instant image from the camera then operating with human face detection using Haar features and Haar classifier cascade profile to determine the face position [3]. The software also detects the human eye due to obtain the definite position to provide the face preprocessing image due to the face recognition is vulnerable with some difference in lighting condition, face orientation, face expression, and so forth. For eye detection, it is used to crop the input image to show only the approximate eye region. We then perform the preprocessing image combining geometrical transformation and crop; separate histogram equalization for left and right side; and smoothing image using a bilateral filter; and elliptical mask to remain hair and background from the face image. Eventually we perform the face recognition using Fisherfaces algorithm with our training data and face label. Fisherfaces is the popular algorithm for face recognition using LDA which is to improve the capability of classification from PCA which is the core of the Eigenfaces method that is a powerful method to represent data. The simple concept of LDA is that the same classes should group closely together whereas different classes are more distance [4]. In this section we provide the environment for recognized face tracking. For the next section, we would experiment this system to apprehend the performance and other specific of this system.

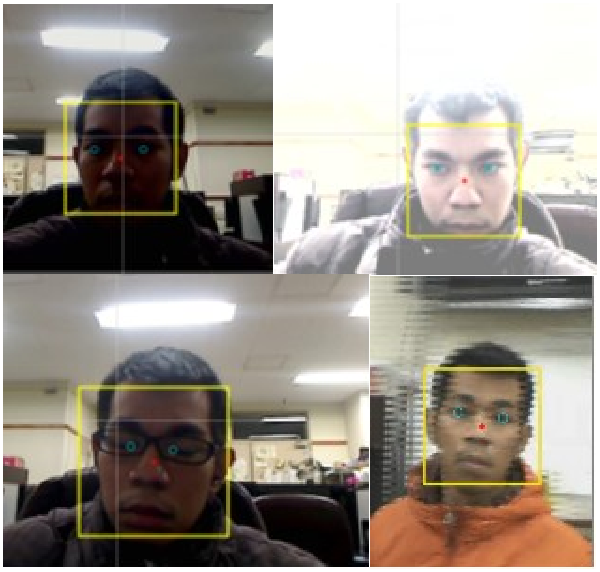


Fig.5. Example image of face and eye detection in different case; the rectangle is the face region and the circle is the position of the left and right eye.

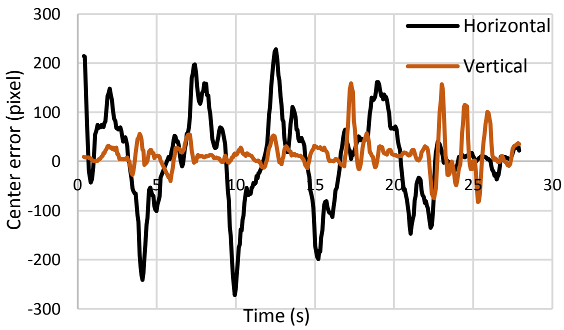


Fig.6. Distance error between the image center and tracking object center in horizontal and vertical.

III. EXPERIMENT

In the system that we already provided, we experiment this system divided in two main sections; the first is face detection and recognition; and the second is the tracking system or the robot gaze following the recognized face. The face detection is implemented using local base pattern feature cascade classifier with the restricted condition of the smallest size image as 25×25 pixels which can detect the face. For the eye detection, we used Viola-Jones object detection framework with Haar feature-based cascade classifier that is an effective and popular object detection method [3]. We experiment the performance of the face and eye detection in different condition with some results as shown in Fig. 5. The face detection also is implemented to incorporate with the robot to experiment for the robot face tracking. The robot face tracking perform the experiment for the robot tracking performance as shown in Fig 6. In this experiment, we test the robot tracking the face while the human moving and facing the camera, the first we move from left to right side and return, the second we move up and down to test. The distance between the human while testing is around 1.5 meters. The experiment finds out the performance and the characteristic system such as the average frame rate as 16 frames per second, the average error as 5, 15.6 pixels, the minimum error as -272, -83 pixels and the maximum value as 228, 159 pixels in horizontal and vertical axis respectively.

The face recognition algorithm test with the laptop and the camera in the robot incorporating with the tracking control. In the experiment, 6 members of our lab have contributed the sample videos to construct the face recognition algorithm. The training program acquired the image by capturing each contributed videos. During the video recording using the camera in the head robot, the member would face the camera first then turn left, right and move up down his face in front of the camera slightly. Every image would perform the face detection using LBP-based cascade classifier profile and preprocessing image to preparing for training in Fisherfaces algorithm to construct the face recognition model [2], [5]. The training image size after resize and preprocessing would be 120×120 pixels.

Fig. 7 illustrates the example of the experimental results face recognition and tracking. There are two member lab testing the system, left side image which is the recognized face then the robot would track the face whereas the right side image which is the unknown face then the robot would not track the face. Fig 8 shows the results of the recognized face tracking such as the error of the center image between the center of the face region and the center of the camera image in horizontal axis and vertical axis. In this condition we move the face on the right, left, up and down to find out the characteristic and performance of the tracking control. In this experiment, we find out the average frame rate at 13 frames per second, a minimum error as -121, -59 pixels, a maximum error as 158, 84 pixels, and the average error as 2.7, 13.8 pixels in horizontal and vertical respectively.

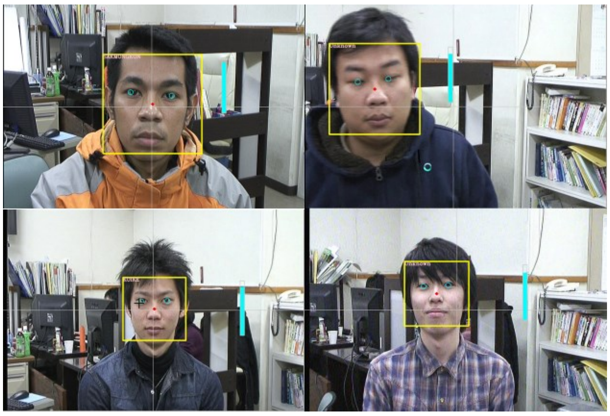


Fig.7. Example images in the robot tracking of recognized face. The left side show the recognized face tracking and the right side show unknown face.

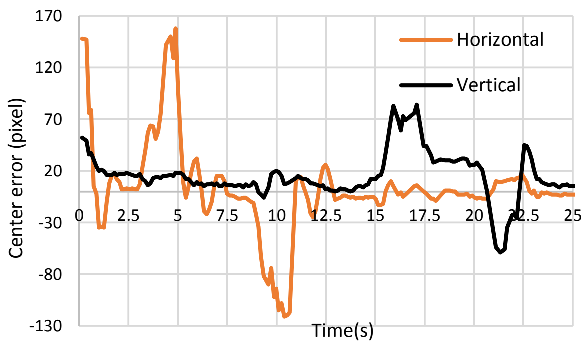


Fig.8. Distance error between the image center and tracking object center in horizontal and vertical.

IV. CONCLUSION

This paper presents the face recognition combining with image tracking by CONBE robot for an experimental personal robot in our concept to develop the CBA. In the proposed visual tracking and the system processes around 16 frames per second for face detection and 13 frames per second for face recognition incorporating with robot face tracking. For the face detection and tracking the robot is able to track the face following pretty fast but in the recognized face tracking, the robot following pretty slow as the face recognition algorithm processing time and depending on the difficulty of the facial similarity. Because while testing the recognized face tracking, this system could not recognize the face every frame as the image is a blur image from the movement reason.

As the results, we first proceed the system to improve the robust face recognition in real time as well as possible. We then develop the facial express recognition to recognize the emotion expressing on the face to fulfill the part of the CBA operating with the face.

ACKNOWLEDGEMENT

We would like to thank the members of our lab who contributed the image along with for testing the system and thank Mr. Masato Koike from Department of Mechanical Information Science and Technology, Kyushu Institute of technology who contributed the sample source code for the robot eye expression.

REFERENCES

[1] Hayashi E, Ueyama K, Yoshida M (2011), Autonomous motion selection via consciousness-based architecture. In Ubiquitous Robots and Ambient Intelligence (URAI), 2011 8th International Conference on, pp. 401-402

[2] Belhumeur P. N., Hespanha J. P., Kriegman D. J. (1997), Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. In Pattern Analysis and Machine Intelligence, IEEE Transactions on, pp. 711-720.

[3] Viola P., Jones M. (2001), Rapid object detection using a boosted cascade of simple features. In Computer Vision and Pattern Recognition, Proceedings of the 2001 IEEE Computer Society Conference on Vol. 1, pp. 511-518

[4] Turk M, Pentland A (1991), Eigenfaces for recognition. In Journal of cognitive neuroscience, pp. 71-86.

[5] He DC, Wang L. (1990), Texture Unit, Texture Spectrum, and Texture Analysis. In Geoscience and Remote Sensing, IEEE Transactions on, vol. 28, pp. 509 - 512.

Eye detection Using Composite Cross-Correlation form Face Images in Varied llumination

Kutiba Nanaa[1], Mohamed Rizon[2] and Mohd Nordin Abd Rahman[1]
[1] School of Computer Science, Faculty of Informatics and Computing,
[2] School of Manufacturing Technology, Faculty of Design and Engineering Technology,
Universiti Sultan Zainal Abidin (UniSZA), Terengganu, Malaysia

Abstract — Facial feature detection is essential process in the field of face recognition. Template match based approach is widely used to implement the issue of facial detection. However, many obstructions adversely affect in template based implementation such as varied illumination, size, face pose and the status of eye in either open case or close case. the objective of this paper is to detect left eye and overcome the obstruction of low illumination in given image using composite cross-correlation. Two single templates form the proposed composite template. First single template is created from an eye image in ordinary illumination while the second single template is created from an eye image in low illumination. The experimental results is concluded by applying our templates on the PICS Database whose varied illumination images and show that applying composite templates gives a better result than applying the both single templates individually.

Introduction

Template match based approaches are widely used for finding small parts of an image which match a given template image. In the field of facial features detection, many methods have been proposed based on template match such as deformable template which links edges, peaks and valleys to find the best place of eye. However, in creating templates verse, several obstructions that may adversely affect on the result of template match based approches such as how much eye open is, the variety in sizes and illuminations between eye templates and the eyes in the given face images, and the pose of face in the given image. In the fact, the behavior of creating eye templates plays a fundamental role in the performance and result of applying template match based approaches.

However, Dong and Wu have proposed a template match method to detect the eye whether it is fully open or semi open in the paper aimed to determine the fatigue driver by compute the distance of the eyelids during driving [1] while Bhoi and Mohanty have employed match template to detect the eye in both cases open eye and close eye regardless the prior [2]

In order to overcome the obstruction of verious sizes between the eye in template and the eye in given image, *Peng et al.* have proposed a method to expect the eye size in the given image to resize the eye template [3]. Multi-scale templates is also proposed by *Florea et al.*, to detect the eye in deferent size of given images in the study aimed to analyze the upper facial expressions [4].

Furthermore, in order to overcome the obstruction of verious poses between the eye in template and the eye in given image, Chen *et al.*, have proposed a method to detect the human face in deferent poses. In the fact, they created the face template from only half face. By this way, the implementation of the method has be accelerated and it has gaven the same results for the frontal face images and has given batter results in the case of deference in face poses. On the same criterion, creating eye template from only half part of eye overcome the obstruction of face poses as suggested by Nanaa *et al.*, in composite cross correlation [5].

One of the ideas that have been realized by creating a template of from only the iris [6]. Kuronuma and Suzuki has employed this template to detect the pupil in vehicle operations [7]. and it has achieved good result because of the ease of preparation template due to the regular circular shape of iris.

In our paper we present a method to overcome the various of illumination between eye templates and the given image. We employed composite cross correlation which have been presented in [5] in order combine the impact of ordinary and low illimunation templates in composite eye templates.

Proposed Method

In general, detecting eye form given image I by using cross-correlation requires an eye template for eye T of size of q × p. as shown in [2]. Both of template and given image is converted into gray images. cross correlation is defined as in Equation (1).

$$Co_{(x,y)} = \frac{\sum_s \sum_t \delta_{I_{(x+s,y+t)}} \delta_{T_{(cx+s,cy+t)}}}{\sqrt{\sum_s \sum_t \delta^2{}_{I_{(x+s,y+t)}}}\sqrt{\sum_s \sum_t \delta^2{}_{T_{(cx+s,cy+t)}}}} \qquad (1)$$

Where:

$$\delta_{I_{(x+s,y+t)}} = I_{(x+s,y+t)} - \bar{I}_{(x,y)};$$
$$\delta_{T_{(cx+s,cy+t)}} = T_{(cx+s,cy+t)} - \bar{T};$$
$$s \in \left[\frac{-q}{2}, \frac{q}{2}\right]; \;\; t \in \left[\frac{-p}{2}, \frac{p}{2}\right];$$
$$x \in \left[\frac{q}{2}, n - \frac{q}{2}\right]; \;\; y \in \left[\frac{p}{2}, m - \frac{p}{2}\right];$$
$$cx = round\left(\frac{q}{2}\right); \;\; cy = round\left(\frac{p}{2}\right);$$
$$\bar{I}_{(x,y)} = \frac{1}{q \times p} \sum_s \sum_t I_{(x+s,y+t)};$$

$$\bar{T} = \frac{1}{q \times p} \sum_s \sum_t T_{(cx+s,cy+t)}$$

In this correlation matrix,. the element that contains the maximum correlation value refers to the center of the eye in the face image I. However, Elements' value in correlation matrix locate in the range [-1,1]. In order to increase the contrast in the correlation matrex we apply Equation (2) .

$$Co''_{(x,y)} = \frac{Co'_{(x,y)}}{\max(|Co'_{(i,j)}|)} \qquad (2)$$

Where:

$$x, i \in [0, n]; \quad y, j \in [0, m]$$

In our methodology we used an eye template in ordinary illumination for PICS database. From incorrect detection result due to low illumination we used the eye in the worst detection result as new template and combine this templates together to form composite cross correlation as presented in [5]. Equation (3) defined the composite cross correlation for k number of templates.

$$Co'''(I,T)_{(X,Y)} = \prod_h^k Co''(I,[T_h])_{(X,Y)} \qquad (3)$$

Where:

$$T = [T_1, T_2, \dots, T_k]$$
$$h \in [0, k]$$

In accordance with our templates, we rewrite the formula of cross correlation in Equation (1) for the ordinary illumination template T_1 and low illumination templatT_2 considering to the Equality of templates' size as defined in Equation (4).

$$Co_{(x,y)} = \frac{\gamma_{1(x+s,y+t)} \times \gamma_{2(x+s,y+t)}}{\rho_1 \times \rho_1 \times \sum_s \sum_t \delta^2{}_{I(x+s,y+t)}} \qquad 4$$

Where:

$$\gamma_{1(x+s,y+t)} = \sum_s \sum_t \delta_{I(x+s,y+t)} \delta_{T_1(cx+s,cy+t)};$$
$$\gamma_{2(x+s,y+t)} = \sum_s \sum_t \delta_{I(x+s,y+t)} \delta_{T_1(cx+s,cy+t)};$$
$$\rho_1 = \sqrt{\sum_s \sum_t \delta^2{}_{T_1(cx+s,cy+t)}};$$
$$\rho_2 = \sqrt{\sum_s \sum_t \delta^2{}_{T_2(cx+s,cy+t)}}.$$

Fig. 1 demonstrates the use of the composite templates by applying two eye templates on a given image I, T_1 is eye template in ordinary illumination and T_2 is eye template in low illumination. The cross correlation is shown as Co(I,[T_1]) and Co(I,[T_2]) for T_1 and T_2, respectively. The composite cross correlation for both templates is shown as Co(I, [T_1, T_2]). However, dark pixels in cross-correlation image shows the point whose a low corrolation with the template while the bright pixels shows the point whose a high correlation with the template. In composite cross-correlation image Co(I, [T_1, T_2, we noticed that the eye position has much brightness than same positionin single composite-correlation images

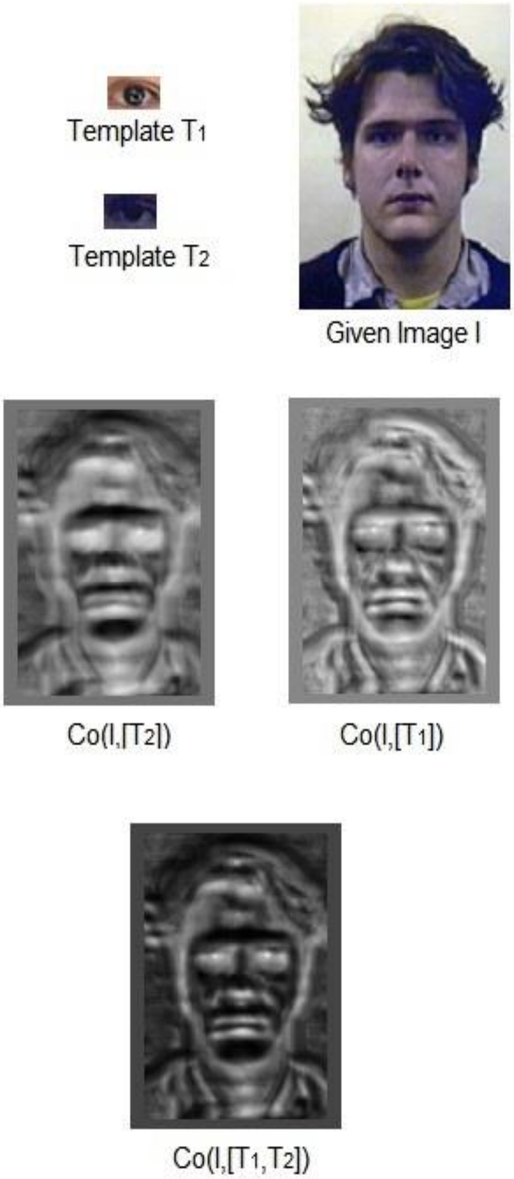


Fig. 1.Using of Ordinary and low Illumination eye template in composite cross correlation.

EXPEIMENTAL RESULTS

. The experimental results has been implemented applied on PICS database [8], a well-known free database of faces using two eye templates, ordinary and low illumination templates to form composite left eye templates [T_1,T_2]. However Fig. 1 shows the templates of T_1 and T_2. PICS Database includes 90 individual faces in various illuminations. Each individual face has a number of images between 1 to 18. The size of images is varied from 336×480 to 624×544.

However, the results demonstrate that the ordinary illumination template exactly detect the eye in the ordinary given images while low illumination template produce a good results in low illumination given images.

Furthermore, applying composite templates produces the correct detection from detection results that produced by applying single template (T_1 or $T_{2)}$, individually. Amazingly the composite template recorrect the detection result although the failure in detection by both of single templates T_1 and T_2, (See the forth row in Fig. 2.

Generally, applying singal templates is the same method used in [2] and it gives 90% as detection rate on PICS database while it gives 98% as detection rate by using our proposed compisite template on the same database.

Fig. 2 shows samples of the results in varied illumination. From left to right, first column shows a given image. Second column shows the correlation results of ordinary illumination template T_1. Third column shows the detection results of applying T_1. The detection result refers to the point whose the much brightness value in correlation result (second column). Forth column shows the correlation results of low illumination template T_2 and its detection results is shown in fifth column . Correlation produced form the composite [T_1,T_2] is shown in sixth column and in the last column is the final left eye detection by applying the composite templates.

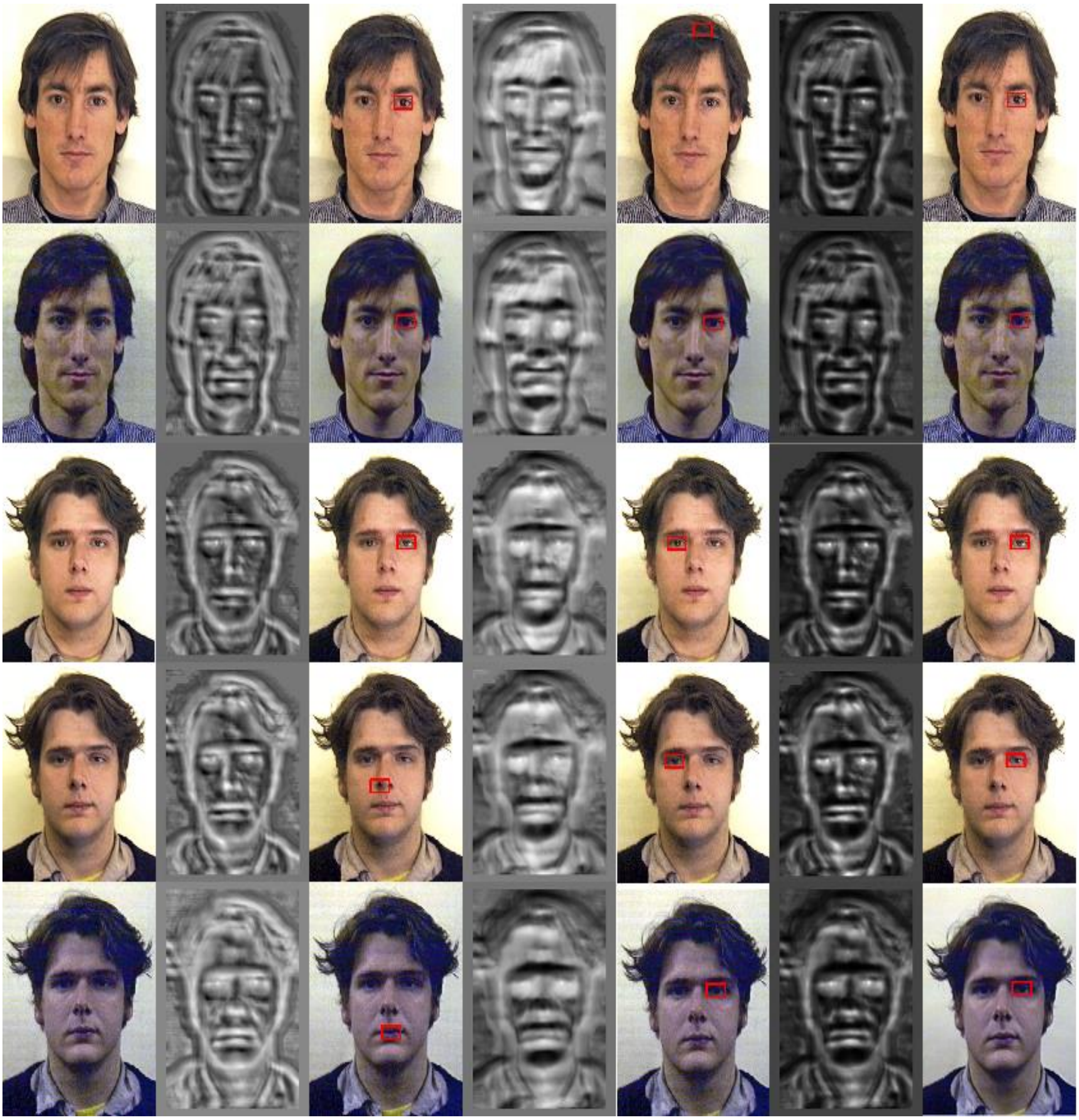


Fig. 2. sample of template match based appoache to detect the left eye by using ordinary illumination eye template, low illumination eye template and composite eye template

CONCLUSION

In this paper we demonstrate eye detection by using cross correlations. We have presented the previous study used template match based method and the improvement to overcome the obstructions in varied size, varied orientation and varied status of eye in both cases (open and close). The objective in this paper is to overcome the obstruction of varied illumination by using composite template in order to detect the eye in varied illumination conditions.

The proposed composite template is compound of two single templates, an eye template in ordinary illumination and another eye template in low illuminations. The experimental results show that applying composite template produces a correct eye detection that produced from applying the appropriate templates and it re-correct the detection in the case of failure of both detection that produced by applying single templates. Comparing the results, using our proposed composite template rises the eye detection in PICS Database form 90% up to 98%.

REFERENCES

[1] W. Dong and X. Wu, "Fatigue detection based on the distance of eyelid," *VLSI Design and Video Technology, 2005. Proceedings of 2005 IEEE International Workshop on*, pp. 365-368, 2005.

[2] N. Bhoi and M.N. Mohanty, "Template matching based eye detection in facial image," *Int. J. Of Computer Applications*, vol. 12, pp. 15-18, 2010.

[3] K. Peng and L. Chen, "A Roust Algorithm for Eye Detection on Grav Intensity Face without Spectacles," *JCS&T,* Vol. 5 No. 3, 2005.

[4] L. Florea, C. Florea, C. Vertan and B. Ionescu, "Upper face description by comparative analysis of gray-scale image pairs," *Signals, Circuits and Systems (ISSCS), 2011 10th International Symposium on*, pp.1-4, 2011.

[5] K. Nanaa, M. Rizon, M. N. A. Rahman, A., Almejrad, A. Z. A. Aziz and S. B. Mohamed,. "Eye Detection Using Composite Cross-Correlation," *American Journal of Applied Sciences*, vol. 10, no. 11, pp. 1448-1456, 2013.

[6] T. Kawaguchi and M. Rizon,"Iris detection using intensity and edge information," *Pattern Recognition*, vol. 36, no. 2, pp. 549-562, 2003.

[7] K. Kuronuma, S. Suzuki and H. Igarashi, "Estimation Method of Awareness Level Using An Iconic Model for Support Robots," *In The First IFToMM Asian Conference on Mechanism and Machine Science*, vol. 21, 2010.

[8] PICS: hltp://pics.psych.stir.ac.uk.

Development of an autonomous-drive personal robot "An object recognition system using Self-Organizing Map"

Keisuke Ito and Eiji Hayashi

Department of Mechanical Information Science and Technology
Faculty of Computer Science and Systems Engineering,
Kyushu Institute of technology, 680-4, Kawazu ,Iizuka, Fukuoka 820-8502, Japan
Tel : 0948-29-7793
ito@mmcs.mse.kyutech.ac.jp

Abstract: We have been developing an autonomous personal robot, it's able to perform practical tasks in a human environment based on information derived from the CCD camera images and the Laser Range Sensor (LRS), It is very important that the robot be able to move autonomously in a human environment, and to select a specific target object from among many objects. Thus, in this paper, we developed the system in order to determine and decide the type of an object using Bag of Features (BoF) and Self-Organizing Map (SOM). The BoF method can extract local features and create a one-dimensional histogram, the SOM can cluster the input data and create a neural network map by using one-dimensional histogram data of BoF. Because of these methods are able to create a neural network and determine the type of the object using histogram of local features. In addition, we focus on recognition of plastic bottle, evaluate its performance and discuss our experimental results.

Keywords: Personal robot, Image processing, Object recognition, Self-Organizing Map

I. INTRODUCTION

In the near future, the autonomous self-driving robots are expected to provide various services in human living environments. For this to occur, the robots will need to gain a grasp in human environments. Therefore, the systems to provide environmental recognition based on image information are being widely studied. However, it is very difficult to recognize all driving environments from image information only; so far, no prospects for such a system have emerged. The robot's system is composed of an autonomous run system for movement and an object recognition system for the recognition and grasping of an object. The object recognition system is composed an object-recognition processing part and a location-information acquisition processing part, both of which use the monocular camera and the LRS. In this research, we focus on the object recognition process by using a CCD camera and special recognition of plastic bottle. In previous research, we developed object recognition system using Speed-Up Robust Features (SURF) that can progress the robust extraction with the bottle features. As a result, the recent robot can determine the angle of an object and operate with the target object from the image provided by the CCD camera. In order to build the robots acting more autonomously, the system need to combine an unsupervised learning with a template matching algorithm. Therefore, in this paper, we developed the system in order to determine the type of an object by using BoF and SOM. This system can visualize the relationship of the object and determine the type of the object.

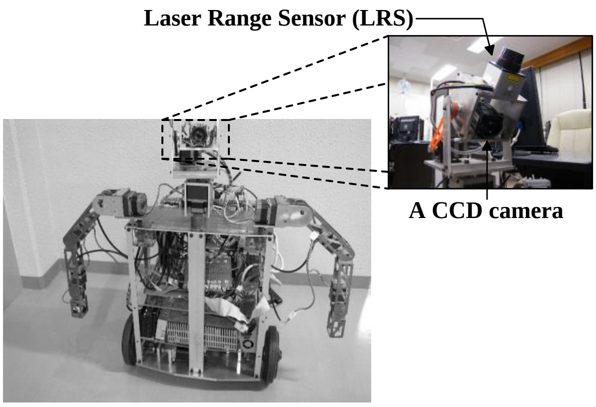


Fig.1 Our developed robot

II. SYSTEM OF AUTONOMOUS ROBOT

The figure 1 shows the appearance of our robot, that has a drive mechanism consisting of two front and two back wheels. The front wheels are attached to the motors which means that the wheels on either side are driven independently, while the back wheels function as castor wheels. This method has the advantage of allowing a small turning radius. In addition, to acquire image information, both a single CCD camera with approximately 2,000,000 pixels and an LRS are installed on the head of the robot and can be rotated to all sides by two motors. DC servo motors are used for the robot's drive mechanism, position and speed control are achieved by the control system of the drive mechanism. The robot also has two arms and hands equipped with sensors, which enable it to respond to the various demands of humans. Finally, an installed wireless LAN can provide remote control for humans. All devices are controlled by a personal computer, and lead-acid batteries are used as power sources in the robot's system.

III. OBJECT RECOGNIZE SYSTEM

1. Outline

This system is composed of BoF and SOM process as shown in Fig.2. These processes permit in the system for learning and determining the objects by using learning data. We explain the database learning system. First, the system load an image data, extract a local feature point and calculate one-dimensional histogram using BoF method. Second, The SOM create an output layer using output data from BoF process. Finally, this system store both BoF data and SOM data. We explain decision process. Next, I explain object recognition process by using created BoF data. First, this system acquires object image data and calculates one-dimensional histogram using same Visual Words created by learning process. Second, this system compares histogram, calculates the Euclidean distance and selects SOM's node which is most nearest against input data.

At the end, this system decides the kind of object by using nearest input data. We will present the algorithm of processing below. We think about design recognition of plastic bottle.

2. Bag of Features (BoF)

A. Outline

Bag of Features (BoF) is a kind of expression of local features of brightness gradient[1]. BoF can convert all local features in the image created to one-dimensional histogram by using the center of gravity, which features called visual words. BoF method uses

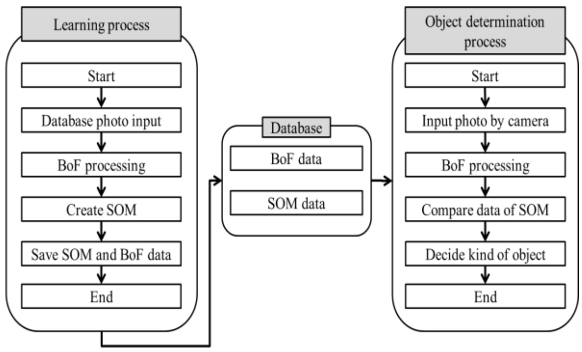


Fig.2 Flowchart of process

Speed-Up Robust Features (SURF) which is a standard local feature extraction and K-means method which is a standard data clustering method. SURF can extract local features quickly and accurate. K-means method can cluster feature data made from SURF and calculate visual words quickly. BoF stores two data, visual words and result of BoF. In the case of the estimation kind of object, by using the same visual word, this system can estimate kind of object. The processing algorithm is as follows.

B. Learning algorithm of BoF method

Step1: Load images and Extract features

The images are loaded from database and extract local features from each image and gather all features by using SURF.

Step2: Clustering features

By using K-means method, cluster all local features of the input images and create visual words.

Step3: Create a one-dimensional histogram

By using Visual words and the SURF method in eac h database image, each feature in images vote most simi lar visual words and create a one-dimensional histogram.

Step4: Store the hitogram data into the database

Save all one-dimensional histogram data and Visual words as Extensi ve markup Language (xml) data.

C. Decide the kind of an object by BoF algorithm

Step1: Extract features

Load images from database and extract local features. Load the BoF database, visual book and one-dimensional histogram.

Step2: Create one-dimensional histogram

By using visual words which are same learning

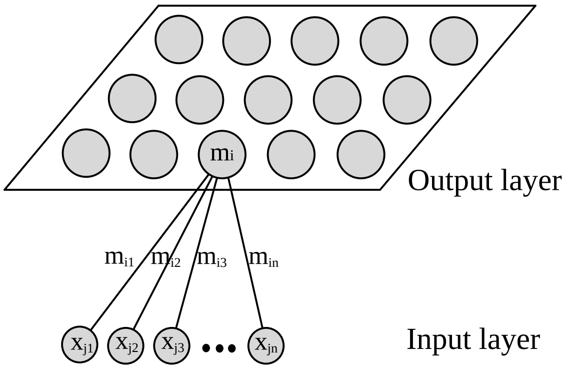


Fig.3 Structure of SOM

process and doing vote of nearest vote, create a one-dimensional histogram.

Step3: Compare a histogram and the decision object

By calculating the Euclidean distance between input histogram and database histogram in order to decide a kind of object.

3. Self Organizing Map (SOM)

A. Outline

A Self Organizing Map (SOM) is unsupervised machine learning method developed by Teuvo Kohonen [2]. The SOM has 2-layer neural network composed of the input layer and the output layer. The SOM can calculate feature by using the Euclidean distance between input data and output data and be placed close to the network. This characteristic allows a mapping from high dimensional feature vector to low dimensional layer without impairing the relationship. We will present the algorithm of SOM processing as below.

B. Learning Algorithm of SOM method

Step1: Initialize network

Set the learning processing times T , the number of output layer's neuron i and the initial learning rate α_{init} . The connection weights between the input layer x and the output layer m are initialized with small random values .

Step2: Present the input vector

The database histogram is used to load into the input layer. In this case, we use one-dimensional histogram as input data.

Step3: Calculate the Euclidean distance

Calculate the Euclidean distance between the input data and the output layer's neurons by Equation (1). And find the winner unit c (Best Matching Unit, BMU) which has the minimum Euclidean distance.

Fig.4 Types of plastic bottles

$$\left| x - m_c \right| = \min \left\| x - m_i \right\| \tag{1}$$

Step4: Adjusting the weights

The connecting weight vectors of all neurons are updated by Equation (2). $h_{c,i}(t)$ is the neighborhood function that will decrease in time. The most commonly used neighborhood function is the Gaussian function which is represented by Equation (3) .

$$m_i(t+1) = m_i(t) + \alpha(t) h_{c,i}(t)[x(t) - m_i(t)] \tag{2}$$

$$h_{c,i}(t) = \exp\left(-\left\| r_i - r_c \right\|^2 \Big/ 2\sigma(t)^2 \right) \tag{3}$$

Where, $r = \left\| r_i - r_c \right\|$ is calculated between unit i and winner unit c and $\sigma(t)$ is called the neighborhood radius. In addition, $\alpha(t)$ is the decay of the learning rate is calculated each iteration using the follow Equation (4).

$$\alpha(t) = \alpha_{init} \left(\frac{T - t}{t} \right) \tag{4}$$

Step5: Repeated learning

The time t increases to t+1. If $t < T$ then go to step 2; otherwise stop the training.

Step6: Decide class of object

Using the average data of each class, decide class of output neuron. The class which is the minimum Euclidean distance in neuron class.

IV. EXPERIMENTAL RESULTS

1. Simulation of the classification by SOM

A. Outline

In this simulation, we evaluated the effectiveness of the development model when evolving a SOM. We confirmed that a reference vector depended on an input value. Table 1 shows each parameter used in the SOM network part of the proposed model.

Table 1. Parameters used to classify simulations.

Number of visual words	300
Output layer’s size	15 x 15
Initial learning rate	0.5
Learning number	1000

In this system, the database has 13 classes, the number of each class assigned to the types of plastic bottles as shown in Fig.4. In each class has 12 images because we use those picture each 30 degrees.

B. Result and discussion

The classification on the map shows in Fig.5, which is derived 13 types of plastic bottles by dividing the cells on the map based on unsupervised learning. In the SOM learning, the similar node gets together in the near location on the map. This result shows SOM’s characteristic that similar object gather near. However, collections of class don’t gather near. For instance, class1 and class 2 (Coca Cola and Coca Cola Zero) and class 8, class 9 and class 10 (Oi-otya, Oi-otya genmai and Oi-otya-koi). We consider it occurred this result that latest database has bias of input data. For instance, tea is 8 kinds. So we need to increase database object and consider the best setting of BoF and the SOM.

2. Experiment of object recognition system

A. Outline

We try to experiment and evaluate the accuracy of this object type determination system. Regarding the experiment’s condition of BoF and the SOM learning, we use the same condition as experiment 1. The input data has 3 objects such as Oi-otya (8), Oi-oyta-koi (9) and Lemon (5), and 6 images per object.

Table 2. Result of object recognition

Input object	The recognition rate [%]
Oi-otya (8)	66 %
Oi-otya-koi (9)	100 %
Lemon (5)	66 %

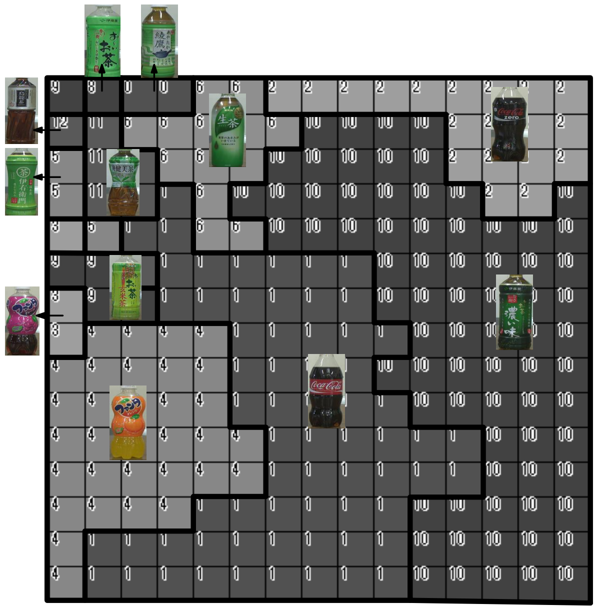


Fig.5 The classification on the map

B. Result and discussion

This result shows that recognition system at Table2 that can recognition difference of between similar objects. For this system more accurate, to store several objects and to detect features which can divide conclusively the type of object are necessary. So, we need to consider adding database object and setting parameter of BoF and SOM system.

V. CONCLUSION

We propose the object recognition system using the Self-Organizing Map and the Bag of Features. This syst e m can extract local feature and create a one-dimensional histogram by using visual words which is the center of gravity features. The SOM can cluster input a one-dimensional histogram and create a neural network map by using of the data. We have shown experimentally this advantage. In order to improve the performance of the system, we will add input data and consider the best setting of BoF and SOM.

REFERENCES

[1] Keiji Y., et al., CVIM tutorial series computer vision leading-edge guide 2(Shinjuku,Tokyo,Japan: Advanced Communication Media CO.,LTD., 2011) , p.85-110.

[2] Kohonen, T., Self-Organizing Map, trans. Heizou, T., et al., (Bunkyo, Tokyo, Japan: Springer-Verlag Tokyo, 1998), p.102-171.

Comparison of Feature Detectors for Obstacles Detection

Shaohua Qian[1], Joo Kooi Tan[1], Hyoungseop Kim[1], Seiji Ishikawa[1], Takashi Morie[2] and Takashi Shinomiya[3]

[1]*Department of Mechanical & Control Engineering, Kyushu Institute of Technology*
Sensuicho 1-1, Tobata, Kitakyushu, 804-8550, Japan
{qian,etheltan,ishikawa}@ss10.cntl.kyutech.ac.jp; kim@cntl.kyutech.ac.jp

[2]*School of Brain Science & Engineering, Kyushu Institute of Technology*
Hibikino 2-4, Wakamatsu, Kitakyushu, 804-8550, Japan
morie@brain.kyutech.ac.jp

[3]*Department of Business Administration, Japan University of Economics*
Sakuragaokacho24-5, Shibuya-ku, Tokyo, Japan
t.shinomiya@aa.cyberhome.ne.jp

Abstract: We have already proposed an obstacles detection method using a video taken by a vehicle-mounted monocular camera. In this method, correct obstacles detection depends on whether we can accurately detect and match feature points. In order to improve the accuracy of obstacles detection, in this paper, we make a comparison between four most commonly used feature detectors; Harris, SIFT, SURF and FAST detectors. The experiments are done using our obstacles detection method. The experimental results are compared and discussed, and then we find the most suitable feature point detector for our obstacles detection method.

Keywords: Feature detectors, Harris, SIFT, SURF, FAST, car vision.

I. INTRODUCTION

Detection of obstacles in a video sequence is a basic task in autonomous collision avoidance systems of intelligent vehicle. Accurate obstacles detection will improve the performance of obstacles tracking, recognition, classification and motion analysis. We have already proposed an obstacles detection method using a video taken by a vehicle-mounted monocular camera [1]. But this method detects 2D and 3D objects simultaneously. Since these 2D objects are not dangerous to driving, they will reduce the accuracy of detection if they are detected as obstacles. In order not to detect these 2D objects, we also have proposed a method for classifying 2D objects and 3D objects [2]. In this 2D and 3D objects classification method, correct classification depends on whether we can accurately estimate the camera motion parameters: Using the camera parameters, the method calculates the coordinates of 3D points in the world coordinate system using two corresponding feature points in two consecutive images. The first step of the camera motion estimation is corresponding feature points detection in two consecutive images: It involves the feature points detection and matching.

Feature points detection and matching is the task of establishing correspondences between two images of the same scene. This is an important problem in computer vision with applications in object detection, object recognition and structure from motion. While feature points detection and matching has been studied extensively for various applications, our interest is to reliably match two images in real time for camera motion estimation. Accurate feature points detection and matching can improve the accuracy of 2D and 3D objects classification, and improve the accuracy of obstacles detection ultimately.

Currently the most commonly used feature detectors are Harris, SIFT, SURF and FAST detectors. In this paper, we will make a comparison between these four different feature detectors. The experiment will be done using the 2D and 3D objects classification method [2]. The experimental results will be compared and discussed, and then we find the most suitable feature point detector for our obstacles detection method.

The structure of this paper is as follows. First four feature detector methods are overviewed in section 2. Experimental results are shown in section 3. Finally the paper is concluded in section 4.

II. FEATURE DETECTORS

1. Harris

The Harris detector [3] is based on the local auto-correlation function of a signal: The local auto-correlation function measures the local changes of the signal with small windows shifted by a small displacement in different directions.

Let $E(u,v)$ be the change of intensity for the small shift $[u,v]$:

$$E(u,v)=\sum_{x,y} w(x,y)[I(x+u,y+v)-I(x,y)]^2 \tag{1}$$

where $w(x,y)$ is a smooth circular window function, for example a Gaussian.

This change can be concisely written as

$$E(u,v)=[u \quad v]M\begin{bmatrix} u \\ v \end{bmatrix} \tag{2}$$

where *M* is a 2*2 matrix computed from image derivatives:

$$M=\sum_{x,y} w(x,y)\begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \tag{3}$$

Let λ_1, λ_2 be the eigenvalues of matrix *M*. Then, if λ_1 and λ_2 have large positive values, a corner is found at the position (*x*,*y*).

2. SIFT

SIFT or Scale Invariant Feature Transform [4] is a feature detector which is invariant to image rotation and scale. SIFT consists of four main steps to find the features.

A. Scale-space extrema detection

This step is to find the salient points in the input image. First the Difference of Gaussian images are created by building an image pyramid of Gaussian-blurred image of the input image with different scales. Then in these DoG images the extrema are found by looking for each pixel in eight neighborhood pixels of the current scale image and in nine neighborhood pixels in the neighboring scale images. If the pixel is a minima or maxima, the pixel is a candidate key point.

B. Key point localization

Once a key point candidate has been found in the previous step, the next step is to eliminate the low contrast key points and the key points which have a strong edge response. For low contrast key points, we can eliminate them using the Taylor expansion of the Difference of Gaussian scale space function. For the key points which have a strong edge response, we can eliminate them using the principal curvature. The principal curvature can be computed from the eigenvalues of the Hessian matrix.

C. Orientation assignment

By assigning a dominant orientation to each key point, the key point descriptor can be represented relative to this orientation and therefore achieve invariance to image rotation. An orientation histogram is formed from the gradient orientations of sample points within a region around the key point. The peaks in this orientation histogram are the dominant orientations. If there are multiple dominant orientations, new key points are added at the same location and scale as the original key points.

D. Descriptor

This step is to compute a descriptor for each key point. The descriptor is formed from a vector containing the values of all the orientation histogram entries. Usually, for one key point, we use a 4*4 array of orientation histograms with 8 orientation bins in each to describe it. Therefore, the descriptor of each key point is a 4*4*8 = 128 element feature vector.

3. SURF

SURF or Speeded Up Robust Features [5] is developed to speed up the feature detecting process using SIFT as a basic algorithm.

A. Detector

The detector is based on the Hessian matrix. It uses a very basic approximation of the Hessian matrix which relies on integral images to reduce the computation time. Therefore it is called the 'Fast-Hessian' detector.

Given a point $X=(x,y)$ in an image *I*, the Hessian matrix $H=(X,\sigma)$ at *X* with a scale σ is defined as follows

$$H(X,\sigma)=\begin{bmatrix} L_{xx}(X,\sigma) & L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma) \end{bmatrix} \tag{4}$$

where $L_{xx}(X,\sigma)$ is the convolution of the Gaussian second order derivative with the original image *I* at point *X*, and similarly for $L_{xy}(X,\sigma)$ and $L_{yy}(X,\sigma)$. The Gaussian second order derivative approximation is further approximated using box filters. The approximations are denoted by D_{xx}, D_{xy} and D_{yy}. The approximation determinant of the Hessian matrix is given by

$$\det(H_{approx})=D_{xx}D_{yy}-(0.9D_{xy})^2 \tag{5}$$

The approximated determinant of the Hessian matrix represents the blob response in the image at location *X*. These responses are stored in a blob response map.

A pyramid is created by applying the box filters which have different sizes on the original image.

The localization of the key point is done by finding the local maxima over a 3*3*3 neighborhood. The maxima of the approximated determinant of the Hessian matrix are then interpolated in the scale and image space.

B. Descriptor

The descriptor is based on the neighborhood of the location of the key point. A square region is centered at the key point, and oriented along the dominant orientation. The size of the window is 20. This region is split up in 4*4 sub-regions. For each sub-region, Haar wavelet responses d_x and d_y are computed in horizontal and vertical directions. The wavelet responses are summed up over each sub-region to form a first set of the feature vector. Also added to this vector are the sums of the absolute values of the responses. Therefore, each sub-region has a four-dimensional description vector of the form

$$v = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|) \qquad (6)$$

This results in total vector length of 4*4*4 = 64. At last, the descriptor is converted into a unit vector.

4. FAST

FAST or Features from Accelerated Segment Test [6] is built on the SUSAN detector. The most promising advantage of FAST corner detector is its computational efficiency.

A. Segment test detector

FAST corner detector uses a circle of 16 pixels around the corner candidate *p* to classify whether a candidate point *p* is actually a corner. Each pixel in the circle is labeled from integer number 1 to 16 clockwise. If a set of *N* contiguous pixels in the circle are all brighter than the intensity of candidate pixel *p* plus a threshold value *t* or all darker than the intensity of candidate pixel *p* minus threshold value *t*, then *p* is classified as a corner. *N* is chosen to be 12 because it admits a high-speed test which can be used to exclude a very large number of non-corners: the test examines only the four pixels at the four compass directions.

B. Machine learning of a corner detector

The ID3 algorithm is used to select the pixels which yield the most information about whether or not the candidate pixel is a corner. This is measured by the entropy of the positive and negative corner classification responses based on this pixel. The process is applied recursively on all three subsets and terminates when the entropy of a subset is zero. The decision tree resulting from this partitioning is then converted into C-code, creating a long string of nested if-then-else statements which is compiled and used as a corner detector.

C. Non-maximal suppression

Finally non-maximal suppression is applied on the sum of the absolute difference between the pixels in the circle and the center pixel.

III. EXPERIMENTAL RESULTS

In this section, we carry out 2D and 3D objects classification, which is described in [2], using these four feature detectors. For the experimentations we use two different image sets. The four feature detectors are all tested with those image sets. The results from these experiments contain information about the detection time, number of detected feature points and the precision of 2D and 3D classification.

First, we analyze the detection time and the number of detected feature points. **Figure 1** shows the detection time and **Fig. 2** gives the number of detected feature points. From these two figures, we see the FAST detector can find a lot of feature points, and also cost the least time. It seems the FAST detector is the best detector in these four detectors.

Next, we evaluate the precision of 2D and 3D objects classification. The result of the evaluation is shown in **Fig. 3**. From Fig. 3, we see SIFT and SURF are the best detectors for this 2D and 3D objects classification method. So why can the FAST detector

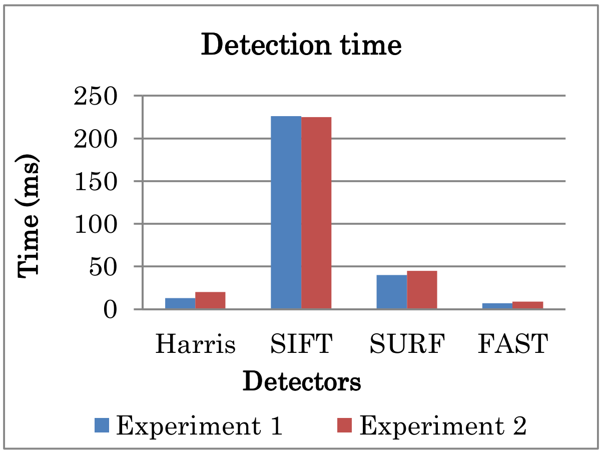


Fig.1. The detection time.

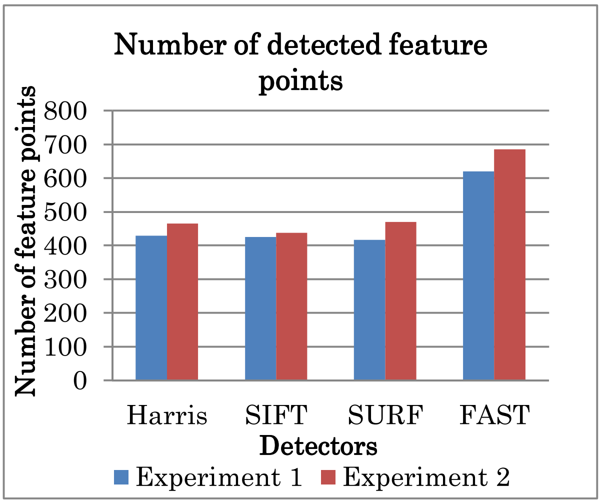


Fig.2. The number of detected feature points.

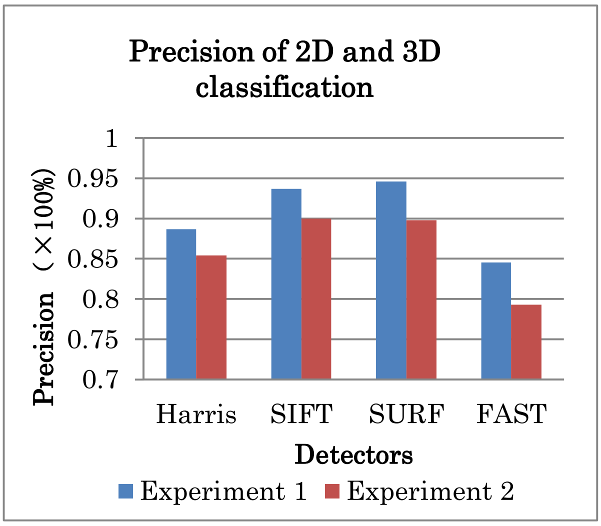


Fig.3. The precision of 2D and 3D classification.

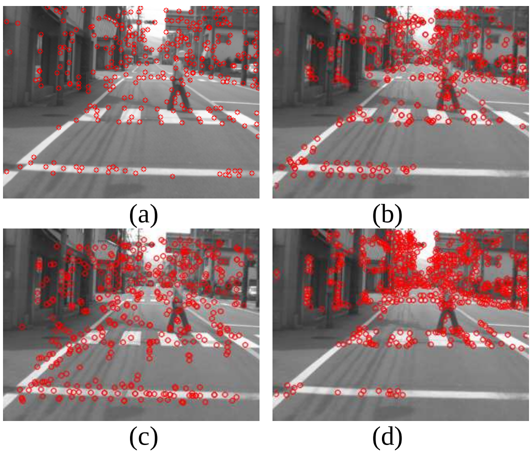


Fig.4. The detected feature points. (a) Harris detector, (b) SIFT detector, (c) SURF detector, (d) FAST detector.

find the most feature points, but get the lowest precision? We can find this reason in **Fig. 4**. Figure 4 shows the detected feature points. Fig.4 (d) shows the detected feature points using FAST detector. In this result, the detected feature points are concentrated in the upper part of the image: Just a few are located in the road region. But in the 2D and 3D classification method, we need to use the feature points which are located in the road region to estimate the parameters of the road plane. So if we use the FAST detector, we cannot estimate the equation of the road plane accurately, this results in the low classification precision.

According to the precision of 2D and 3D classification (shown in Fig. 3), although SIFT and SURF have almost the same precision, according to the detection time (shown in Fig. 1), it's easy to get this conclusion: The SURF detector is the most suitable feature detector for our obstacles detection method.

IV. CONCLUSION

In this paper, we performed comparison among four most commonly used feature detectors; Harris, SIFT, SURF and FAST detectors. The experiments are done using our 2D and 3D objects classification method [2]. The experimental results are compared in different ways, and we find that the SURF detector is the most suitable feature point detector for our 2D and 3D objects classification method from the points of the detection time and the precision of 2D and 3D classification. This will also improve the accuracy of obstacles detection (described in [1]) ultimately.

REFERENCES

[1] Qian S, Tan JK, Kim H, Ishikawa S, Morie T, Shinomiya T(2013), Road region estimation and obstacles extraction using a monocular camera. International Journal of Innovative Computing, Information and Control, 9(9): 3561-3572.
[2] Qian S, Tan JK, Kim H, Ishikawa S, Morie T, Shinomiya T (2013), Classifying 2D and 3D objects on a road employing the road plane. Proc. SICE Annual Conference 2013, 1689-1692.
[3] Harris C, Stephens M (1988), A combined corner and edge detector. Proc. Alvey Vision Conference, 147-151.
[4] Lowe DG (2004), Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2): 91-110.
[5] Bay H, Tuytelaars T, Gool LV (2006), SURF: Speeded up robust features. Proc. the European Conference on Computer Vision, 404-417.
[6] Rosten E, Drummond T (2006), Machine learning for high-speed corner detection. Proc. the European Conference on Computer Vision, 430-443.

A Decision Making System of Robots Introducing a Re-construction of Emotions Based on Their Own Experiences

Shogo Watada[1], Masanao Obayashi[2], Takashi Kuremoto[3], Shingo Mabu[4] and Kunikazu Kobayashi[5]

[1-4]Yamaguchi University, Japan, [5]Aichi Prefectural University, Japan

{[1]t003we, [2]m.obayas, [3]wu, [4]mabu}@yamaguchi-u.ac.jp, [5]kobayashi@ist.aichi-pu.ac.jp

Abstract: We already proposed a decision making system for autonomous robots, which is based on a Markovian emotional model proposed by Banik. In our proposed system, decision making of the robot is designed automatically by two processes. The first is a construction of emotions based on learning of the input stimuli. The second is optimization of behavior selection policy for emotions by optimization of system parameters which define behavior selection probability of each emotion. In our above previous studies, a construction of emotions based on learning of the input stimuli was processed using learning samples of input stimuli on offline in advance. In this study, we introduce an experiences-based reconstruction of emotions to our decision making system as an approach to an automatic construction of decision making online learning.

Keywords: Emotion model, Decision-making, Optimization problem.

I. INTRODUCTION

There exists much literature on the study which focuses on the concept of emotions that animals have, for developing autonomous systems in various fields[1]. We have already proposed a decision making system for autonomous robots, which is based on a Markovian emotional model proposed by Banik[2]. The Markovian emotional model is a probabilistic model that consists of a finite state machine[1].

In our proposed system, a decision making system of the robot is designed automatically by two processes. The first is a construction of emotions based on learning of input stimuli. We proposed use of a SOM (Self-organizing map) to associate input stimuli with internal parameters that represent emotional states. The second is optimization of behavior selection policy for emotions by optimization of system parameters which define behavior selection probability of each emotion.

In our previous studies, a construction of emotions based on learning of the input stimuli was designed only using learning-samples of input stimuli on offline in pre-learning. Therefore, it made useless emotional areas corresponding to the unrealistic input stimuli in the SOM.

In this study, we introduce an experiences-based reconstruction of emotions in our decision making system as an approach to an automatic construction of decision making online learning. That is, we thought that more effective associations with input stimuli and parameters decided by considering actual frequency and measurement values of input stimuli is achieved by online learning.

We confirmed the usefulness of our method by computer simulation for an E-puck mobile robot. The task for the robot is to maximize floor areas which the robot has passed. In the simulation, the result showed that additional learning of input stimuli makes unevenness of frequency of use of each node in the SOM decreases and it also makes more efficient associations between input stimuli and emotional states.

This paper is organized as follows. Section 2 describes the Markovian emotional model proposed by Banik. In Section 3, the proposed behavior selection system is described. In Section 4, the simulation results of the proposed method are presented. Finally, Section 5 summarizes the effectiveness of the proposed method and future works.

Ⅱ. The Markovian emotional model

The Markovian emotional model referenced in this paper is a probabilistic model that consists of a finite state machine[1]. It consists of four basic emotions: joy, anger, fear and sadness. Emotional states are presented Y as shown in Eq.1.

$$Y = [y_{joy} \quad y_{anger} \quad y_{sadness} \quad y_{fear}] \qquad (1)$$

The emotional state transition at time k is given by Eq.2.

$$Y_{k+1} = CY_k \qquad (2)$$

Here C is the emotional state transition matrix, which can be expressed as Eq.3. $P_{A/B}$ is the probability of transition from state B to state A. $q_{*/*}$ is a prior probability of emotion transition, which should be set in advance (See Eq. 4). These values $P_{*/*}$ can be changed online by the influence of emotion-inducing factors: α, β, γ and δ for joy, anger, fear, and sadness, respectively. These factors vary depending on the internal state and information given by certain other robots and have the role to stimulate each emotion. For example, the probability of state transition from joy to other state expressed by Eq.4.

$$C = \begin{bmatrix} P_{joy/joy} & P_{joy/anger} & P_{joy/sad} & P_{joy/fear} \\ P_{anger/joy} & P_{anger/anger} & P_{anger/sad} & P_{anger/fear} \\ P_{fear/joy} & P_{fear/anger} & P_{fear/sad} & P_{fear/fear} \\ P_{sad/joy} & P_{sad/anger} & P_{sad/sad} & P_{sad/fear} \end{bmatrix} \quad (3)$$

$$\left.\begin{aligned} P_{anger/joy} &= q_{anger/joy} + (\beta - \alpha) q_{anger/joy} \\ P_{sad/joy} &= q_{sad/joy} + (\delta - \alpha) q_{sad/joy} \\ P_{fear/joy} &= q_{fear/joy} + (\beta - \alpha) q_{fear/joy} \\ P_{joy/joy} &= 1.0 - (P_{anger/joy} + P_{sad/joy} + P_{fear/joy}) \end{aligned}\right\} \quad (4)$$

Ⅲ. Proposed behavior selection system

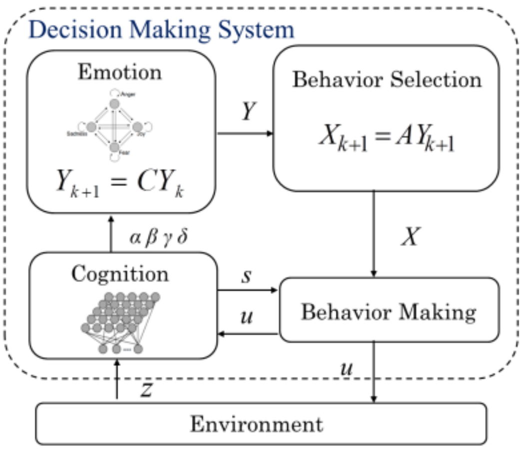
We have already proposed a decision making system of a robot, which is using a Markovian emotional model proposed by Banik. The overall structure of the system can be shown in Fig.1. The system consists of four modules: Cognition, Emotion, Behavior-selection, and Behavioral-making. Z is input stimuli, α, β, γ and δ are emotion-inducing factors, X_k is a behavior selection probability vector at time k, s is robot states and u is a control input for the robot.

1. The function of each module

The function of each module is as follows; the Cognition module transforms input stimuli Z such as sensor values into emotion-inducing factors α, β, γ and δ. The Emotion module is a Markovian emotional model proposed by Banik, and this module updates emotional states Y using α, β, γ and δ. The Behavior-selection module determines the behavior selection probability vector X using Y by Eq.5.

$$X_{k+1} = AY_{k+1} \quad (5)$$

Elements of X are selection probability of each behavior of the robot, for example, X is defined such as Eq.6. T



Z: Measurement Y: Emotional state
C: Emotional state transition matrix A: Behavioral state transition matrix
s: Situation Behavior state α, β, γ, δ: Emotion-inducing factors

Fig.1. A proposed behavior selection system

means transpose of a vector.

$$X = \begin{bmatrix} x_{drive} & x_{left} & x_{right} & x_{Back} \end{bmatrix}^T \quad (6)$$

A is a transition matrix consists of probabilities $P_{behavior/emotion}$ which means a probability of transition from *emotion* to *behavior*, and represented by such as Eq.7.

$$A = \begin{bmatrix} P_{Drive/joy} & P_{Drive/anger} & P_{Drive/sad} & P_{Drive/fear} \\ P_{Left/joy} & P_{Left/anger} & P_{Left/sad} & P_{Left/fear} \\ P_{Right/joy} & P_{Right/anger} & P_{Right/sad} & P_{Right/fear} \\ P_{Back/joy} & P_{Back/anger} & P_{Back/sad} & P_{Back/fear} \end{bmatrix} \quad (7)$$

Finally, the Behavior-making module generates the control input u based on X taking into account states of the robot.

2. Clustering of input stimuli in Cognition module

In our proposed system, we used a SOM in the Cognition module in order to cluster input stimuli such that similar input stimuli correspond with similar emotion-inducing factors. In this method, we regard the position of each node in the SOM as value of emotion-inducing factors. Since it is considered that groups of similar input vectors are compressed in the output layer as weight vectors after learning, the value of emotion-inducing factors is uniquely determined to a stimulation input vector if we consider that the two-dimensional output layer is a two-dimensional map consisting of emotional inducing factors.

The structure of the Cognition module in this study is shown in Fig.2. Each node in the input layer of the SOM corresponds to N input stimuli. The Cognition module constitutes the association between emotions and input stimuli by clustering of input stimuli to the output layer. Then, after learning, the Cognition module can convert input stimuli into emotion-inducing factors by determining a winner node and normalizing the position of it into range from 0 to 1.

In previous studies, learning of the SOM was processed only using learning-samples of input stimuli in advance. Therefore, it generated useless emotional areas corresponding to the unrealistic input stimuli in the SOM. In this study, we propose that the association between input stimuli and emotion-inducing factors becomes more efficient by additional learning of realistic input stimuli on online during the task.

3. Optimization of behavior policy by GA

In our Behavior-selection module, the policy of behavior selection depends on transition matrix A. So, it becomes possible to make various methods for behavior selection by adjusting parameters being included in A. In this work, we use GA for setting A. GA is a meta-

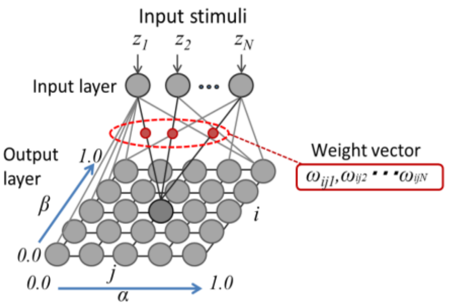


Fig.2. New cognition module proposed in this paper.

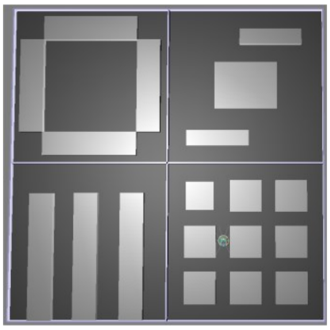


Fig.3. Environments used in simulations.

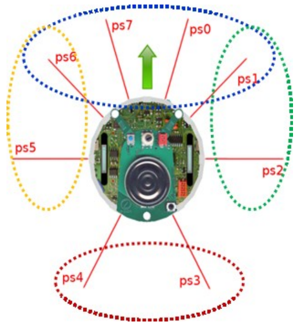


Fig.4. Sensors of e-puck

heuristic algorithms for searching the approximate solution. The detail explanation about GA is omitted.

Ⅳ. Simulation

For evaluating our system, we simulated the problem of maximizing areas that the robot passed. In such as house cleaning robots, it is demanded that avoiding obstacles using own sensors, robots pass more large areas in the environment. In this work, we showed our proposed system generate a behavior policy automatically that a two-wheel mobile robot passes larger areas in fixed time. The simulation for an e-puck robot was carried out using a Webots simulator which make possible to simulate in a very realistic way.

1. Simulation setting

The simulation was executed through four environments with different structures which are shown in Fig.3. Each environment is a square of 1 m on a side, and there are walls or obstacles (brighter areas in Fig.3). The e-puck is equipped with 8 infrared sensors which measure proximity of obstacles in a range of 4 cm. We defined four behaviors of the robot; straight, left, right and back, and defined the motor torques of left and right wheels on each behavior as follows: straight: [100, 100]; left [-50,50]; right: [50, -50]; back: [-100,-100]. The Behavior-making module calculates output torques by weighted syntheses of torques with rate of each behavior probability in the probability vector X.

The number of input stimuli is 6, and there are shown in Table1. The pre-learning of SOM used learning samples which were generated randomly based on prior knowledge of devices such as the range of the sensor.

On the basic setting of simulation, the pre-learning of the input stimuli runs at first. Then, the decision making policy is generated by optimization of the transition matrix A using GA. In the GA, each individual executes the task in all 4 environments, and the fitness value of the individual is determined as the sum of above 4 areas the robot passed. The setting of the SOM and the GA are shown in Table 2.

Table 1. Input stimuli of the robot.

z_0	The average distance value of four sensors on front (see Fig.4).
z_1	The average distance value of two sensors on right (see Fig.4).
z_2	The average distance value of two sensors on left (see Fig.4).
z_3	The average distance value of two sensors on rear (see Fig.4).
z_4	The total moving distance of robot from 12 seconds ago.
z_5	The total rotation angle of robot from 12 seconds ago.

Table 2. Parameters setting of SOM and GA.

SOM		GA	
Number of learning	500	Number of Individuals	24
Number of nodes	30×30	Final Generation	50
Learning coefficient η	0.0001	The limited time in one environment	600 (s)
$\sigma(0)$ *	24.0		

* The neighborhood function: $\sigma(step) = \sigma(0)(1.0 - step/maxStep)$

2. Simulation result

In this simulation, the additional learning runs while the generation is young in the process of GA after pre-learning. Altogether, the SOM is updated on online using input stimuli which are occurred in first 10 generations of 50 generations by using input stimuli which occur in the task. Since updating of the SOM is running every steps, the total number of updates is 576,000.

Since the position of each node of the SOM correspond to values of emotion-inducing factors, it is easy to understand the relationship between input stimuli and emotions by making the emotional map. Fig.5 shows emotional map made from the SOM after the additional learning. The input stimuli are represented in each circle on the map. A distance to obstacle is represented as size of an ellipse, and magnitudes of a moving distance and a rotation angle are represented as the length of the up arrow and the horizontal direction of the arrow, respectively. For example, Joy and Anger correspond to situations of input stimuli such that there are no obstacles around the robot, and these situations occur frequently during the task. In the pre-learning, SOM learns equally all situations such as the robot is surrounded by wall. However, input stimuli corresponding to the situation of going straight or turning left are more often. This result shows that the SOM reflects realistic input stimuli by additional learning.

Observing behaviors of the robot using final parameters in GA, the behavior policy were such as follows: go straight bending slightly the left when there are no obstacles around the robot; turn left when there're obstacles in front or right side of the robot; turn right when there're obstacles in left size of the robot. In our method, the behavior of are mentioned is explained by parameters of A shown in Eq.8. For example, in probabilities of transition from Joy which correspond to a free space according to Fig.5, $P_{dri/joy}$, 0.64, and $P_{lef/joy}$, 0.23, are higher than others. Therefore, the robot went straight bending slightly the left when the robot is in the open space.

$$A = \begin{bmatrix} P_{Dri/joy} & P_{Dri/ang} & P_{Dri/sad} & P_{Dri/fea} \\ P_{Lef/joy} & P_{Lef/ang} & P_{Lef/sad} & P_{Lef/fea} \\ P_{Rig/joy} & P_{Rig/ang} & P_{Rig/sad} & P_{Rig/fea} \\ P_{Bac/joy} & P_{Bac/ang} & P_{Bac/sad} & P_{Bac/fea} \end{bmatrix} = \begin{bmatrix} \underline{0.64} & 0.44 & 0.00 & \underline{0.34} \\ \underline{0.23} & 0.16 & \underline{0.80} & 0.12 \\ 0.12 & 0.21 & 0.00 & \underline{0.42} \\ 0.01 & 0.19 & 0.20 & 0.10 \end{bmatrix} \quad (8)$$

Next, the result using additional learning was compared with the result using only pre-learning. Fig.6 shows transitions of fitness values at two results in GA, and the vertical dashed line on 10 steps means the generation which additional learning finishes. It shows that the fitness value of the system using additional learning is higher than the other.

Fig.7 shows histograms of referenced number as a winner node in each system where white nodes were referenced frequently. In the case of only prior learning, areas of white nodes are concentrated, and it means that emotion-inducing factors changes little. On the other hand, in the case of additional learning, areas of white nodes are spread because that useless node areas of SOM decrease by additional learning. The performance of decision making process of the robot is improved by the more efficient association between emotion-inducing factors and input stimuli.

V. CONCLUSION

In our previous studies, a construction of emotions was processed only using learning-samples of input stimuli on offline. Therefore, it generated useless emotional areas in the SOM.

In this work, we proposed reconstruction of emotions by additional learning input stimuli obtained during the task. Then, associations between input stimuli and parameters more effective by additional learning, using actual input stimuli were made.

The proposed method was evaluated by a simulation of a problem which maximizes areas a robot passed. The result showed that since an additional learning

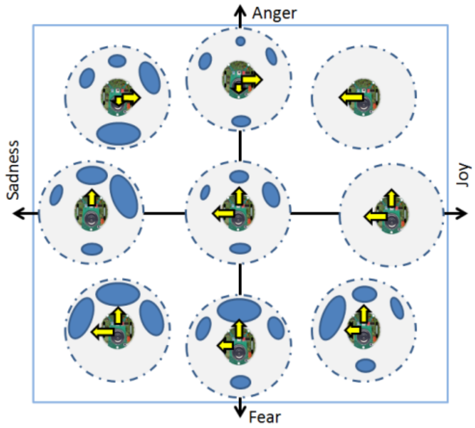


Fig.5. Emotional map.

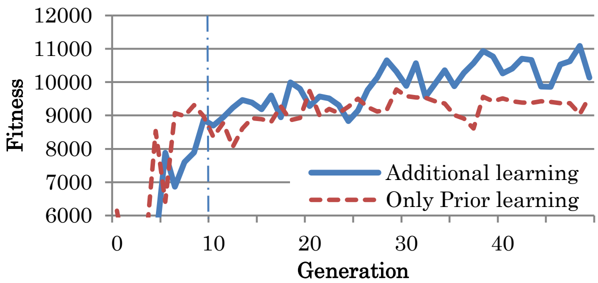


Fig.6. Transition of fitness value.

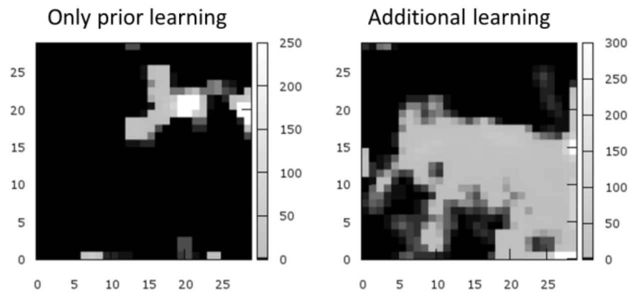


Fig.7. The histogram of the number of times each node is selected.

decreases useless emotional areas, it improved decision making of the robot. In addition, it was impossible to explain the decision-making policy of robot by analyzing generated the SOM and parameters of A.

For further work, we will attempt to consider a method which optimizes parameter setting on online and robot tasks that is possible to evaluate the system on online instead of additive learning of using GA.

REFERENCES

[1] Sajal Chandra, Banik, Keigo Watanabe, Kiyotaka Izumi (2008), Improvement of group performance of job distributed mobile robots by an emotionally biased control system", Artificial Life Robotics(12), pp.245-249

[2] Shogo Watada, Masanao Obayashi, Takashi Kuremoto, Kunikazu Kobayashi, Shingo Mabu (2013), Behavior selection method of robots based on a Markovian emotional model (in Japanese). IEEJ Transactions on Electronics, C, Vol.134, No.1

Visual Servoing and Sound Localization in a Surveillance Robot

So-Yeon Park, Yeoun-Jae Kim, and Ju-Jang Lee

Electrical and Electronic Engineering KAIST, Eu-Seong Gu Daehak-Ro 291 KAIST EE 4231
Tel : 82-42-350-5432; Fax : 82-42-350-5432
e-mail address: arashi@kaist.ac.kr, lethkim@kaist.ac.kr, jjlee@ee.kaist.ac.kr

Abstract: The purpose of this paper is to explain the control and implementation details of the visual servoing and sound localization in a tank shaped surveillance robot. With an eye in hand camera configuration, the key point of the visual servoing is to measure the depth to the target object. In this paper, we suggested the depth estimation technology by moving the camera in a lateral plane. In sound localization, the hardware and software details are fully explained to implement the sound localization technology. Especially the electronic filter design is implemented by utilizing the genetic algorithm.

Keywords: visual servoing, sound localization, surveillance, robot**.**

I. INTRODUCTION

The visual servoing technique can be divided into two categories. One is image based visual servoing (IBVS) and the other is position based visual servoing (PBVS). IBVS uses the image information directly for the visual servoing and PBVS extracts 3D position data from the image information and performs the visual servoing technique by the 3D position data. Most of the eye in hand shaped robot has only an eye camera on its manipulator, so, IBVS technique is used in this kind of robot. The main concern of IBVS technique is how to extract the depth information [1]. In most cases, authors deal with interaction matrix for IBVS and combined IBVS control with plant control algorithm. In this paper, we suggested a simple depth estimation algorithm without combining IBVS with plant control algorithm. The suggested algorithm is verified by experiment of the surveillance robot.

Comparing with the visual servoing, sound localization is not so popular technique. However, many authors researched this area [2]. The humans have only two ears but in robot, with two microphones it is difficult to precisely measure the direction of the sound source. Usually three or more microphone array is needed. In this paper, we implemented the state of the art sound localization technique with three microphone array on the surveillance robot. The detailed hardware and software implementation is explained later and experimental results shows that the sound localization technique is well implemented on the surveillance robot.

The rest of paper is organized as follows. Section 2 explains the surveillance robot in which the visual servoing and sound localization functionality is implemented. The visual servoing without pre-depth information is explained in section 3 and In section 4, the detailed explanation of sound localization technique in the surveillance robot is presented. Conclusion is drawn in section 5.

II. SURVEILLANCE ROBOT

The surveillance robot with which the visual servoi ng and sound localization functionality is implement ed is a tank shaped robot which performs surveilla nce work in normal time and rescue humans in em ergency time. The caterpillar is implemented in this robot instead of a wheel for its movability in outdo or environment and robustness against an external s hock. The length of the tank robot is about 1.5m a nd the width is about 1m. It has a manipulator whi ch has a 5 DOF(Degree of Freedom) for picking u p and investigating a suspicious object. The robot h as two cameras. One camera is mounted at the top of the tank body for watching distance scenery. Th e other camera is mounted at the bottom of the ma nipulator gripper for visual servoing purpose. Overal l appearance of the surveillance robot is in Fig. 1.

Figure 1 The surveillance robot-overall view

.III. VISUAL SERVOING WITHOUT PRE-DEPTH INFORMATION

The camera at the end-effector of the surveillance robot is a single lens camera, not a stereo camera. For this reason, with the vision signal from the single lens camera, the surveillance robot cannot determine the 3D position of the target object. Therefore some algorithm to determine the depth information is needed. We divided the visual servoing movement into two parts. First part is to move the manipulator in lateral direction to determine the depth length and the second part is to move the manipulator forward to the target object with this determined depth information.

1. Image-based visual servoing

In order to place end-effector of the manipulator near the target, the position of the target object must be determined by using the visual information of the camera.

$$X_c = \frac{(u-c_u)Z_c}{f\alpha} \qquad Y_c = \frac{(v-c_v)Z_c}{f} \tag{1}$$

In the equation (1), u and v are the coordinates of the target object in pixel units and c_u and c_v are the principal point in pixel units. X_c , Y_c , Z_c is 3-D coordinates of the object with respect to the camera frame. The 3-D coordinates of the object with respect to the camera frame can be transformed to the 3-D coordinates of the object with respect to the inertial frame by equation (2).

2. Estimating the depth

As mentioned previously, the objective of the visual servoing is to move the end-effector of the manipulator near to the target object. That is, the coordinates of the

$$\begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{bmatrix} = {}^0A_1\,{}^1A_2\,{}^2A_3\,{}^3A_c\,{}^cA_4\,{}^4A_5 \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{2}$$

center of target object (u,v) expressed in pixel units should be located at principal points (c_u , c_v) at goal position. Therefore, moving distance of the target object with respect to camera frame ΔD^c_{target} can be computed as equation (3) when the manipulator end-effector moves in lateral direction.

$$\Delta D^c_{target} = \begin{bmatrix} \Delta X^c_{target} \\ \Delta Y^c_{target} \\ \Delta Z^c_{target} \end{bmatrix}$$
$$\Delta X^c_{target} = \frac{(c_u-u)Z_c}{f\alpha} \qquad \Delta Y^c_{target} = \frac{(c_v-v)Z_c}{f} \qquad \Delta Z^c_{target} = 0 \tag{3}$$

As in Fig.2, the principal point of the camera (black dot) moves in the reverse direction of the target object (blue dot). Therefore, the moving distance of the camera with respect to camera frame ΔD^c_{camera} is computed as in equation (4)

$$\Delta D^c_{camera} = \begin{bmatrix} \Delta X^c_{camera} \\ \Delta Y^c_{camera} \\ \Delta Z^c_{camera} \end{bmatrix}$$
$$\Delta X^c_{camera} = -\Delta X^c_{target} = \frac{(u-c_u)Z_c}{f\alpha} \qquad \Delta Y^c_{camera} = -\Delta Y^c_{target} = \frac{(v-c_v)Z_c}{f} \qquad \Delta Z^c_{camera} = 0 \tag{4}$$

Then, the moving distance of the camera with respect to the base frame ΔD^0_{camera} can be computed as in equation (5)

$$\begin{bmatrix} \Delta D^0_{camera} \\ 1 \end{bmatrix} = {}^0A_1\,{}^1A_2\,{}^2A_3\,{}^3A_c\,{}^cA_4\,{}^4A_5 \begin{bmatrix} \Delta D^c_{camera} \\ 1 \end{bmatrix} \tag{5}$$

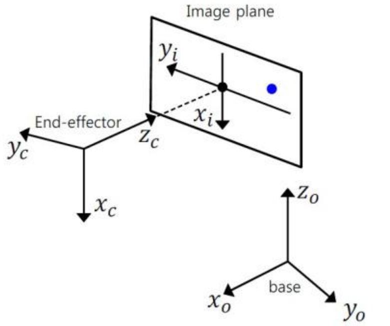


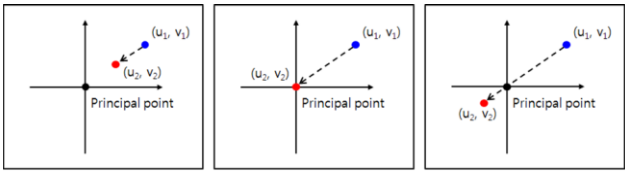
Figure 2. Base and camera (end effector) frame

As mentioned previously, since Z_c cannot measured directly by single lens camera, it must be estimated. In our research, moving repetitively the end-effector in lateral direction and returning to the initial position, we estimate the Z_c .

As seen in Fig. 3, the lateral movement of the target point depends on estimated Z_c. (u_1, v_1) is the position of the target object in the image plane before the lateral movement of the manipulator and (u_2, v_2) is the position of the target object in the image plane after the lateral movement of the manipulator. As can be seen in Fig. 3 (a) when the estimated value is smaller than the actual value, the location of the target object before and after moving the manipulator is at the same quadrant in the image plane. However, when the estimated value is bigger than the actual value, the location of the target object before and after moving the manipulator is at the different quadrant in the image plane as in Fig. 3 (c). If the estimated value matches the actual value, the location of the target object is just at the principal point as can be seen in Fig 3 (b). Therefore, the control policy of Z_c estimation is as follows.

-If target object before and after moving is located in a different quadrant, the estimated Z_c is bigger than the actual value. And it should be decreased.

-Otherwise the estimated Z_c is smaller than the actual value. And it should be increased.



(a) (b) (c)

Figure 3 Lateral movement of the end-effector

IV. SOUND LOCALIZATION ALGORITHM

For the sound localization, a hardware capturing the sound signal and an algorithm calculating the direction of arrival is needed. As for the hardware, it can be largely divided into three groups, microphone array, analog circuit and digital data acquisition device. Regarding to the algorithm, it can be largely divided into two part, signal processing and direction of the arrival estimation. As for the algorithm used in the sound localization, it can be divided into two part, signal processing and direction of arrival (DOA) estimation. Theoretical background for the algorithm can be obtained in [3].

1. Signal processing

In signal processing, a sound signal obtained from the hardware part, which went through some editing, goes through further process. It is transformed into frequency domain, filtered by more sophisticated filter. Because ideal infinite time signal cannot be used in the real-time application, time domain signal must be chopped up appropriately. Frame duration between 10 and 50ms is recommended. In this paper, we used 30ms duration. Framing can be presented as the multiplication of the infinite signal with a rectangular window as follows

$$x_n(t) = x(t) w_{rectangle}(t,n) = \begin{cases} x(t), & |t - nT| \le \dfrac{T}{2} \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

Where T stands for time duration. After extracting frame, the signal is converted from time-domain to frequency-domain using Fourier transform (FT) as follows.

$$X(f) = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi ft} dt \quad (7)$$

When we change the framed signal into frequency domain, there exists a signal leakage because the rectangular window mentioned above has an infinite spectrum. Therefore in order to reduce error due to the finite duration processing, weighting function which falls sharply as frequency bin increases has been suggested. Hann weighting, one of the most widely used window, is used in this paper.

$$w_{Hann}(t) = \begin{cases} 0.5(1+\cos\dfrac{\pi t}{T}), & |t| \le T \\ 0, & |t| > \mathrm{T} \end{cases} \quad (8)$$

Moreover, most of meaningful sound, that is, audible sound lies between 20Hz and 20kHz in frequency band. Therefore in order to reduce possible noise and concentrate on audible sound, c-weighting is also used in the project.

$$G_c(s) = \frac{5.918 \times 10^9 s^2}{(s+129.4)^2 (s+76655)^2} \quad (9)$$

When estimating direction of arrival, voice activity detector (VAD) detects the presence of voice in a mixed signal consisting of voice plus noise. Its application range can vary from simple binary classification, “noise only” or “voice present”, for entire frame to presence probability for every frequency bin. VAD itself can be usefully applied but it can be included in most of noise suppression. VAD detects presence of voice based on an assumption that noise changes more slowly than voice. When capturing the signal through the microphone, the captured signal has signal x(t) and noise d(t) at the same time. Since signal and noise are unobservable we have to estimate original signal through observed signal and some a priori-knowledge.

In noise suppression, noise reduction process can be considered as an application of a time-varying, non-negative, real-valued gain H to very frequency bin of observed signal in order to get the estimation of original signal. While the noise filters included in the hardware part passes or stops the signal according to designed pass band and stop band frequencies, noise suppression learns background noise in a working environment. For example, if the noise filter from hardware is set to pass frequency band between 100Hz and 10 kHz, it cannot filter the siren signal which mainly lies in 1 kHz. But noise suppression can learn siren signal as noise and remove the appropriate frequency bins. In noise suppression, at first, time-domain signal x(t) passes through framing, weighting, and fast FT (FFT). Then converted frequency-domain signal is divided into phase component and magnitude component. Phase component is kept during the noise suppression. Meanwhile after magnitude component passes through per-frame soft VAD, rough noise model is built. Again magnitude passes through per-bin soft VAD. Then precision noise model is built. And finally suppression rule is made based on rough model and noise is removed. Finally if needed, magnitude and phase get together and after inverse FFT, estimated time-domain signal can be obtained.

2. Direction of arrival estimation

In this paper, we use an interaural time difference in order to estimate sound orientation. It is caused by the difference in distance where the sound must travel to reach each microphone. Theoretically time difference between two microphones is represented as follows.

$$\tau = \frac{d \sin\theta}{c} \qquad (10)$$

Meanwhile time difference in two microphones can be practically measured using cross-correlation function of two signals. Many of the methods which have been suggested so far, and a generalized cross-correlation using phase transform (PHAT) weighting is used in the project as follows.

$$R_{12}(t) = iFFT[\psi_{PHAT} \cdot X_1(f) X_2^H(f)]$$
$$\psi_{PHAT}(f) = \frac{1}{|G_{X_1X_2}(f)|} \qquad (11)$$

PHAT weighting eliminates magnitude and gives equal weight to the phases in each frequency been. Therefore it can work in a reverberant environment.

As for the microphone array used in the project, there exist three pairs of time difference between microphone 0 and 1, between 1 and 2, and between 2 and 0. Time difference means distance difference between source and microphones. Therefore stands for time difference microphone j based on microphone I as follows.

$$d_{ij} = \left\|\overline{M_j S}\right\| - \left\|\overline{M_i S}\right\| \qquad (12)$$

We divided 2-dimensional space into 3 areas as shown in Figure 4.

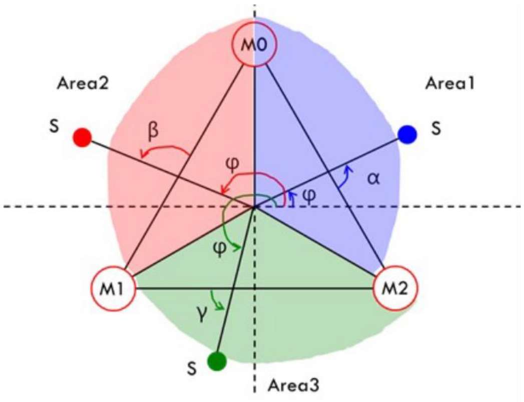


Figure 4. Microphone deployment

V. CONCLUSION

In this paper, we implemented visual servoing without depth information and sound localization algorithm in a surveillance robot.

ACKNOWLEDGEMENT

This work was supported by the Industrial Strategic technology development program, 10045252, "Development of Robot task intelligence technology that can perform task more than 80% in inexperience situation through autonomous knowledge acquisition and adaptation knowledge application" funded By the Ministry of Trade, industry & Energy (MI, Korea)

REFERENCES

[1] H. Fujimoto (2003), “Visual servoing of 6 dof manipulator by multirate control with depth identification”, Proceedings of 42nd IEEE Conference on Decision and Control, pp. 5408-5413.

[2] K. Nakadai, H. G. Okuno, and H. Kitano (2002), “Real-time sound source localization and separation for robot audition”, Proceedings IEEE International Conference on Spoken Language Processing, pp. 193-196

[3] Ivan Tashev (2009), Sound Capture and Processing, John Wiley and Sons.

Study on the improvement of flexibility for an industrial robot based on machine vision

Jiwu Wang 1, Xianwen Zhang 1, Weining Zhang 1, Sugisaka Masanori 2
1 Department of Mechanical and Electrical Engineering, Beijing Jiaotong University, Beijing, China，100044
2 Nippon Bunri University, Oita 870-0397, Japan

1 jwwang@bjtu.edu.cn
2 sugisaka3@mocha.ocn.ne.jp

Abstract: In order to improve the applications for an industrial robot, it is necessary to increase its flexibility and control accuracy. Many algorithms and control strategies are developed, in which machine vision gets more and more attention. This is because the CCD camera can provide more information than other sensors, which makes the whole control process simple. But it requires that the target position should be extracted accurately and robustly. Moreover, the frequency of image processing should be satisfied with the control requirements. Here an industrial robot arm is designed and set up for experiment simulation with machine vision. The target extraction method with fusion of color space and moment invariants is developed. The experiments results showed that the developed image processing algorithms are robust, and the flexibility of an industrial robot can be improved by machine vision.

Keywords: Arm robot, moment invariants, image processing**.**

I. INTRODUCTION

With the development of new technology, the requirements of product quality and production efficiency get higher, and experience workers can not be trained in short time and their number are not enough. With more and more production lines are used in modern companies, many automation machines are required to operate like a skill worker. Compared with a worker, the robot seldom makes mistakes if its working conditions are not changed. And its efficiency and reliability is higher than a worker. Especially in the danger or high working intensity environment, it is more appropriate[1]. Then various robots are applied in the most advanced company.

Generally, most robots can only be used for some simple, repeating and single jobs, which is not appropriate to the further applications for its high cost. On the other hand, the repeat positioning accuracy can not be satisfied if a robot arm is used for a longer time[2]. Although some feedback or sensor fusion technology is adopted, the effective solutions are not enough. .

With the development of the CCD camera technology, machine vision gets more attention. The CCD camera not only works like human being eyes, but also can provide more information than other sensors. With the pattern recognition, the CCD camera can provide multiple target positions in a captured image. Moreover, it can calculate the end position of the robot arm in a captured image if necessary. In a word, with a CCD camera, the previous problems for the arm robot can be solved[3] .

In order to get robust and reliable image processing results, the lighting conditions and image processing efficiency should be considered carefully. In this paper, considering the practical applications, the fusion technique of true color space and moment invariant is applied in the pattern recognition[4].

The structure of the arm robot is shown as figure 1. A is the CCD camera, B is the control box, C is a step motor to control the upper arm, D is a step motor t control the forearm, E is a step motor to control the hand, F is a support just like the shoulder, G is the upper arm, H is the forearm. The length of the upper arm and forearm is 400mm. The height between the camera and ground is 700mm.

The relationship among all hard wares is illustrated in Figure 2. The CCD camera is used to determine the target position. Three motors (C, D, E) is controlled by the PC through motion control card. One control process is: 1) After the CCD camera detect one target, the target will be extracted by the image processing algorithms in PC, and its position will be calculated. 2) Based on the determined target position, three motors will run together and make the hand reach the target.

3) The arm go to its original position, and one movement cycle is performed.

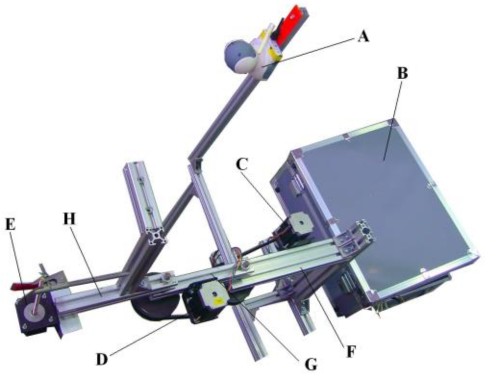
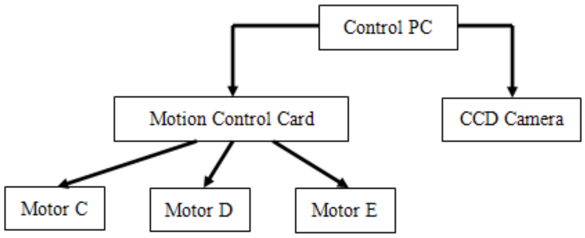


Fig.1. The illustration of the structure of the arm robot



.Fig.2. The illustration of the relationship about all hard wares

The developed software system is shown in figure 3

Fig.3.The illustration of the developed software interface

II. Target extraction method

In order to extract targets from the captured image at a higher speed, the color based target detection method is developed. At the same time, to reduce recognition errors, the moment invariants are applied to separate the previous extracted objects again[5].

Although the triple components (R, G, B) in a common RGB image can be used to extract the target, it is more difficult to determine an appropriate color measure range suitable for different lighting conditions. This will bring some mistakes.

In order to reduce the influence of different lighting conditions, a color model in the chromatic color space was built[6]. Chromatic colors, also known as pure colors in the absence of luminance, are defined by a normalization of the triple components (R, G, B). The process to build guideline color model is as follows:

1) Take common RGB color images of the target under different lighting conditions as sample images, and extract the typical RGB target colors manually from these sample images as sample colors.
2) Transform these sample colors from the RGB color space to the chromatic color space by a normalization process. Here color green is redundant because $r+g+b=1$
3) Calculate the mean values and covariance of sample colors in the chromatic color space.

where, x is a pair of normalization components (r, b) for each pixel in the chromatic color space, $\bar{x}$ is the mean value and C is the corresponding covariance.

4) The possibility of the target color for one

$$r = R/(R+G+B)$$

$$b = B/(R+G+B)$$

pixel in the detected image can be calculated as follows

$$\bar{x} = \frac{\sum_{n} x}{n}$$

$$C = \frac{\sum_{n} (x-\bar{x})(x-\bar{x})^T}{n-1}$$

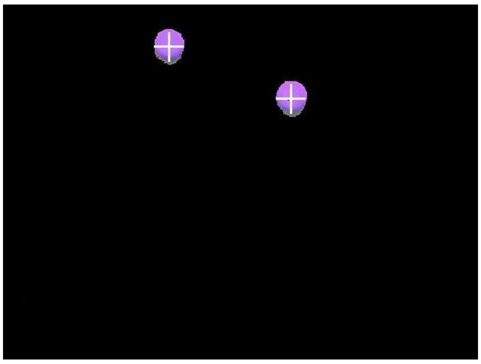
When we use this color model to extract target, we should firstly transform each pixel from the RGB color representation to the chromatic representation. Then a pair of chromatic color values for each pixel is input the color model to decide if that pixel has the same color as the target [7]. Then the geometric features of the targets calculated with

$$P_x = \exp[-0.5(x-\bar{x})C^{-1}(x-\bar{x})]$$

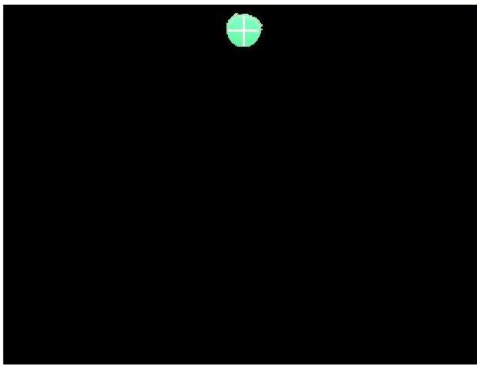
moment invariants, is used for further verification about the extracted regions

III. Multiple targets extraction and position control experiments

The experiments are performed by two steps: First, determine the multiple target positions by CCD camera. Based on the developed image processing method, the sample targets with different colors were verified using the arm robot system[8]. The original image is shown in figure 3, the extracted results is shown in figure 4.



a) The extraction of two purple marks



b) The extraction of green mark

Fig.4 Illustration of multiple targets extraction based on image processing

After the target position is obtained, the robot will move its arm to the target position. Figure 5 is an example to show the arm robot move to the above of one blue target.

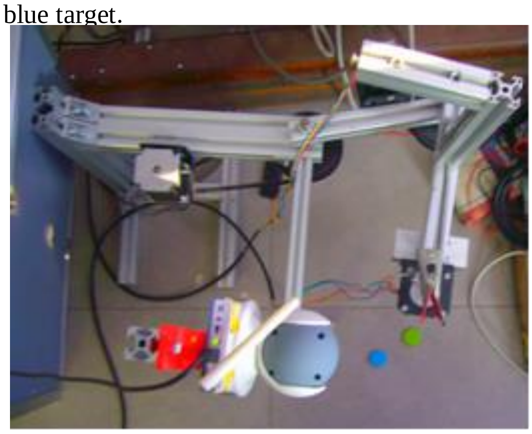


Fig.5. The illustration of the arm robot performing target moving operation

IV. CONCLUSIONS

In this paper, an arm robot and the position control system based on CCD camera is developed. By fusion of color based target detection and moment invariants, the target can be extracted efficiently and its position can be calculated. Then the robot can reach target position accurately.

Later, the further research, such as error analysis etc., will be carried on.

REFERENCES

[1] Masanori Sugisaka, Xin Wang, Intelligent control of a mobile vehicle using on-line learning, System science, Vol. 25, 1999: 41-50.

[2] Masanori Sugisaka, Xin Wang, Ju-Jung Lee, Intelligent control strategy for a mobile vehicle, Applied mathematics and computation, 1998: 91-98.

[3] S. O. Belkasim, M. Shridhar, M. Ahamadi, Pattern recognition with moment invariants: a comparative study and view results, Pattern recognition, Vol.24, No.12,1991: 1117-1138.

[4] Alireza Khotanzad, Yaw Huahong, Rotation invariant image recognition using features selected via a systematic method, Pattern recognition, Vol.23, No.10,1990: 1089-1101.

[5] Hagan, M.T., H.B.Demuth, and M.H.Beale, Neural network design, Boston, Ma: PWS publishing, 1996

[6] Moller, M.F., A scale conjugate gradient algorithm for fast supervised learning, Neural networks, Vol. 6, 1993: 525-533

[7] Riedmiller, M., and H.Braun, A direct adaptive method for faster backpropagation learning: The RPROP algorithm, Preceeding of the IEEE international conference on neural networks, 1993

[8] Vogl, T.P., J.K.Mangis, etc., Accelerating the convergence of the backpropogation method, Biological cybernetics, vol.59, 1988:256-264

Improved Map Generation by Addition of Gaussian Noise for Indoor SLAM using ROS

Khairul Salleh Mohamed Sahari and Barry Loh Tze Yuen

Centre for Advanced Mechatronics and Robotics, College of Engineering, Universiti Tenaga Nasional, Jalan IKRAM-UNITEN, 43000, Kajang, Selangor, Malaysia
Tel : 60-3-89212020 Fax : 60-3-89212116
khairuls@uniten.edu.my

Abstract: Rao-Blackwellized Particle Filter (RBPF) is used in this paper to solve the Simultaneous Localization and Mapping (SLAM) problem. RBPF algorithm uses particle filter where each particle carries an individual map of the environment. With the usage of Robot Operating System (ROS), GMapping package was used as a basis for map generation and SLAM. To improve the map generation, Gaussian noise was introduced to the data from laser range finder and also the odometry from the robot Pioneer P3AT's pose. The introduced algorithm was successful in decreasing the uncertainty as well as increased the knowledge of each particle in the estimation of the robot's pose, proven through practical experiment. Exploration experiments were also carried out to test the performance of P3AT based on our proposed method.

Keywords: SLAM, ROS, Gaussian Noise, map generation, exploration**.**

I. INTRODUCTION

Over the years, the rapid development of technology has contributed to the development of robots. The production of many new sensors including laser range finders and Kinect camera provide more precise, reliable and faster readings.

Robots equipped with sensors allow them to learn about their surrounding environment. This enables the robot to conduct various tasks that include map generation, localization and exploration, indoor services like moving from one room to another, gripping and picking up objects. Moreover, robots should be able to conduct Simultaneous Localization and Mapping (SLAM) to achieve even greater results.

SLAM, as the name suggests, allows the robot to conduct mapping of the environment while localizing itself in the map being created on-the-fly. An addition to SLAM algorithm is to utilize more than one sensor for more accurate results. Castellanos et al. [1] presented a 2D version of SPmodel, a probabilistic representation model and EKF integration system for the usage of multi-sensor fusion. Results obtained proved that fusing two sensor data improved the localization of mobile robots even more accurately. Besides that, Arras et al. [2] proved that using feature extraction and the Kalman filter to fuse data from a laser range finder and monocular camera increased the precision of localization significantly. On the other hand, utilization of open-sourced systems and software are becoming very common. Zaman et al. [3] presented their findings of utilizing open-sourced system, Robot Operating System (ROS) for mapping, localization and autonomous navigation. Gong et al. [4] also presented their research on utilization of ROS for the purpose of object localization, using RFID and laser scans. Despite being open-sourced and free, there are also drawbacks.

Hence this paper is produced to test the reliability and precision of already programmed SLAM algorithms found in open-sourced software ROS [5] using a common research robot. Based on the results obtained from experimentations, improvements were carried out to further enhance the SLAM and map generation. By studying the algorithms written by the open-sourced community, a portion of the code was found that the addition of a Gaussian type noise will implement the usage of Kalman filter to improve the knowledge of the particle filters of the map. Besides that, exploration experiments were carried out on the robot to further test the reliability and performance of the proposed improvement. To achieve this objective, a Hokuyo URG-04LX laser range finder was used together with a Pioneer 3-AT equipped with wheel encoders only. A common laptop was used to provide the algorithms and data collection throughout the process.

II. CURRENT WORK

In this research, a ROS-based control system was used for the Pioneer 3-AT robot for localization and creation of maps in unexplored controlled and uncontrolled indoor environments.

1. ROS setup for Pioneer 3-AT

The two sensors used in this research are the Hokuyo URG-04LX, and the wheel encoders of the mobile robot. Linux (Ubuntu 12.04 LTS) based computer system equipped with ROS was used to control the robot. For manned maneuvers of the robot, a keyboard was used. Both the laser and the mobile robot are connected to the computer through USB ports. Based on ROS, the packages and stacks used in this research are the P2OS, Hokuyo_node, SLAM_Gmapping, Navigation stack and the Exploration stack. The P2OS and Hokuyo_node package serve as the driver to establish connection and data transfer between the laser range finder and the mobile robot to the computer. Not only that, the P2OS package provide a 3D model of the Pioneer 3-AT that can be viewed in the RVIZ Visualization tool and also allows the movement of the mobile robot to be controlled by a joystick or a keyboard.

SLAM_Gmapping package provides a highly efficient Rao-Blackwellized particle filter (RBPF) to create occupancy grid maps. The RBPF algorithm not only incorporates data fusion between the laser and robot’s pose, but also solves the SLAM problem. However, the utilization of particles to carry individual map of the environment possess a big threat to the performance of the computer. This is because large computations have to be conducted. Grisetti et al. [6] proposed adaptive techniques to be included in the GMapping algorithm to reduce the number of particles in a RBPF for learning the occupancy grid maps. Gmapping also computes an accurate proposal distribution, considering not only the movement of the robot, but also the most recent observation. This produces drastic decrease of the uncertainty in the robot’s pose, in the prediction step of the filter. Re-sampling operations were also carried out selectively which reduces the problem of particle depletion.

Following up the algorithm of Gmapping, individual importance weightage is assigned to each particle based on the significance of the sampling principle. The weightage allows resampling that performs particle filtering when the target distribution differs from the proposal distribution. Instinctively, the more efficient proposal distribution approximation of the target distribution, the better is the performance of the filter. Montemarlo et al. [7] presented RBPF that uses a Gaussian approximation of the improved proposal in the context of landmark-based SLAM. However, Grisetti et al. [6] found that GMapping uses the same principle of Gaussian approximation of improved proposal but in the context of dense grid maps.

The Gaussian computes for each particle using a Kalman filter that estimates the pose of the robot. This technique is utilized whenever the map is represented by a set of features and the errors affecting the feature detection is assumed to be Gaussian. From the understanding, to fully utilize the Kalman filter, a Gaussian distribution or also known as normal distribution type of noise was added to both the laser range finder sensor data and the wheel encoders’ sensor data. The Box-Muller transform was used to add a Gaussian distribution noise to the sensor data. In [8], Box and Muller mentioned that the Box-Muller transform is a pseudo-random number sampling method to generate pairs of standard, independent, normally distributed random numbers when provided uniformly distributed random numbers. The Box-Muller transform can be expressed in two different forms, the basic and polar form. The polar form calculations were adapted in the source codes due to its advantages of being arguably faster and cheaper to the performance of the CPU.

The navigation stack provides a costmap that is built around the mobile robot as it moves to provide information of any nearby obstacles. The navigation stack also provides Dijkstra algorithm, Trajectory Rollout and Dynamic Window Approach for local and global path planning, and obstacle avoidance.

The exploration stack implements a frontier-based entropy approach with loop closing abilities for further stability. The exploration node scans the environment around the current robot’s position for reachable empty spaces within the map. Nearest goals will be selected while unreachable areas are ignored. The exploration node work very closely with the costmap created by the navigation stack.

Throughout the experiments, the configuration files for all the packages were tweaked to provide improvements and enhance the results obtained.

2. Experiments

Two experiments were carried out in different types of environments. The map area, as shown in Figure 1, is a self-created controlled environment. It is an empty room, filled with cardboard boxes, with a dimension of 7m x 5m. The purpose of the controlled environment is to ensure a safe and high performance working automatic robot system before proceeding with more complex situations. Figure 2 shows the map generated which only took about 90 seconds.

Figure 1: Self-created controlled environment

Figure 2: Map generated of the controlled environment

The other experiment was carried out in an uncontrolled environment, that is the South Wing corner of the College of Engineering, UNITEN with total estimated dimensions of about 41m x 25m. During this experiment, to prove that the map generation will be increased further, the mobile robot was manually driven twice. Once without the addition of Gaussian noise to the sensor data, and another time with the addition of Gaussian noise. Few parameters like the configurations, path taken by the mobile robot and the speed of the mobile robot were kept constant throughout the entire experiment to obtain results with integrity. Figures 3 and 4 show the map generated without addition of Gaussian noise and addition of Gaussian noise, respectively.

For the final experiment, the mobile robot was equipped to utilize the exploration stack in the same uncontrolled environment. During this experiment, the mobile robot was allowed to be independent in conducting path planning and obstacle avoidance. The entire map generation took about 35 minutes and shown in Figure 5. Throughout all the experiments, there were many issues faced and observations made that will be discussed further in the next section.

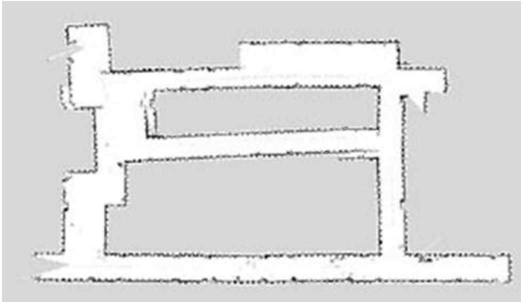


Figure 3: Map generated for uncontrolled environment without addition of Gaussian noise

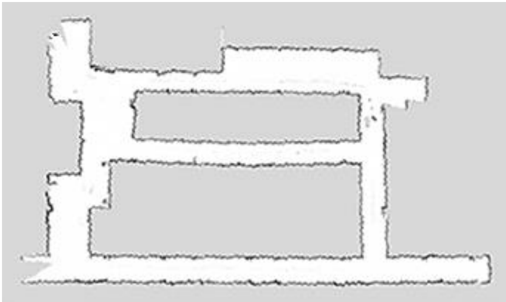


Figure 4: Map generated for uncontrolled environment with addition of Gaussian noise

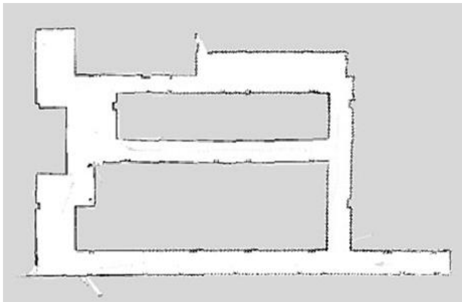


Figure 5: Map generated for uncontrolled environment with addition of Gaussian noise and exploration stack

III. DISCUSSION

1. Observations

In smaller environments, the mobile robot is able to navigate around smoothly and generate accurate maps. This is mainly due to the fact that in that short period of time, not many errors were accumulated by the wheel encoders. Not only that, the laser range finder is powerful enough for the small environments. However, in bigger environments like the uncontrolled environment, the map generated tends to have a lower resolution and is inconsistent. As the mobile robot continues to navigate and generate map of its surrounding environments, the wheel encoders tend to accumulate more and more errors, mainly due to slippage. Hence, to counter the accumulating errors, the sensor data from the wheel encoders were fused together with sensor data from the laser range finder. However, as shown in Figure 3, the effort still may not be enough as the map generated is not very satisfactory. With the addition of Gaussian noise to both the sensor data, the map generation improved quite significantly due to the usage of Kalman filter.

On the other hand, Figure 5 displays one of the best maps generated. The exploration stack from ROS implements a frontier-based exploration algorithm, which also includes loop closing methods. It formulates the entropy of the environment as the mobile robot is navigating. Higher entropy generated signifies a higher uncertainty of the knowledge of the environment. This allows the mobile robot to recalculate and navigate through the most optimal path with low entropy to generate the most accurate map.

2. Issues

Despite providing the algorithms and hardware for optimal map generation, there were still some factors that were affecting the results. Most of these factors involve the aspects related to both the environment and the mobile robot. These issues include:

A. Height and orientation of the laser range finder

The laser range finder was placed 20cm above the ground and onto the platform of the mobile robot. The laser range finder also scans the environment at a plane parallel to the ground. Due to that, there were certain occasions whereby the mobile robot could not navigate away from an obstacle as the obstacle’s height was below the scanning plane of the laser. Furthermore, the laser range finder has a blind spot at the back and would not be able to detect obstacles behind the robot.

B. Objects with wider lower parts

Certain objects in the environment like chairs and tables affect the navigation of the mobile robot. The mobile robot tends to have the perception that the path planned is clear. However, the chair in Figure 6 has its frame sticking out and the mobile robot has a very high chance of actually colliding with it.

C. Surface difference and inclination

As shown in Figure 7, the difference in surface type and its inclination caused a disturbance in the sensor data of the laser range finder, generating maps with structural irregularities. Besides the laser range finder, the wheel encoders also produce a much higher error count whenever the mobile robot is navigating through this area. Moreover, the main plug hole covers which are not properly fitted a shown in Figure 7 (right) also contributed to inconsistency in the map generated.

Figure 6: Chair with a wider lower part

Figure 7: Left picture shows the different surface types and inclination. Right picture shows the main plug hole covers

D. Overbearing dynamic obstacles

While conducting the exploration experiment, there may be times whereby the mobile robot may stop for a long period of time before being able to navigate itself. When there are too many people walking pass by in the room, the costmap generated by the navigation stack tends to get very congested and not being refreshed quickly. The mobile robot will then assume that there are many obstacles very close to it in all directions.

IV. CONCLUSION

Maps of the environment are very important to the mobile robot. With accurate maps, the mobile robot may be able to conduct proper path planning, obstacle avoidance and precise localizations.

The GMapping node from ROS utilizes the RBPF SLAM. Not only that, it also uses the EKF for multi-sensor fusion to reduce the amount of errors accumulated by the sensors. Based on the algorithm, it detects that if the sensor data contains error of Gaussian distribution, a Kalman Filter is used to compute each particle that estimates the robot's pose. Given the linearity and advantages of Kalman Filter, the estimation would be more accurate to correct the localization. In turn, it will generate more precise and consistent maps as the particles contain more valuable information than before. Hence, two nodes were written in C++ language to perform addition of uniform distribution values using the Box-Muller transform to both the laser range finder sensor data and the wheel encoders sensor data. Experiments were then carried out to prove the improvement of the map generated which turned out to be successful as shown in Figures 4 and 5.

However, this improvement is only limited to applications using RBPF-based grid-mapping algorithm and the usage of GMapping in ROS. Further implementations in the future may include a visual sensor for calculating depths to enhance further the knowledge of the particles and reduce the accumulated errors. Not only that, it is also recommended to look for improvements in the costmap generated, reliability of the sensor data, and even various mapping strategies that could be implemented.

ACKNOWLEDGMENT

The authors would like to thank Ministry of Education, Malaysia for the funding of this work.

REFERENCES

[1] J. Neira, J.Horn, J. Tardos and G. Schmidt, "Multisensor mobile robot localization", in Robotics and Automoation, 1996. Proceedings. 1996 International Conference on, vol. 1, pp. 673-679, 1996.

[2] K. Arras, N. Tomatis and R. Siegwart, "Multisensor on-the-fly localization using laser and vision," in Intelligent Robots and Systems, 2000. (IROS 2000). Proceedings, 2000 IEEE/RSJ International Conference on, vol. 1, pp. 462-467, 2000.

[3] S. Zaman, W. Slany and G. Steinbauer, "ROS-based mapping, localization and autonomous navigation using a Pioneer 3-DX robot and their relevant issues", Electronics, Communications and Photonics Conference (SIECPC), 2011 Saudi International, pp. 1-5, 2011.

[4] S. Gong, H. Liu, Y. Hu, J. Zhang, "ROS-based object localization using RFID and laser scan," Information and Automation (ICIA), 2012 International Conference on, PP. 406-411, 2012.

[5] M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler and A. Y. Ng, "Ros: an open-source robot operating system," in ICRA Workshop on Open Source Software, 2009.

[6] G. Grisetti, C. Stachniss and W. Burgard, "Improved techniques for grid mapping with rao-blackwellized particle filters," Robotics, IEEE Transactions on, vol. 23, no. 1, pp. 34-46, 2007.

[7] M. Monteerlo, S. Thrun, D. Roller and B. Wegbreit, "Fastslam 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges," in Proceedings of the 18th international joint conference on Artifical Intelligence, IJCAI'03, (San Francisco, CA, USA), pp 1151-1156, Morgan Kaufmann Publishers Inc., 2003.

[8] G. E. P. Box and M. E. Muller, "A note on the generation of random normal deviates," The Annals of Mathematical Statistics, vol. 29, pp. 610-611, 1958.

The recollection characteristics of generalized MCNN using different control methods

Shun Watanabe[1], Takashi Kuremoto[1], Kunikazu Kobayashi[2], Shingo Mabu[1], and Masanao Obayashi[1]

[1]Graduate School of Science and Engineering, Yamaguchi University, Yamaguchi, Japan
Tel: +81-836-85-9005; e-mail { s005we, wu, mabu, m.obayas }@yamaguchi-u.ac.jp
[2]School of Information Science and Technology, Aichi Prefectural University, Aichi, Japan
Tel: +81-561-64-1111; e-mail kobayashi@ist.aichi-pu.ac.jp

Abstract: As an auto-association memory model, a chaotic neural network (CNN) proposed by Adachi et al. is able to recollect multiple stored patterns dynamically. Kuremoto et al. proposed a multi-layer chaotic neural network (MCNN) combined multiple CNN layers to realize mutual association of plural time series patterns. However, recollection simulation of MCNN was limited in a two-layer model, the recollection characteristics c oncerning with the different external inputs, and the control methods were not investigated. In this paper, we extend the MCNN to be a general from (GMCNN) with more layers and use particle swarm optimization (PS O) to select internal parameters of CNNs to improve the recollection performance of GMCNN. The recollecti ng characteristics of different GMCNNs using different methods to decide internal parameters were invested b y computer simulations of multiple time series pattern association.

Keywords: chaotic neural network, association memory, time-series pattern, particle swarm optimization (PSO)

I. INTRODUCTION

Hopfield model has characteristics of a transfer of network energy to a stable state and an ability to recollect a stored pattern stably [1]. On other hand, a chaotic neural network (CNN) proposed by Aihara et al. is able to recollect multiple stored patterns dynamically because CNN consists of neurons which internal states perform unstably and chaotically [2, 3]. Generally, chaotic neural network models proposed in the 1990s are known as their recollection characteristics of "chaotic itinerancy", which means stored patterns are recollected one by one in aperiodic states [2-4]. CNNs are applied to not only dynamic association memory [3] but also optimization [5] and complex computing [6]. Additionally, Kuremoto et al. have combined multiple CNNs to be a multi-layer chaotic neural network (MCNN) [7-11]. MCNN is able to recollect mutually multiple time series patterns by controlling the unstable behaviors of CNN. Control methods for MCNN are proposed by various approaches, e.g., by observing the amount of dynamical state changes [9, 10], or upper and lower limit of recalling times [12,13], and meta-heuristics (e.g., genetic algorithm, particle swarm optimization (PSO)) [14]. MCNN has been applied to a mathematical model of hippocampus to realize a transform of long-term memory [8-11]. However, in these works, the number of layers in MCNN was limited to two layers. So we modified MCNN to be a generalized MCNN (GMCNN) and have investigated recollect characteristics of GMCNN using direct control method (DC) [15]. However, characteristics of GMCNN using other control methods was not investigated and compared.

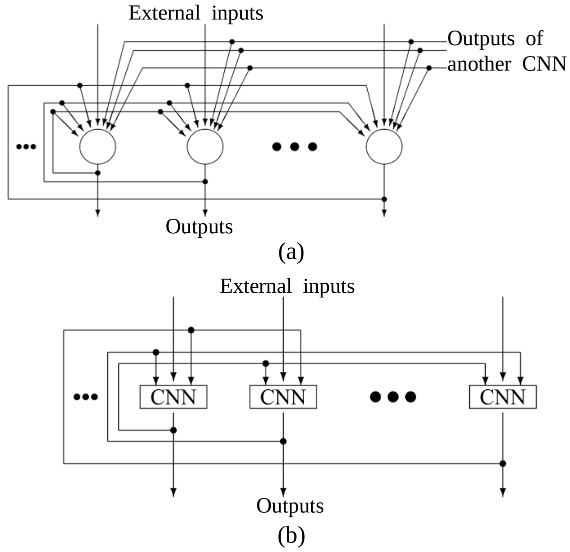


Fig. 1 Association memory models: (a) a chaotic neural network (CNN) [2, 3], (b) a generalized multi-layer chaotic neural network (GMCNN) [7-11].

In this paper, a meta-heuristics control method using PSO (PSOC) for GMCNN was proposed. Computational simulations of 2, 3, and 4 layers GMCNN were reported and the recollection characteristics of GMCNNs using DC [15] and PSOC were compared by the simulation results.

II. ASSOCIATION MEMORY MODEL

In this section, a generalized multi-layer chaotic neural network (GMCNN) which consists of plural chaotic nerual networks (CNNs) of Aihara et al. [7-11] and its storage/recollection methods of multiple time series patterns are introduced.

1. Generalized multi-layer chaotic neural network

GMCNN is structured by interconnecting chaotic neurons of CNN layers which have same network structures [15]. In Fig.1 (a), a CNN is shown and in Fig.1 (b) a GMCNN which has multiple CNN layers is shown. The chaotic neurons in CNN layer connect to each other as complete connected recurrent network as same as Hopfield network [1]. CNN layers in Fig.1 (b) connect to each other and play either role of "display layer" or "association layer". "display layer" means a state of the layer keeps static. "association layer" means the layer output patterns dynamically. GMCNN is able to recollect multiple time series patterns by switching the roles between CNN layers. The dynamics of the chaotic neuron of CNN layer in GMCNN is defined as follows:

$$\eta_i^{\langle n\rangle}(t+1) = k_1^{\langle n\rangle}\eta_i^{\langle n\rangle}(t) + \sum_j^N w_{ij}^{\langle n,n\rangle} x_j^{\langle n\rangle}(t)$$

$$\zeta_i^{\langle n\rangle}(t+1) = k_2^{\langle n\rangle}\zeta_i^{\langle n\rangle}(t) - k_3^{\langle n\rangle} x_i^{\langle n\rangle}(t) + a_i^{\langle n\rangle}$$

$$\xi_i^{\langle n\rangle}(t+1) = k_4^{\langle n\rangle}\xi_i^{\langle n\rangle}(t) + \sum_{m\neq n}^M \sum_j^N w_{ij}^{\langle n,m\rangle} x_j^{\langle m\rangle}(t) \quad (1)$$

$$y_i^{\langle n\rangle}(t+1) = \eta_i^{\langle n\rangle}(t+1) + \zeta_i^{\langle n\rangle}(t+1) + \gamma\xi_i^{\langle n\rangle}(t+1)$$

$$x_i^{\langle n\rangle}(t+1) = \begin{cases} f\left(y_i^{\langle n\rangle}(t+1)\right) & \text{if} \quad R^{\langle n\rangle}(t), \\ x_i^{\langle n\rangle}(t) & \text{otherwise} \end{cases} \quad (2)$$

where, t is an association time, $n,m \in \{1,\ldots,M\}$ is CNN layer number, $i,j \in \{1,\ldots,N\}$ is neuron number of CNN layer, M is the number of CNN layers in GMCNN, N is the number of the neurons of one CNN layer, $\eta_i^{\langle n\rangle}(t), \zeta_i^{\langle n\rangle}(t), \xi_i^{\langle n\rangle}(t)$ are internal states for feedback inputs from self-layer, refractoriness and inputs from another layers respectively, $w_{ij}^{\langle n,m\rangle}$ is a connection weight from jth neuron of mth layer to ith neuron of nth layer, $k_1^{\langle n\rangle}, k_2^{\langle n\rangle}, k_4^{\langle n\rangle}$ are decay parameters for feedack inputs from self-layer, refractoriness and inputs from another layers respectively, $k_3^{\langle n\rangle}$ is a refractory scaling parameter, $a_i^{\langle n\rangle}$ is a summation of threshold and external input, γ is a rate of effectiveness from another layers, $y_i^{\langle n\rangle}(t)$ is a internal state, $\mathbf{x}^{\langle n\rangle}(t) = \{x_1^{\langle n\rangle}(t),\ldots,x_N^{\langle n\rangle}(t)\}$ is an output of CNN, $R^{\langle n\rangle}(t)$ is a propositional function (nth CNN layer performs as "association layer" if $R^{\langle n\rangle}(t)$ is fulfilled, otherwise the layer performs as "display layer"). $f(\cdot)$ is a output function of the neuron:

$$f(y) = \frac{1}{1+e^{-y/\varepsilon}} \quad (3)$$

where, ε is a steepness parameter. $\mathbf{k}^{\langle n\rangle} = \{k_1^{\langle n\rangle},\ldots,k_4^{\langle n\rangle}\}$ are called internal parameters in this paper.

2. Storage process of time series patterns

GMCNN stores each partial pattern of plural time series patterns into the multiple CNN layers alternately. Hebbian learning rule [16] for storing the plural time series patterns are defined as follows:

$$g(t_l) = t_l \bmod M + 1$$

$$\tau^{\langle m\rangle}(t_l) = M\left\lfloor \frac{t_l - m + 1}{M} \right\rfloor + m - 1$$

$$\delta_{ij}\left(t_l, \tau^{\langle m\rangle}\right) = \beta(2\chi_i(t_l) - 1)\left(2\chi_i\left(\tau^{\langle m\rangle}\right) - 1\right) \quad (4)$$

$$\Delta w_{ij}^{\langle n,m\rangle}(t_l) = \begin{cases} 0 & \text{if} \quad n \neq g(t_l), \\ 0 & \text{if} \quad n = m \wedge i = j, \\ \delta_{ij}\left(t_l, \tau^{\langle m\rangle}\right) & \text{othrewise} \end{cases}$$

$$w_{ij}^{\langle n,m\rangle} \leftarrow w_{ij}^{\langle n,m\rangle} + \Delta w_{ij}^{\langle n,m\rangle}(t_l) \quad (5)$$

where, t_l is a storage time, $\chi(t_l) = \{\chi_1(t_l),\ldots,\chi_N(t_l)\}$ is a partial pattern of the time series patterns at t_l, $g(t_l)$ is the layer number to store a pattern, $\tau^{\langle m\rangle}(t_l)$ is a storage time for mth layer at t_l, $\lfloor\cdot\rfloor$ is floor function, β is a storage rate, $\Delta w_{ij}^{\langle n,m\rangle}(t_l)$ is a correction mass of the weight, Eq.(5) is an equation to renew state at t_l.

3. Network energy

GMCNN has a characteristic of transfer network energy (NE) to a local minimal state when the network outputs its stored pattern (also called "recollection") [1] because GMCNN store the patterns by Hebbian learning rule (which is also used in Hopfield network [1], CNN [2,3], and MCNN [8-11]). In this paper, the equation of NE is defined as a follow:

$$E(t) = -\frac{1}{2}\sum_{n,m}^M \sum_{i,j}^N w_{ij}^{\langle n,m\rangle} x_i^{\langle n\rangle}(t) x_j^{\langle m\rangle}(t) \quad (6)$$

III. CONTROL METHODS

The meta-heuristics control method using PSO (PSOC) is proposed for controlling the state of each CNN layer and switching the roles of “display layer” and “association layer” for GMCNN. The controlled the parameter of CNN layer are $\mathbf{k}^{\langle n\rangle}$. The propositional function $R^{\langle n\rangle}(t)$ of Eq.(5) is defined as $n=n'(t)$ in order to make only one CNN layer as "association layer" at all times. $n'(t)$ is the layer number as "association layer" at time *t* and is defined a follow:

$$n'(t+1)=\begin{cases} n'(t) \bmod M+1 & \text{if} \quad Q(t), \\ n'(t) & \text{otherwise} \end{cases} \tag{7}$$

where, $Q(t)$ is a propositional function for switching the role of CNN layer to be "display layer" or "association layer" at time *t*. The output of GMCNN for observing the output of "association layer" is defined as a follow:

$$z_i(t)=\begin{cases} 1 & \text{if} \quad x_i^{\langle n'(t)\rangle}(t)\geq 0.5, \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The direct control method (DC) [15] controls "association layer" in GMCNN by switching the internal parameters $\mathbf{k}^{\langle n'(t)\rangle}$ to different parameters from the amount of decrease of state changes of the layer. However, the switching parameters were set arbitrarily. In this paper, optimal parameters are found by a meta-heuristic method described in the next subsection and they are used to "association layer" to control the layer.

1. Particle swarm optimization control method (PSOC)

Particle swarm optimization (PSO) is a meta-heuristic method base on an exploratory behavior of a group of birds or fishes proposed by Kennedy and Eberhart [17, 18]. An exploration of optimal parameters by PSO is run according to a joint ownership parameter in a search space. Here, PSO is used to find the optimal parameters of each CNN layer in GMCNN. In [14], Kuremoto et al. used PSO to find the optimal parameters of MCNN, however, to evaluate fitness of each particle, it is used stored patterns during recalling process. To eliminate this supposition, the stability of NE given by Eq.(6) can be used to evaluate the fitness of each particle. An algorithm for the control method using PSO named "PSOC" is shown as follow:

[Algorithm for PSOC]

Step 1 Initializations of the particles are defined as follows:

$$\begin{aligned} & s \leftarrow 0 \\ & u_{pd}(s)\leftarrow U\left(u_d^{\langle \min\rangle},u_d^{\langle \max\rangle}\right) \\ & v_{pd}(s)\leftarrow U\left(v_d^{\langle \min\rangle},v_d^{\langle \max\rangle}\right) \end{aligned} \tag{9}$$

Step 2 $\mathbf{x}^{\langle n'(t)\rangle}(t+1)$ is calculated by Eq. (1)-(2) using $\mathbf{k}^{\langle n'(t)\rangle}$ given by $\mathbf{u}_p(s)$. The fitness of *p*th particle is defined as follows:

$$\begin{aligned} & \Delta E\left(t,\mathbf{u}_p(s)\right)=E(t)-E(t+1)\big|_{\mathbf{k}^{\langle n'(t)\rangle}\leftarrow \mathbf{u}_p(s)} \\ & F\left(\mathbf{u}_p(s)\right)=\max\left\{\Delta E\left(t,\mathbf{u}_p(s)\right),0\right\} \end{aligned} \tag{10}$$

Step 3 Equations for renewing the particles are defined as follows:

$$\begin{aligned} v_{pd}(s+1)\leftarrow & \ \omega u_{pd}(s) \\ & +c_1 r_1\left(\hat{u}_{pd}(s)-u_{pd}(s)\right) \\ & +c_2 r_2\left(\tilde{u}_d(s)-u_{pd}(s)\right) \\ u_{pd}(s+1)\leftarrow & \ u_{pd}(s)+v_{pd}(s+1) \end{aligned} \tag{11}$$

Step 4 The particles are reevaluated by Eq.(10).

Step 5 This process is repeated from Step 3 if an iteration step *s* fulfills $s<s^{\langle \max\rangle}$. Otherwise, this process ends and "association layer" is renewed by Eq. (1)-(2) ($t\leftarrow t+1$) after $\tilde{\mathbf{u}}(s)$ is applied to $\mathbf{k}^{\langle n'(t)\rangle}$ ($\mathbf{k}^{\langle n'(t)\rangle}\leftarrow \tilde{\mathbf{u}}(s)$). $Q(t)$ of Eq. (17) is defined as $E(t)\leq E(t+1)$.

Where, $U(a,b)$ is an uniform random number in a range $[a,b]$, $u_d^{\langle \min\rangle}$ and $u_d^{\langle \max\rangle}$ are the minimum value and the maximum value in a *d*-dimensional range $\left[u_d^{\langle \min\rangle},u_d^{\langle \max\rangle}\right]$, $v_{pd}(s)$ is a velocity of $d\in\{1,\ldots,D\}$ th parameter of $p\in\{1,\ldots,P\}$ th particle at step *s*, $E(t+1)\big|_{\mathbf{k}^{\langle n'(t)\rangle}\leftarrow \mathbf{u}_p}$ is the NE of Eq.(6) in "association layer" renewed by Eq.(1)-(2) in a case that parameters $\mathbf{u}_p(s)=\{u_{p1}(s),\ldots,u_{pD}(s)\}$ of *p*th particle at step *s* are applied to $\mathbf{k}^{\langle n'(t)\rangle}$ of the layer at time *t*, $F(\mathbf{u}_p)$ is the fitness of *p*th particle, ω is an inertia coefficient, c_1,c_2 are coefficients for local and global search, r_1,r_2 are uniform random numbers in a range $[0,1]$, $\hat{\mathbf{u}}_p(s)=\{\hat{u}_{p1}(s),\ldots,\hat{u}_{pD}(s)\}$ and $\tilde{\mathbf{u}}(s)=\{\tilde{u}_1(s),\ldots,\tilde{u}_D(s)\}$ are the best parameters in all past of *p*th particles and all particles and defined as follows:

$$\begin{aligned} & \hat{\mathbf{u}}_p(s)=\underset{\mathbf{u}_p(s')|0\leq s'\leq s}{\arg\max}\, F\left(\mathbf{u}_p(s')\right) \\ & \tilde{\mathbf{u}}(s)=\underset{\hat{\mathbf{u}}_p(s)|0\leq p\leq P}{\arg\max}\, F\left(\hat{\mathbf{u}}_p(s)\right) \end{aligned} \tag{12}$$

IV. COMPUTATIONAL SIMULATIONS

In this section, the storage / recollection simulations of GMCNN using DC [15] and PSOC and their results are reported. The neurons of all CNN layers are stimulated by the external input $a_i^{\langle n\rangle}$ in each simulation. And recollections of partial patterns of time series patterns are observed as the output of GMCNN. The external input is defined as a follow:

$$a_i^{\langle n\rangle}=c_e\left(2\hat{\chi}_i-1\right)$$

Table 1 Parameters setting.

Name	Symbol	Value
Number of neurons in a CNN	N	100
Initial internal states	$\eta_i^{\langle n \rangle}(0), \zeta_i^{\langle n \rangle}(0), \xi_i^{\langle n \rangle}(0)$	0
Initial connection weights	$w_{ij}^{\langle n,m \rangle}$	0
Initial patterns	$x_i^{\langle n \rangle}(0)$	Random pattern
Steepness parameter	ε	0.015
Dimension of parameters	D	4
Number of particle	P	10
Maximum step	$s^{\langle \max \rangle}$	20
Search space range	$\mathbf{u}^{\langle \max \rangle}$	{1,1,20,1}
	$\mathbf{u}^{\langle \min \rangle}$	{0,0,0,0}
Search velocity range	$\mathbf{v}^{\langle \max \rangle}$	{1,1,20,1}
	$\mathbf{v}^{\langle \min \rangle}$	{-1,-1,-20,-1}
Inertia coefficient	ω	1
Search coefficients	c_1, c_2	1
Initial association layer	$n'(0)$	1

Where, c_e is a scaling coefficient for external input, $\hat{\chi} = \{\hat{\chi}_1, \ldots, \hat{\chi}_N\}$ is an external input pattern given to each CNN layer.

Table 1 shows parameter setting in all simulations. Table 2 shows parameter setting in 3 kind of the simulation environments: Env.1 with 2 CNN layers, Env.2 with 3 CNN layers and Env.3 with 4 CNN layers.

In this paper, the number of kinds of the time series patterns stored in GMCNN is 2 defined as follows:

$$\boldsymbol{\chi}(t_l) = \begin{cases} \mathbf{r}(0,1) & \text{if} \quad 0 \le t_l \le 4 \\ \mathbf{r}(0,1) & \text{if} \quad 7 \le t_l \le 11 \\ \mathbf{0.5} & \text{otherwise} \end{cases} \tag{13}$$

$$\boldsymbol{\chi}(t_l) = \begin{cases} \mathbf{r}(0,1) & \text{if} \quad 0 \le t_l \le 4 \\ \mathbf{r}(0,1) & \text{if} \quad 7 \le t_l \le 11 \\ \mathbf{r}(0,1) & \text{if} \quad 14 \le t_l \le 18 \\ \mathbf{0.5} & \text{otherwise} \end{cases} \tag{14}$$

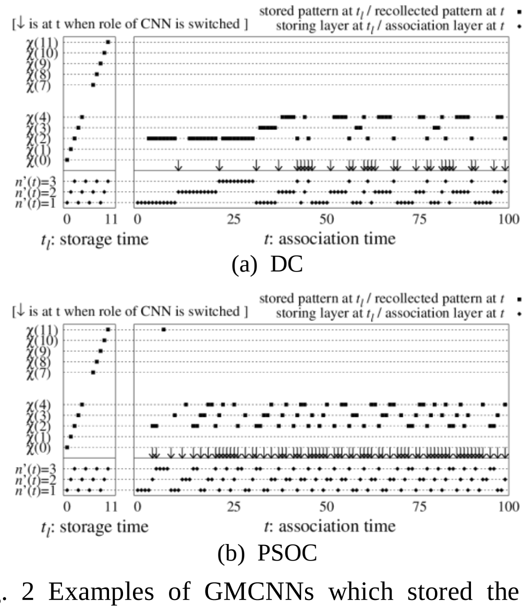
Where, $\mathbf{r}(0,1)$ is a binary pattern randomized with 0 or 1, $\boldsymbol{\chi}(t_l) = \mathbf{0.5}$ means no pattern can be observed (because $\delta_{ij}(t_l, \tau^{\langle m \rangle}) = 0$ in Eq. (4)). Eq. (13) gives 2 time series patterns ($0 \le t_l \le 4$, $7 \le t_l \le 11$) which are consisted of 5 random binary patterns. An interval between these time series patterns are 2 from storage time $t_l = 5$ to $t_l = 6$. On the other hand, Eq.(14) gives 3 time series patterns consisted of 5 random binary patterns which intervals between these time series patterns are 2 respectively ($t_l = 5,6$, $t_l = 12,13$).

In computational simulations, recollection processes of the stored plural time series patterns were observed from $t = 1$ to $t = 100$ in Env.1-Env.3 and in GMCNNs using DC [15] and PSOC. The recollection processes were repeated 100 times (as 100 samples) to evaluate a performance of GMCNN by each average stored pattern retrieval times in Env.1-Env.3.

Table 2 Parameter setting of each simulation environment.

Environment number	The number of layer	Rate of effectiveness	Learning rate	Scaling coefficient
	M	γ	β	c_e
Env.1	2	1	0.2	6
Env.2	3	0.5	0.3	9
Env.3	4	1/3	0.4	12

Fig. 2 shows storage processes and recollection processes of GMCNN with 3-CNN layers (Env.2). The plural time series patterns stored in GMCNN is the patterns of Eq.(13). Fig. 2 (a) is using DC. Fig. 2 (b) is using PSOC. The external input to GMCNN was a random pattern as same as the stored pattern at the storage time $t_l = 2$ given by Eq.(13). The left column of Fig.2 (a)-(b) shows the storage processes of GMCNN for storing 2 time series pattern $\boldsymbol{\chi}(0), \ldots, \boldsymbol{\chi}(11)$ of Eq.(13) by Eq.(4)-(5). The right column of Fig.2 (a)-(b) shows the recollections of the stored patterns ($\boldsymbol{\chi}(0), \ldots, \boldsymbol{\chi}(4)$ and $\boldsymbol{\chi}(7), \ldots, \boldsymbol{\chi}(11)$) on an above ten lines and the role of "association layer" on an under three lines in GMCNN with 3-CNN layers from the association time $t = 1$ to $t = 100$. From Fig.2, we can confirm that the switching of roles of CNNs worked constantly and $\boldsymbol{\chi}(2)$, $\boldsymbol{\chi}(3)$ and $\boldsymbol{\chi}(4)$ were recollected in different association times. And PSOC resulted more role switching times of 3-CNN layers than DC, comparing the role switching times of DC and PSOC (↓



(a) DC

(b) PSOC

Fig. 2 Examples of GMCNNs which stored the times series patterns of Eq. (13) in Env.2 and given $\boldsymbol{\chi}(2)$ of Eq. (13) as external input.

Table 3 The average retrieval frequencies of the part of the stored time series patterns, the average of their totals, the average total switching frequencies and the expectancies of the number of the patterns recollected for an association time from t=1 to 100 (and their standard deviations) in GMCNN which store Eq. (13) in Env.2 (in 100 sample).

(a) DC

Stored pattern	External input pattern: $\hat{\chi}$								
	$\chi(0)$	$\chi(1)$	$\chi(2)$	$\chi(3)$	$\chi(7)$	$\chi(8)$	$\chi(9)$	$\chi(10)$	random
$\chi(0)$	**18.89** (2.15)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.05 (0.50)	0.00 (0.00)
$\chi(1)$	**6.22** (4.28)	**20.31** (5.81)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.01 (0.10)
$\chi(2)$	**31.56** (6.19)	**19.66** (8.51)	**23.50** (8.61)	0.26 (2.59)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.61 (1.77)
$\chi(3)$	0.00 (0.00)	**20.63** (5.09)	**16.75** (9.95)	**28.75** (8.71)	0.00 (0.00)	0.09 (0.63)	0.00 (0.00)	0.38 (2.35)	1.29 (2.58)
$\chi(4)$	0.13 (0.93)	0.89 (1.56)	**17.82** (7.01)	**22.62** (13.88)	0.00 (0.00)	0.10 (0.70)	0.00 (0.00)	0.09 (0.63)	1.68 (3.51)
$\chi(7)$	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	**21.92** (5.08)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
$\chi(8)$	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	**5.88** (4.39)	**22.54** (7.39)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
$\chi(9)$	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	**24.73** (6.05)	**12.96** (7.41)	**17.43** (1.80)	0.00 (0.00)	0.50 (2.63)
$\chi(10)$	0.00 (0.00)	0.00 (0.00)	0.09 (0.90)	0.06 (0.42)	0.02 (0.20)	**13.87** (5.49)	**24.27** (11.57)	**28.70** (8.34)	0.39 (1.60)
$\chi(11)$	0.00 (0.00)	0.00 (0.00)	0.85 (3.93)	0.43 (2.54)	0.26 (1.02)	1.26 (2.07)	**29.30** (3.56)	**22.97** (14.46)	2.70 (5.20)
Total retrieval	56.80 (8.46)	61.49 (12.31)	59.01 (17.39)	52.12 (20.34)	52.81 (11.30)	50.82 (14.38)	71.00 (11.81)	52.19 (19.70)	7.18 (8.71)
Total switching	34.07 (3.51)	26.15 (3.29)	28.06 (3.46)	32.98 (4.43)	32.31 (3.22)	26.30 (3.21)	29.30 (3.22)	32.93 (4.40)	23.11 (3.67)
expectancy	**3**	**3**	**3**	**2**	**3**	**3**	**3**	**2**	**0**

(b) PSOC

Stored pattern	External input pattern: $\hat{\chi}$								
	$\chi(0)$	$\chi(1)$	$\chi(2)$	$\chi(3)$	$\chi(7)$	$\chi(8)$	$\chi(9)$	$\chi(10)$	random
$\chi(0)$	**13.82** (6.70)	0.25 (0.90)	0.22 (0.70)	0.00 (0.00)	0.22 (0.61)	0.14 (0.79)	0.00 (0.00)	0.38 (1.14)	0.45 (1.27)
$\chi(1)$	**13.14** (7.60)	**14.56** (6.97)	0.33 (1.00)	0.00 (0.00)	0.15 (0.74)	0.29 (0.98)	0.00 (0.00)	0.33 (1.11)	0.65 (1.67)
$\chi(2)$	**21.72** (2.31)	**17.29** (7.71)	**17.51** (9.79)	**13.19** (10.83)	1.56 (4.82)	1.38 (4.39)	0.12 (1.19)	3.73 (7.31)	5.18 (8.55)
$\chi(3)$	**9.97** (6.53)	**17.40** (7.54)	**15.74** (10.00)	**22.94** (6.20)	1.89 (4.97)	1.90 (4.78)	0.15 (1.49)	4.69 (7.87)	5.57 (8.39)
$\chi(4)$	**7.49** (6.29)	**6.71** (5.60)	**15.68** (9.82)	**20.39** (7.11)	1.70 (4.61)	1.69 (4.45)	0.12 (1.19)	4.00 (7.41)	5.20 (8.28)
$\chi(7)$	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	**4.93** (3.19)	0.04 (0.31)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
$\chi(8)$	0.00 (0.00)	0.00 (0.00)	0.02 (0.20)	0.20 (1.32)	**3.68** (3.18)	**8.61** (6.40)	0.00 (0.00)	0.00 (0.00)	0.11 (0.82)
$\chi(9)$	0.00 (0.00)	1.08 (3.67)	3.27 (6.96)	0.52 (2.55)	**16.59** (7.18)	**14.39** (8.26)	**26.88** (3.13)	8.63 (10.03)	4.29 (7.98)
$\chi(10)$	0.01 (0.10)	1.22 (3.99)	3.24 (6.95)	0.63 (2.84)	**13.55** (6.65)	**14.63** (8.34)	**25.88** (3.24)	**14.14** (10.03)	4.38 (8.12)
$\chi(11)$	0.00 (0.00)	1.24 (4.02)	3.42 (7.12)	1.46 (3.98)	**12.79** (6.45)	**8.18** (7.05)	**26.10** (3.26)	**12.53** (10.18)	4.67 (8.42)
Total retrieval	66.15 (10.98)	59.75 (21.49)	59.43 (24.16)	59.33 (20.26)	57.06 (17.61)	51.25 (23.35)	79.25 (6.27)	48.43 (26.10)	30.50 (29.43)
Total switching	62.87 (5.39)	62.66 (5.48)	65.94 (6.09)	64.43 (6.27)	59.13 (4.93)	60.29 (5.24)	73.83 (4.97)	61.37 (6.19)	58.12 (5.73)
expectancy	**5**	**4**	**3**	**3**	**5**	**4**	**3**	**2**	**0**

in Fig.2).

To investigate the relationship between the external input and the output of GMCNN (retrieval patterns), on columns in Table 3, we show the average retrieval frequencies of 10 stored patterns ($\chi(0),\ldots,\chi(4)$ and $\chi(7),\ldots,\chi(11)$) in GMCNNs with Env.2 given respective 9 kinds of the external input patterns ($\chi(0),\ldots,\chi(3)$, $\chi(7),\ldots,\chi(10)$ and a non-stored random pattern). In Table 3, above data in cells mean the average retrieval times, and under data in parentheses are deviations. **Bold number** in Table 3 means the average retrieval frequency is larger than its deviation (that is the recollection of the stored pattern is expected and it is called "expectancy" in this paper). Height of the "expectancy" indicates higher recollection performance of GMCNN. GMCNN using PSOC showed higher expectancy, comparing GMCNNs using DC in Table 3(a) and PSOC in Table 3(b).

Total retrieval frequencies of all stored patterns, total switching frequencies and expectancies for each external input pattern in Table 3 and other simulations are illustrated in Fig. 3. Fig. 3 (a) is in a case of 2 time series patterns given by Eq. (13) in 2-layer GMCNN (Env.1). Fig. 3 (b) and (c) is in 3-layer GMCNN and in 4-layer GMCNN. Fig. 3 (d)-(f) is in a case of 3 time series patterns given by Eq. (14) in 2, 3 and 4-layer GMCNN. The better performance of GMCNN using PSOC than DC (marked with ★) is concluded, comparing the total retrieval frequencies, total switching frequencies and expectancies in respective external input patterns. Therefore, PSOC proposed in this paper is suggested preferentially as the control method of the switching the role of CNN layers as "display layer" and "association" in GMCNN.

V. CONCLUSION

In this paper, we proposed the control method using PSO as meta-heuristics which finds and uses the optimal parameters in GMCNN. In the computational simulations using 3 kinds of parameter setting, 2 kinds of time series patterns, different external input patterns and 2 kinds of the control methods (DC and PSOC), the better performance of GMCNN using PSOC was confirmed by comparing the processing results of DC and PSOC. The highest expectancy is in 3-layer GMCNN. As future works, GMCNN is expected to be adopted into long term memory models and the limbic model of the brain.

REFERENCES

[1] Hopfeld JJ (1984) Neurons with graded response have collective computational properties like those of two-state neurons. Proc. Natl. Acad. Sci. U.S.A. 81:3088-3092

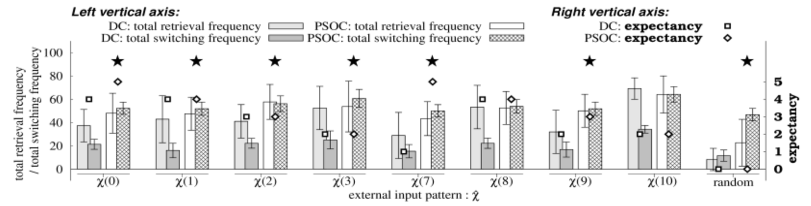
[2] Aihara K, Tanaka T, Toyoda M (1990) Chaotic neural network. Phys. Lett. A 144(6-7):333-340

[3] Adachi M, Aihara K (1997) Associative dynamics in a chaotic neural network. Neural Networks 10(1):83-98

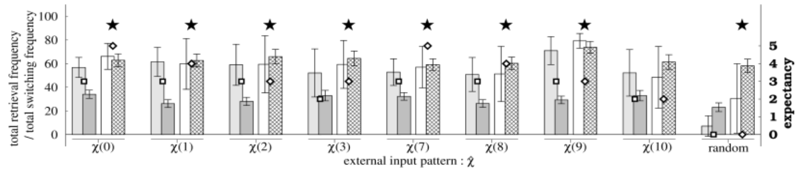
[4] Kaneko K (1990) Clustering coding, switching, hierarchical ordering, and control in a network of chaotic elements. Physica D 41:137-172

[5] Hasegawa M, Ikeguchi T, Aihara K (2002) Solving large scale traveling salesman problems by chaotic neurodynamics. Neural Networks 15(2):271-283

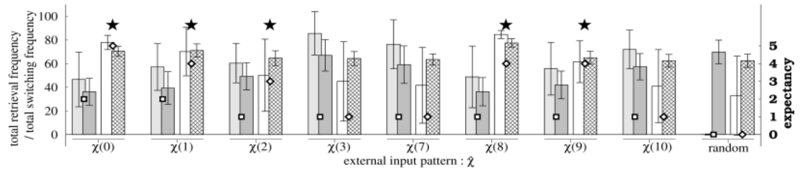
[6] Horio Y, Ando H, AiharaK (2009) Fundamental technology for complex computational systems. IEICE Fundamentals Review 3(2):34-44



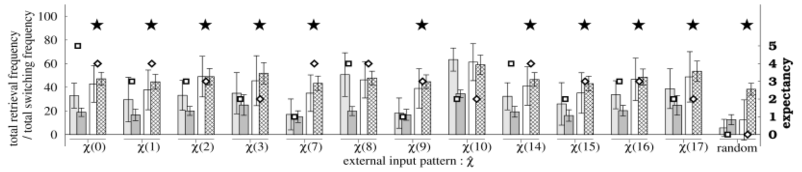
(a) A case of 2 time series patterns in Eq.(13) for a 2-layer GMCNN (Env.1)



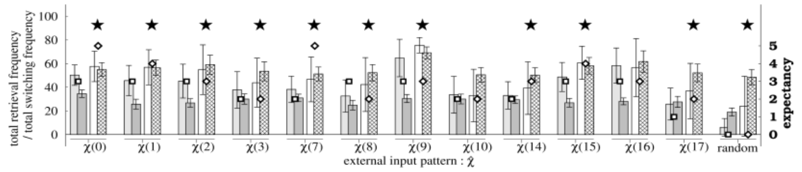
(b) The case of 2 time series patterns in Eq.(13) for a 3-layer GMCNN (Env.2) in Table 3



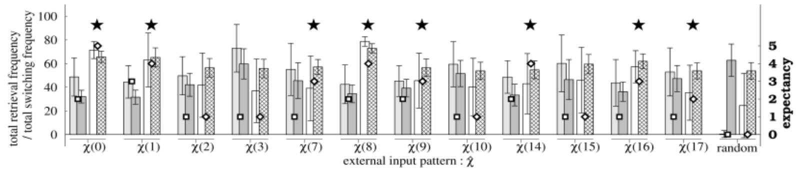
(c) A case of 2 time series patterns in Eq.(13) for a 4-layer GMCNN (Env.3)



(d) A case of 3 time series patterns in Eq.(14) for a 2-layer GMCNN (Env.1)



(e) A case of 3 time series patterns in Eq.(14) for a 3-layer GMCNN (Env.2)



(f) A case of 3 time series patterns in Eq.(14) for a 4-layer GMCNN (Env.3)

Fig. 3 The average total retrieval frequencies, the average total switching frequencies and the expectancies of GMCNNs (for an association time from t=1 to 100 in 100 samples) given by different parameter control methods: DC and PSOC; "★" denotes that over one of total retrieval, total switching and "expectancy" in the GMCNN using PSOC are higher than in one using DC.

[7] Kuremoto T, Eto T, Obayashi M et al. (2005) A multilayer chaotic neural network for associative memory. Proc. of SICE Annual Conf.

[8] Kuremoto T, Eto T, Kobayashi K et al. (2005) A chaotic model of hippocampus-neocortex. Lect. Notes Comput. Sci. 3610:439-448

[9] Kuremoto T, Eto T, Kobayashi K et al. (2007) A hippocampus-neocotex model for chaotic association. in Trends in Neural Computation (Eds: K. Chen and L. Wang), Studies in Computational Intelligence 35:111-133

[10] Kuremoto T, Ohata T, Kobayashi K et al. (2009) A dynamic associative memory system by adopting an amygdala model. AROB 13(2):478-482

[11] Kuremoto T, Ohata T, Kobayashi K et al. (2009) A functional model of limbic system of brain. Lect. Notes Comput. Sci. 5819:135-146

[12] He GG, Cao ZT, Chen HP et al. (2003) Controlling chaos in a neural network based on the phase space constraint. Int. J. Mod Phys B 7(22-24):4209-4214

[13] He GG, Shrimali MD, Aihara K (2008) Threshold control of chaotic neural network. Neural Networks 21:114-121

[14] Kuremoto K, Watanabe S, Kobayashi K et al. (2011) Dynamical recollection of interconnected neural networks using metaheuristics (in Japanese). IEEJ Trans. EIS 131(8):1475-1484

[15] Watanabe S, Kuremoto T, Kobayashi K et al. (2013) The recollection characteristics of a generalized MCNN. Proc. of SICE Annual Conf. 1375-1380

[16] Hebb DO (1949) The organization of behavior. New York: Wiley

[17] Kennedy J, Eberhart RC, Shi Y (1995) Particle swarm optimization. Proc. of the IEEE Int. Conf. on Neural Networks :1942-1948

[18] Kennedy J, Eberhart RC (2001) Swarm intelligence. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA

Revealing Terrorism Contents form Web Page Using Frequency Weighting Techniques

Thabit Sabbah[1], Ali Selamat[2] and Md Hafiz Selamat[3]

[1,2,3]K-Economy Research Alliance & Faculty of Computing,
Universiti Teknologi Malaysia, Malaysia

[1]sosthabit2@Live.utm.my
[2]aselamat@utm.my, [3]mhafiz@utm.my

Abstract: Intelligence and security informatics plays an important role in revealing terrorism in web content. However, the extremist groups are exploiting Internet facilities intensively to stimulate violence, hatred, and terrorism. Accurate classification of web content will increase the opportunities for the effective use of intelligence and security informatics in the early detection of terrorist activities. In this paper, we propose a feature selection method based on term weighting techniques TF, DF, and TF-IDF for revealing terrorism in web content. Firstly, we compare the performance of the individual term weighting techniques in revealing terrorism. Then, we present a method for enhancing the accuracy of classification. The method was tested with selected dataset from Dark Web Portal Forum. The experimental results show that classification using the proposed method is more accurate than the individual technique based classification.

Keywords: Classification, Term Weighting, Terrorism Detection, TF, IDF, TF-IDF.

1 INTRODUCTION

The Internet facilities such as anonymity, cheap development and maintenance environment and the huge potential audience are misused by the extremist groups to incite terrorism, market their ideas, and seek for the support and fund rising[1]. The hidden and covert part of the web used by terrorists or extremist groups online is referred as the Dark Web (DW)[2]. September 11 2001 (9/11) attacks stimulated the information technology researchers' to study and analyze different types of DW content such as text, video, images, etc., in order to reveal the potential terrorism. However, text content is the largest in web data [3], and in DW collection [4], so that many text content analysis based techniques are proposed to detect the terrorism on web, section 2 discusses some of these techniques.

The rest of this paper is organized as follows: section 2 reviews the related work in terrorism detection and term weighting techniques used in this study. In section 3 we compare the performance of individual term weighting techniques in terrorism detection. Then, we present our proposed feature selection method in section 4. While section 5 discusses the experimental results, and finally, we conclude this study in section 5.

2 RELATED WORKS

2.1 Terrorism Detection

Over the recent years, text analysis has been used intensively in detecting potential terrorism on the web based on statistical methods [3]. Commonly, in statistical methods, the text is represented as vectors of weighted features (words). TF-IDF term weighting algorithm is the most commonly weighting scheme used to determine the significant words in the text [3]. In some other cases, text is represented statistically base on lexical, syntactic, domain specific bag of words, and n-grams [5-10] features. However, the inability of statistical methods to understand the semantic meanings of text written by human [3] cause the low performance of these methods in text classification. In general, the performance of TF-IDF and other statistical methods is not sufficient [3, 11, 12]. So that some other techniques are proposed to overcome the deficiency in using individual statistical methods. Techniques proposed by [3, 13, 14] utilize the use of knowledge base tools that provide the conceptual hierarchal interconnections such as WordNet and Wikipedia to measure the semantic relations between features (concepts) in order to achieve higher performance.

In this research we compare the performance of individual term weighting techniques such as Term Frequency (TF), Document Frequency (DF), and Term Frequency Inverse Document Frequency TF-IDF as a feature selection methods, and propose a method for enhancing the accuracy of dark web detection.

2.2 Term Weighting Techniques

Term weighting techniques such as Term Frequency (TF) is used widely [11] in Information Retrieval and web classification domain, besides the DF and TF-IDF. However, there are many other term weighting algorithms proposed and used in the domain [15, 16]. Here is a brief description of the techniques used in this study.

i) Term Frequency (TF)

TF is concerned with the normalized frequency of certain term in the document. Larger TF value indicates the more significance of the term. *tf* of term *t* in document *d* is denoted as $tf_{t,d}$, and it is defined as in equation (1) [17]:

$$tf_{t,d} = \frac{N_{t,d}}{\sqrt{\sum_{t=1}^{n} \left(fr_{t,d} \right)^2}} \tag{1}$$

where the denominator is known as the Euclidean norm of the documents *d* in which $fr_{t,d}$ is the raw frequency of t^{th} term in document *d*, and *n* is the number of distinctive terms in document *d*.

ii) Document Frequency (DF)

DF weight of a term *t* (denoted by DF_t) represents the number of documents within the collection in which the term *t* is found. DF is expressed by equations (2):

$$DF_t = \sum_{d=1}^{D} df_{t,d} \tag{2}$$

where *D* is the total number of documents in the collection, and $df_{t,d}$ is a function defined as in equation (3):

$$df_{t,d} = \begin{cases} 1 & t \in d \\ 0 & t \notin d \end{cases} \tag{3}$$

iii) Term Frequency- Inverse Document Frequency (TF-IDF)

TF-IDF reflects the assumption that less frequent term in the collection is the most term significant, and vise-versa. TF-IDF is used by many information retrieval applications [18]. TF-IDF of a term *t* in document *d* (denoted by $TF\text{-}IDF_{t,d}$) is the dot product of the term frequency (TF) and the inverse document frequency (IDF) of the term, as in equation (4):

$$TF-IDF = TF_{t,d} . IDF_t \tag{4}$$

Where IDF_t is the inverse document frequency, which defined as in equation (5):

$$IDF_t = \log \frac{N}{DF_t} \tag{5}$$

3 COMPARISON OF INDIVIDUAL TERM WEIGHTING PERFORMANCE

To compare the performance of individual term weighting techniques in terrorism detection, we firstly downloaded thousands of Arabic web pages from the Dark Web Portal Forum (DWPF). DWPF is the largest collection of crawled terrorist related documents [4]. The data on DWPF is collected from 17 Arabic forums, and many other forums in other languages such as English, German, French, and Russian languages[4]. However, in this research we focus on Arabic web pages. An expert labeled the documents into two classes (*Dark* and *non-Dark*), a collection of 1000 documents (500 *from each class*) was selected as the dataset for our experiments. Then, the documents preprocessed and represented in the Vector Space Model (VSM) form, based on term weighting techniques included in this study. The preprocessing included removing the stop words, stemming, and filtering the text against non-Arabic characters, numbers, and symbols. Finally, the classification is performed using Support Vector Machine (SVM) classifier based on Top K feature sets for each term weighting technique. In this experiment, we use feature sets consist of the top 100, 200, 300, and 400 features. Following are the results of this experiment.

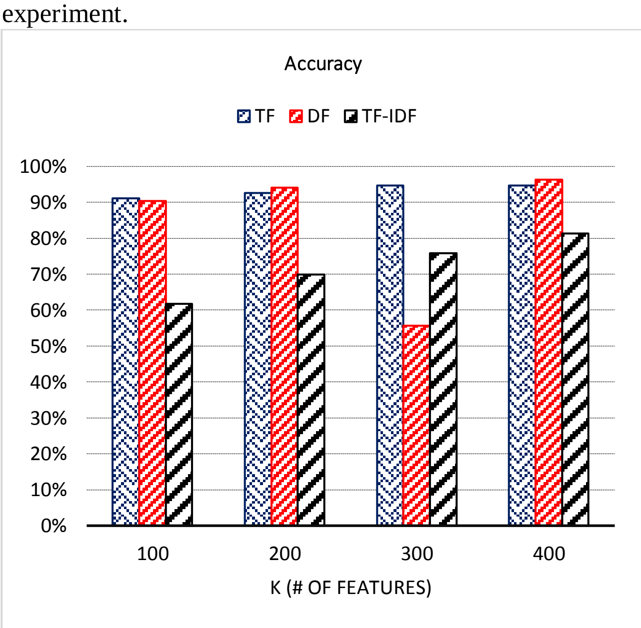


Fig. 1. Classification accuracy for individual term weighting techniques

As we can see from Fig. 1, the accuracy of classification based on TF-IDF is the least for all feature sets with maximum value near to 81% for top 400 features. However, the classification accuracies based on TF and DF are comparable for almost all feature sets except for Top 300 feature set. The highest accuracy was achieved by using the feature set of top 400 feature based on DF term weighting technique.

4 PROPOSED METHOD

4.1 Method Design

Our proposed method aims to enhance the accuracy of terrorism detection by combining the feature sets of individual term weighting techniques. In this study, we use three different term weighting techniques: TF, DF, and TF-IDF, so the combinations will consists of two sets and three sets combinations. The combination of two and three sets is defined as in equations (6-8):

First, let $$S = \left\{ \begin{array}{l} \{TpTF\}, \\ \{TpDF\}, \\ \{TpTF - IDF\} \end{array} \right\} \quad (6)$$

And

$A, B, C \in S$, where $A \neq B \neq C$, then

$$C(A,B) = A \cup B = \{x : x \in AorB\} \quad (7)$$

and

$$\begin{aligned} C(A,B,C) &= A \cup B \cup C \\ &= \{x : x \in AorBorC\} \end{aligned} \quad (8)$$

Where

$\{TpTF\}$ is the Top K features based on TF,
$\{TpDF\}$ is the Top K features based on DF,
$\{TpTF\text{-}IDF\}$ is the Top K features based on TF-IDF,
C(A,B) is the combination of two sets,
And
C(A,B,C) is the combination of three sets.

Based on the previous definitions, the following combinations are generated:

TF_DF, TF_TF-IDF, DF_TF-IDF, and TF_DF_TF-IDF.

4.2 Experiments and Results

To test and evaluate our method we conducted the experiment based on the same dataset and parameters of the individual term weighting techniques comparison, considering top 100, 200, 300, and 400 features.

エラー! 参照元が見つかりません。 shows the number of features and the classification accuracy for all possible combinations of two feature sets for all parameters.

However, there is only one possible combination based on three sets which is (TF_DF_TF-IDF), which generates only three sets according to different K values (i.e. number of features in the base feature sets which equal to 200, 300, and 400), as shown in Table 2.

Table 1 Number of features and accuracy achieved by combinations of two sets

		DF				TF-IDF			
# of features In base set		100	200	300	400	100	200	300	400
TF	100	90.71 **132**				92.18 **148**			
	200		94.23 **360**				92.28 **318**		
	300			95.80 **535**				93.16 **466**	
	400				96.28 **704**				93.45 **539**
DF	100					91.69 **177**			
	200						93.35 **401**		
	300							95.21 **600**	
	400								95.50 **799**

In エラー! 参照元が見つかりません。, the accuracy (in bold lines) varies from one combination to another, as well the number of features in the combination set. The highest accuracy achieved is 96.28% by combining the set of top 400 features generated by TF techniques with the corresponding set generated by DF technique, and the number of features in the combined set is 704 features. Our experiment included only the combination of feature sets that has the same number of features as seen in エラー! 参照元が見つかりません。.

Table 2 Number of features and accuracy achieved by combinations of three sets

Combination	TF_DF_TFIDF			
# of features in base set	100	200	300	400
Accuracy	91.89	93.16	95.60	96.48
# of Features	180	478	526	897

In Table 2, it seen that the highest accuracy achieved is 96.48% with 897 features when combining the feature sets generated based on top 400 features for each term weighting technique.

Last step was joining the results in Fig. 1 with the results in in エラー! 参照元が見つかりません。 and Table 2 (i.e. the accuracies and number of features based on individual term weighting technique with the results based on combining two and three sets respectively). From last step, we get Fig. 2 that compares the results of our proposed method to the results of classical method that depend on individual term weighting technique in feature selection.

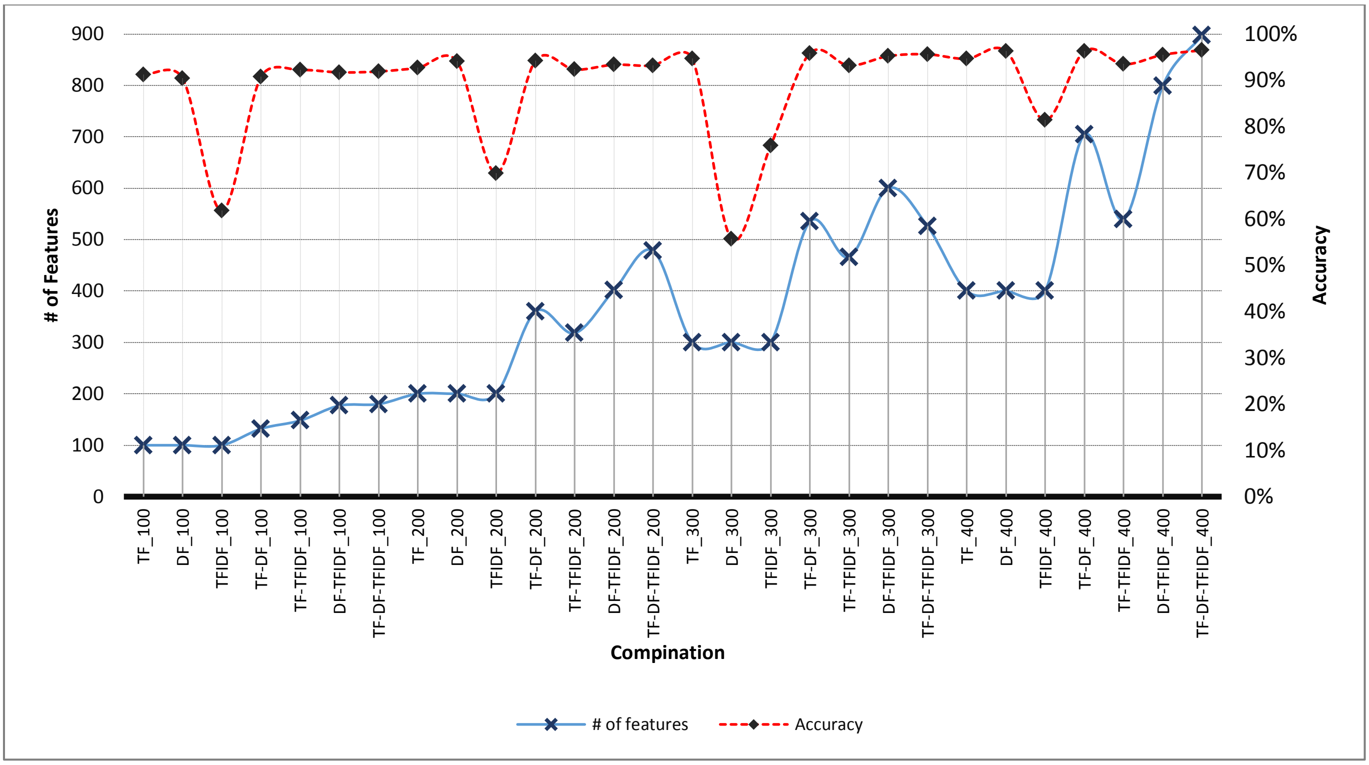


Fig. 2. Performance of proposed method versus classical method in terms of accuracy and number of features

From Fig. 2, It can be seen that the largest number of features among different combinations was around 900 features (in the TF-DF-TFIDF combination), and the accuracy at that point was the maximum (96.48%). Moreover, it can be seen, on one hand, that the lowest accuracy was recorded in the case of using the feature set of **Top 300** features generated by the DF technique individually. On the other hand, it also noticed that the accuracy (96.29%) achieved based on **Top 400** features generated by DF technique was almost comparable to the maximum accuracy achieved by the combination-based method. In addition, it is notable that the lowest accuracies are related to feature sets based on TF-IDF technique, in addition to the minimum accuracy mentioned above.

4 DISCUSSION

This research proposes a feature selection method based on the combination of feature sets generated by different classical term weighting techniques. By comparing the classification accuracy achieved in this research to the accuracies reported in [3, 12], we find that the highest achieved accuracy by our proposed method was greater. The accuracies reported in mentioned studies were 94% and 91.67% correspondingly, while the maximum accuracy achieved in this research is 96.48%. The results shown by this research are very encouraging because the higher accuracy could be achieved in relative to the fewer number of features. Moreover, the results clearly show that high accuracies (>95%) could be achieved by combining less small feature sets that are generated using different basic term weighting techniques.

5 CONCLUSION

In this paper, we compared the performance of the individual term weighting techniques TF, DF, and TF-IDF in revealing terrorism content. Then, we present a method for enhancing the accuracy of classification by combining the Top K feature sets generated individual techniques. The method was tested and evaluated on a selected dataset downloaded from the Dark Web Portal Forum. The results show that higher classification accuracy is achievable by combining different individual feature sets. As mentioned before, in this research, we only combine feature sets of the same sizes, so, our future work will focus on finding the point of equilibrium in which the highest accuracy could be achieved with less number of features.

AKNOWLEDGEMENTS

The Universiti Teknologi Malaysia (UTM) under research grant 03H02 and Ministry of Science, Technology & Innovations Malaysia, under research grant 4S062 are hereby acknowledged for some of the facilities utilized during the course of this research work.

REFERENCES

[1] Abbasi, A. and H. Chen. Affect intensity analysis of dark web forums. in ISI 2007: 2007 IEEE Intelligence and Security Informatics, May 23, 2007 - May 24, 2007. 2007. New Brunswick, NJ, United states: Inst. of Elec. and Elec. Eng. Computer Society.

[2] Zhou, Y., et al., Exploring the dark side of the web: collection and analysis of u.s. extremist online forums, in Proceedings of the 4th IEEE international conference on Intelligence and Security Informatics2006, Springer-Verlag: San Diego, CA. p. 621-626.

[3] Choi, D., et al., Text analysis for detecting terrorism-related articles on the web. Journal of Network and Computer Applications, 2013(0).

[4] Fu, T., A. Abbasi, and H. Chen, A focused crawler for dark web forums. Journal of the American Society for Information Science and Technology, 2010. 61(6): p. 1213-1231.

[5] Abbasi, A. and H. Chen, Writeprints: A stylometric approach to identity-level identification and similarity detection in cyberspace. ACM Transactions on Information Systems (TOIS), 2008. 26(2): p. 7.

[6] Abbasi, A. and H. Chen, Applying Authorship Analysis to Extremist-Group Web Forum Messages. IEEE Intelligent Systems, 2005. 20(5): p. 67-75.

[7] Zheng, R., et al., A framework for authorship identification of online messages: Writing-style features and classification techniques. Journal of the American Society for Information Science and Technology, 2006. 57(3): p. 378-393.

[8] Huang, C., T. Fu, and H. Chen, Text-based video content classification for online video-sharing sites. J. Am. Soc. Inf. Sci. Technol., 2010. 61(5): p. 891-906.

[9] Tianjun, F., H. Chun-Neng, and C. Hsinchun. Identification of extremist videos in online video sharing sites. in Intelligence and Security Informatics, 2009. ISI '09. IEEE International Conference on. 2009.

[10] Choi, D., et al., Building knowledge domain n-gram model for mobile devices. Information, 2011. 14(11): p. 3583-3590.

[11] Ran, L. and G. Xianjiu. An Improved Algorithm to Term Weighting in Text Classification. in Multimedia Technology (ICMT), 2010 International Conference on. 2010.

[12] Greevy, E. and A.F. Smeaton, Classifying racist texts using a support vector machine, in Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval2004, ACM: Sheffield, United Kingdom. p. 468-469.

[13] Hwang, M., C. Choi, and P. Kim, Automatic enrichment of semantic relation network and its application to word sense disambiguation. IEEE Transactions on Knowledge and Data Engineering, 2011. 23(6): p. 845-858.

[14] Choi, D. and P. Kim, Automatic Image Annotation Using Semantic Text Analysis, in Multidisciplinary Research and Practice for Information Systems, G. Quirchmayr, et al., Editors. 2012, Springer Berlin Heidelberg. p. 479-487.

[15] Yang, Y. and J.O. Pedersen, A Comparative Study on Feature Selection in Text Categorization, in Proceedings of the Fourteenth International Conference on Machine Learning1997, Morgan Kaufmann Publishers Inc. p. 412-420.

[16] Crestani, F., et al., Short queries, natural language and spoken document retrieval: Experiments at Glasgow University, 1998.

[17] Salton, G. and C. Buckley, Term-weighting approaches in automatic text retrieval. Information Processing & Management, 1988. 24(5): p. 513-523.

[18] Chiang, D.-A., et al., The Chinese text categorization system with association rule and category priority. Expert Systems with Applications, 2008. 35(1–2): p. 102-110.

Application of Self Organizing Map for analyzing robotic arm's action with Consciousness-Based Architecture module

Wisanu Jitviriya and Eiji Hayashi

Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
Tel : 81-97-597-7760; Fax : 81-97-597-7760
wisanu@mmcs.mse.kyutech.ac.jp, haya@mse.kyutech.ac.jp

Abstract: Our research has been focused on developing human-robot interactions, The robot need to perform a high level of intellectual activity and user compatibility. Therefore, we considered the primary structure of a conscious human/animal action, which can be represented by the sequence process, Perception → Motivation → Action. We have improved a hierarchical structure model, which defines the relationship between the consciousness field and the behavior module. This model is called Consciousness-Based Architecture (CBA). Furthermore, the robot should select the action itself, we have investigated the application of brain-inspired technology so we introduced a Self Organizing Map (SOM) neural network that is trained using unsupervised learning to classify its behavior according to the motivation value in the CBA module. In this paper, we attempt to describe the integration of a Self Organizing Map (SOM) method into CBA module in order to classify and select autonomous behavior. We confirmed the effectiveness of the proposed system by the experimental results.

Keywords: Consciousness-Based Architecture (CBA), Human-robot interactions, Self Organizing Map (SOM)

I. INTRODUCTION

Nowadays, the rapid development of high technology has produced non-industrial robots such as robotic pets for kids, home robots, therapeutic robots and service robots. These robots are designed for entertainment, guide in museums, education, welfare and other purposes. The necessary challenge in the development of these robots is a high level of intellectual ability and user compatibility. "User compatibility" in the future robot is an important for ease of use, robot friendliness, human-like capricious behavior and the robots are more attractive to people if their movements are not endlessly repeated. However, so far the development of the robots has met with problems in the recognition system, memorizing the situation, the action selection and communication etc., when the robots interact with human. Therefore, our research has been focused on developing the robot that has process of the actions and consciousness resembling that of a human or an animal which is able to enhance the liveliness and the likelihood of their interactions with users [1]-[3]. We have constructed a hierarchical structure model, which defines the relationship between the consciousness field and behavior module in the robot. This model is called Consciousness-Based Architecture (CBA). We have improved the integration of a Self Organizing Map (SOM), one kind of unsupervised artificial neural networks (ANNs) proposed by Kohonen [4]-[6], and based on the neuron organization and decision-making process in the human brain into CBA module. In order to realize a practical robot that can autonomously choose its behavior and adapt to its environment.

In this paper proposes an unsupervised classification method or Self Organizing Map (SOM) combined with CBA module for use with a robotic arm. First, we

Fig.1 System Structure of Conbe-I

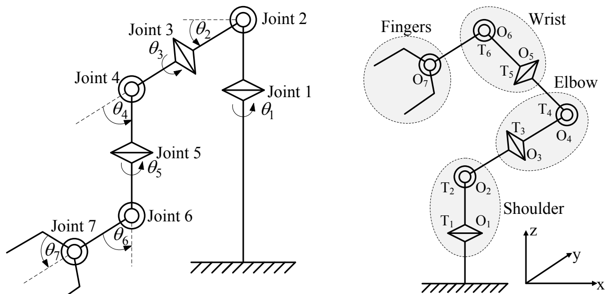


Fig.2 Degree of freedom in the robotic arm and the robotic arm is divided into 4 parts

showed the outline of the proposed overview system and described an autonomous behavior system of a conscious behavior robot (Conbe-I), consisting of a recognition of the external situation using a web camera, calculating the naturally occurring dopamine and determining the robot's motivation, selecting an action using Self Organizing Map combined with CBA module and the drive of the actuator which accepted an action. Next, we described the SOM context that consists of the classification as "liking" behavior and "disliking" behavior when the robot recognizes the green and blue object, respectively, and the normalization process for preparing the input data to be suitable for training. Finally, the kinds of actions such as look around, interest, be alert, approach, catch the object and avoid were classified using a weight matrix derived by dividing the cells on the map based on unsupervised learning, the robot could select an action by finding the maximum similarity of the each action.

II. SYSTEM STRUCTION OF CONBE-I

The figure 1 shows the appearance of a conscious behavior robot (Conbe-I), the robotic arm is divided into 2 parts, an arm and a hand. The arm part has 6 levels of flexibility: the shoulder (Joint1, Joint2), the elbow (Joint3, Joint4) and the wrist (Joint5, Joint6). The hand part has 3 fingers and one level of flexibility. So thus the robotic arm has a total of 7 degrees of freedom as shown in Fig.2 and a small web camera equipped at the palm of the robotic arm which is able to recognize an external situation. The web camera and the multiple actuators are controlled using the RS-485 serial data communication.

III. AUTONOMOUS BEHAVIOR SYSTEM

McCarthy has indicated that a robot will need to consider and introspect in order to operate in the common sense world, it will need to have a consciousness and introspective knowledges and some philosophy [7]-[8], so we describe the overview system by the flowchart of the autonomous behavior system is shown in Fig.3 and the details of each item are described below.

1. Recognize an external situation by a web camera

In this research, The images were simplified by divided into three color groups: green, blue and flesh-color, which were distinguished and recognition in terms of the shape, size and center-of-gravity position as shown in Fig.4. The robot could recognize the position and its distance from the target object in order to calculate the naturally occurring dopamine model.

2. Calculating the naturally occurring dopamine and determining the robot's motivation

When humans and animals take various actions, changes occur in the dopamine in the brain. Dopamine affects brain processes that control movement, emotion response, and ability to experience pleasure and pain, so we used the typical pattern of naturally occurring dopamine for determining the robot's motivation. The waveform of naturally occurring dopamine divided into the raise and fall parts, which are given by Equation (1) and (2) respectively.

$$y''(t) + 2\zeta\omega_n y'(t) + \omega_n^2 y(t) - \omega_n^2 x(t) = 0 \qquad (1)$$

$$Ty'(t) + y(t) = x(t) \qquad (2)$$

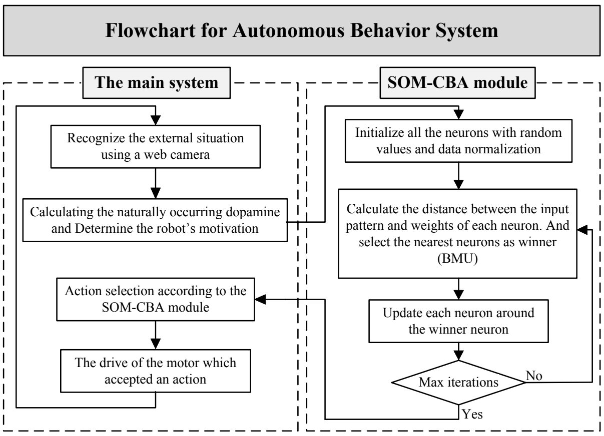


Fig.3 The flowchart of autonomous behavior system

Fig.4 Simplified image

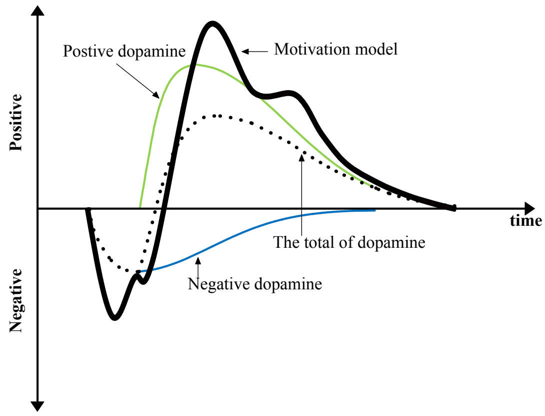


Fig.5 Motivation model

Where, $y(t)$ is the naturally occurring dopamine, $x(t)$ is the input data from image recognition, ω_n is the natural angular frequency, ζ is the damping factor and T is the time constant.

The total of dopamine is calculated by the sum of their positive (the green objects) and negative values (the blue objects), then the total of dopamine is considered as the input variable for determining the robot's motivation as shown in Fig.5 and expressed by a 2nd order system of linear difference equation as shown in Equation (3).

$$\begin{aligned} &motivation''(t) + 2\zeta\omega_n motivation'(t) \\ &\qquad \omega_n^2 motivation(t) = \omega_n^2 u(t) \end{aligned} \qquad (3)$$

Here, $motivation(t)$ is the robot's motivation and $u(t)$ is the total of naturally occurring dopamine.

3. Choose an action using Self Organizing Map method with CBA module

In this section, we are aiming for an integrated development the autonomous behavior system that combines a sample the Self Organizing Map and CBA module in order to classify and select its behavior correspond to its environment and the details are explained below.

A.Self Organizing Map algorithm

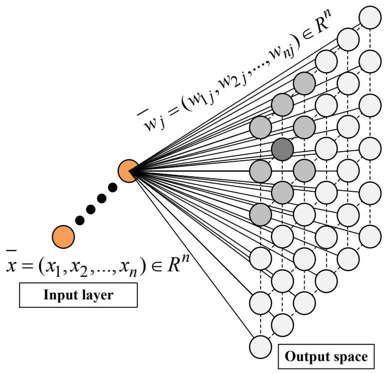


Fig.6 A basic rectangular SOM topology

The basic Self Organizing Map (SOM) consists of the input layer and the output layer which is fully connected with the input layer by the adjusted weights. After learning of the SOM, the weight vector of a particular cell defines the center of a certain class of the input. In Fig.6 a typical two-dimensional rectangular SOM is shown. The input vector $\bar{x} = (x_1, x_2, \dots, x_n) \in R^n$ is connected to cell j and the respective connection has a weight $\bar{w}_j = (w_{1j}, w_{2j}, \dots, w_{nj}) \in R^n$. The learning of the SOM is performed in the following steps:

Step 1. Initially, we assign to the weight vectors $\bar{w}_j(0)$ random values, and we normalize the values of $\bar{w}_j(0)$ in [0,1].

Step 2. Present an input vector $\bar{x}$ to the network, then we compute the cell j that is closest to it by means of the Euclidean distance and then find the winner of unit c (Best Matching Unit: BMU) which has the minimum Euclidean distance as expressed by Equation (4).

$$U_c = \min \left\| \bar{x}_j - \bar{w}_j(t) \right\| \tag{4}$$

Step 3. The adaptation process adjusts the weight vector of cell j and the weight vectors of its neighboring cells by Equation (5).

$$\bar{w}_j(t+1) = \bar{w}_j(t) + \eta(t) \cdot \left[\bar{x}(t) - \bar{w}_j(t) \right] \tag{5}$$

Step 4. The time t increases to t+1, if $t < T$ then go to step 2; otherwise stop the training.

Where, t is the time of iteration, T is the predefined number of iterations and $\eta(t)$ is the learning rate which is decreasing function of time by Equation (6).

$$\eta(t) = \eta_0 \cdot \exp\left(-t/T\right) \tag{6}$$

B.Combine the Self Organizing Map with CBA module

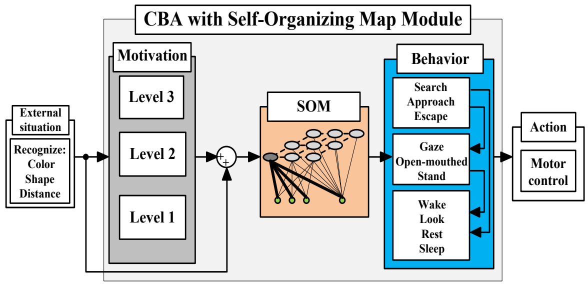


Fig.7 Combine SOM into CBA module

Our previous research, the CBA model decides on the conscious level that the robot must strongly consider and select the behavior corresponding to the conscious level. In this paper, the CBA has been developed as shown in Fig.7, the CBA model has a 3-level hierarchy in the motivation module that connects with a SOM module for clustering actions in a a 3-level hierarchy in the behavior module.

IV. EXPERIMENTAL RESULTS

1. Classification results

The actions of the robotic arm are divided into two characteristic behaviors according to the value of the motivation and the effect of surface color on the object recognition. For instance, when the robot sees a green object (Liking behavior), the motivation value rises and the robot looks around and interests the target object after that the robot attempts to approach or catch the object as shown in Fig, 8. On the other hand, when the robot recognizes a blue object (Disliking behavior), the motivation value drops and the robot should be alert and

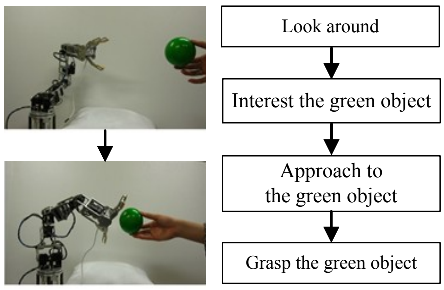


Fig.8 "Liking behavior": recognize the green object

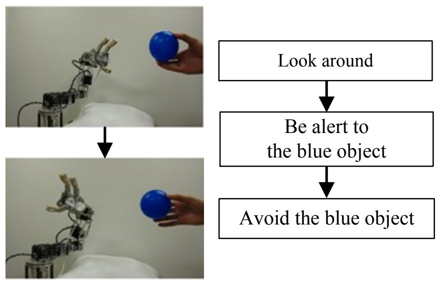


Fig.9 "Disliking behavior": recognize the blue object

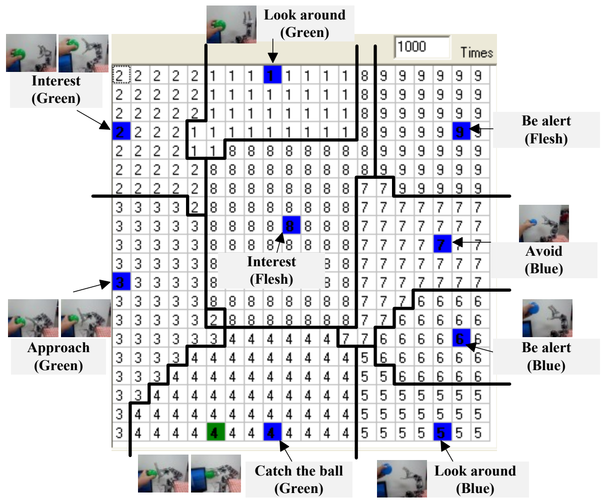


Fig.10 The classification on the map

avoid the object as shown in Fig, 9. In addition, when the robot recognized the flesh color, the actions are clustered around interested behavior and alert behavior. In this paper, the Min-Max Normalization technique is used to scale the input sample in the range of 0 and 1 depending on its relative size. The parameters of SOM as shown in Table 1.

Table 1. Parameters of SOM

	Value	Description
Moti_flag	True of False	The positive motivation value is True = 1, and the negative value is False = 0.
Moti_value	0 to 1	The motivation value
Green	0 to 1	Pixel of green color
Blue	0 to 1	Pixel of blue color
Flesh	0 to 1	Pixel of flesh color
t	1000	The number of learning
η_0	0.35	The initial learning rate
Map size	20 x 20	The rectangular SOM topology size

To make it easy to understand, Fig.10 shown the all actions of the robotic arm which were classified using a weight matrix derived by dividing the cells on the map based on unsupervised learning SOM. For instance, if the input data was Moti_flag = 1, Moti_value = 0.5, Green = 0.8, Blue = 0 and Flesh = 0.1404. The robot recognized the green object and the target object was near the robot's hand, when the system finished the SOM learning. The blue cells were shown in Fig.10, which were the winner node of each behavior and the green cell was the response action from SOM learning, so in this state the robot autonomously selected the "Catch the object" action.

2. Verification experiment

We let the robotic arm recognized a green ball (Liking) and a blue ball (Disliking). We observed the action of the Conbe-I until it took a green ball from the human's hand. The transition of the motivation is shown in Fig.11. After the Conbe-I was started it recognized

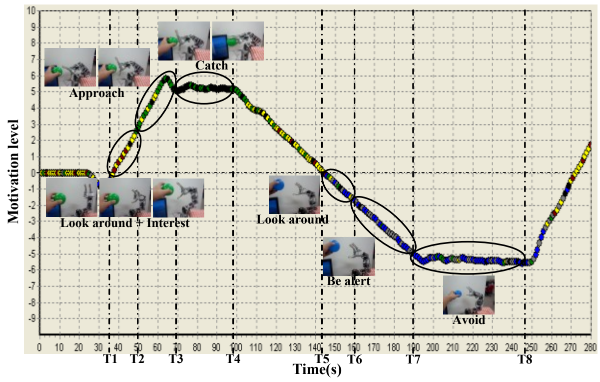


Fig.11 Transition of the motivation

the green ball at T1, then the value of motivation increased steadily and the robot selected the action between looking around behavior and an interesting behavior. From T2-T3, it was selected "Approach" action and it began to approach the green ball and opened the fingers after that the robot came close enough to the green ball to catch it in a time interval (T3-T4), but on the other hand, the blue ball was shown in order to decrease the robot's motivation and the motivation became to the negative value. In the time interval from T5 to T8 that shown the disliking behavior that consists of "Look around", "Alertness" and "Avoidance" respectively.

V. CONCLUSION

In this paper, we present the sample application of Self Organizing Map (SOM) for analyzing and clustering the robotic arm's action. From the experimental results, 9 motions of the robotic arm could be classified on the map of SOM and the robotic arm could autonomously select its behavior and adapt to its surrounding environment. In the future study, our research will continue to develop the system of Conbe-I such as it becomes possible to create an emotional system and apply the memorization system based on an artificial neural network into the robot's system.

REFERENCES

[1] N. Goto and E. Hayashi, "Design of Robotic Behavior that imitates animal consciousness," Journal of Artificial Life and Robotics Vol.12, Springer, pp. 97-101, 2008.

[2] Eiji Hayashi, Takahiro Yamasaki and Koichiro Kuroki, "Autonomous Behavior System Combing Motivation with Consciousness Using Dopamine," IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA2009), pp.126-131.

[3] Motoki Yoshida, Eiji Hayashi, "Design of robotic behavior that imitates animal consciousness: Construction of the user-recognition systems," AROB 17th 2012, Beppu, Oita, Japan

[4] T. Kohonen, "Self-Organizing Maps," Series in Information Sciences, Vol. 30, Springer, Heidelberg. Second ed., 1995.

[5] T. Kohonen, "The self-organizing map," Neurocomputing, Vol.21, pp.1-6, 1998.

[6] T. Kohonen, "Essentials of the self-organizing map," Neural Networks, Vol. 37, pp.52-65, 2013.

[7] McCarthy J (1995) Proceedings of the fourteenth international joint conference on Artifi cial Intelligence, IJCAI 95, pp.2041–2044.

[8] McCarthy J (1996) Making robots conscious of their mental states. In: Muggleton S (ed) Machine intelligence, vol. 15, Oxford University Press, Oxford, pp. 3–17.

Study on Time Synchronization Algorithm of Wireless Sensor Networks Based on Maximum Likelihood Estimation

Dong-Mei Ai[1)], Guang-Ping Zeng[2)], Di He[2)], Xu-Yan Tu[2)]

1) *Mathematics Laboratory, University of Science and Technology Beijing, China*

2) *School of Computer and Communication Engineering, University of Science and Technology Beijing, China*

aidongmei@ustb.edu.cn

Abstract: Time synchronization, a supporting technology for application of WSNs, is critical to the design and application of WSNs. Unlike general computer networks, WSNs face the issue of constraints on energy. Hence, the time synchronization mechanism isn't suitable when extensively applied to traditional network. This paper proposes High-Accurate Energy-Efficient Time Synchronization (HAEE) algorithm based on TPSN algorithm, which significantly improves the accuracy of time synchronization and reduces energy consumption, as proven by software simulation results.

Keywords: Wireless Sensor Networks, time synchronization, maximum likelihood estimation, high accuracy, low consumption

I. INTRODUCTION

Sensing technology is considered to be the essential technology in obtaining information. Along with the progress of modern sensing technologies, outmoded simple means of acquisition of information gradually developed towards integration, miniaturization, and networking. These changes accelerate the development of Wireless Sensor Network (WSN), which integrates technologies of wireless communication, sensing, network interconnection, and distribution. Comprised of a large amount of micro sensor networks, it can collect real-time information, monitor objects within a specific range, and send back usable information immediately to observers for examination and analysis. WSN has a huge potential for application in military safety, environmental monitoring, traffic management, medication, intelligent furniture, etc.

As the supporting technology for WSNs, time synchronization algorithm directly decides on the prospect of WSNs. Unlike traditional networks, WSNs need to deploy a huge number of sensor nodes. It requires cheap and small sizes sensor nodes, so the energy consumption must be lowered.

This paper is a study on several synchronization algorithms and compares their pros and cons. It uses the mathematical method of Maximum Likelihood Estimation (MLE) to improve TPSN synchronization algorithm, which effectively reduces energy consumption while improving accuracy of synchronization. It also describes the building of a software simulation and experiment platform. The simulation result proves the advantages of this improved algorithm over TPSN algorithm.

II. Classic algorithm of time synchronization of WSNs

1. RBS algorithm

Reference Broadcast Synchronization (RBS) [1-3] algorithm was advanced by Elson et al. The basic idea of RBS algorithm is that nodes compare their receipt time, according to reference broadcast sent by a reference node, with each other by exchanging message, as shown in Figure.1. Then, each node calculates its clock offset with all others through the offset matrix [1] (Formula 1). Next, all nodes that have received reference broadcast from the same reference node achieve time synchronization, i.e. subnet synchronization.

$$\forall\ i, j \in [1, n] : offset[i, j] = \frac{1}{m}\sum_{k=1}^{m}(T_{j,k} - T_{i,k}) \tag{1}$$

In RBS algorithm, the features of broadcast channel of wireless data link are used. However, it results in excessive time of message exchange. Simple network with n nodes needs $o(n^2)$ times of message exchange. As the number of nodes increase, the time complexity of RBS algorithm increases rapidly and results in high energy consumption.

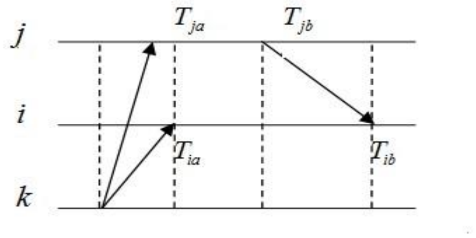


Figure.1. RBS synchronization algorithm message transmission

2. TPSN algorithm

Saurabh Ganeriwal advanced the Timing-Syn Protocol for Sensor Networks (TPSN) [4] synchronization mechanism in November 2003. TPSN adopts the bidirectional paired message synchronization mechanism, through which two modes obtain the offset between each by message exchange. The entire algorithm includes the layer discovery stage where all nodes are divided to layers, and the node synchronization stage, where synchronization between nodes is achieved.

TPSN algorithm significantly increases the energy consumption for time synchronization due to complicated calculation and the use of flooding broadcast hierarchical tree. Once the root node becomes invalid during synchronization in TPSN algorithm, it needs to choose the root node again before proceeding to the layer discovery stage and node synchronization stage, a process that weakens robustness of the algorithm while increasing the work of calculation and energy consumption [5-7].

III. HAEE algorithm

Clock offset between nodes includes two parts: clock-phase offset and clock-frequency offset [8]. Most of the existing time synchronization mechanisms consider only the effect of clock-phase offset on clock synchronization and neglect the effect of clock-frequency offset. In fact, clock-frequency offset, i.e. drift of clock frequency, continually increases the time difference between nodes. So the effect of clock frequency offset should not be neglected. This paper introduces an innovative HAEE algorithm, which introduces the theory of MLE to calibrate clock-phase offset and clock-frequency offset between nodes based on original bidirectional paired message synchronization mechanism [9-11].

HAEE algorithm is divided into three stages, namely, layer discovery stage, leader offset solution stage and node synchronization stage, as shown in Figure.2.

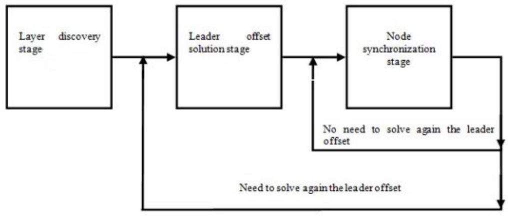


Figure.2. HAEE algorithm framework

1. Layer discovery stage

During the layer discovery stage, every node is assigned a layer number. Besides, each node should have the information of the lower-layer nodes. All sub-nodes in each node build into a group. The upper-layer node randomly assigns one of its sub-nodes as the leader node (Figure.3). After several flooding broadcast, the entire network shall grow to a hierarchical topology.

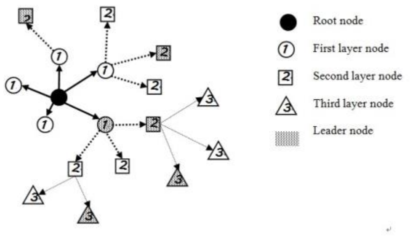


Figure.3. Node topology

2. Leader offset solution stage

The time synchronization between the root node and leader node of first layer nodes is based on bidirectional paired message synchronization mechanism during this stage. The mathematical method of MLE, an important mathematical statistics method whose basic idea is to build a mathematical model based on the density function and to estimate the possible maximum values, is introduced to weaken the effect of indefinite delay and clock frequency offset on clock synchronization.

As shown in Figure.4, node A represents the root node and node B as the leader node of the first layer. θ_1 represents the time difference between nodes A and B at moment $T_{1,1}$ which is kind of clock-phase offset. Meanwhile, θ_2 represents the clock-frequency offset between nodes A and B, which leads in causing increasingly bigger clock offset between two nodes. To simplify the calculation, $T_{1,1}$ is assumed to be the time benchmark, i.e. $T_{1,1} = 0$. The following equation can be deduced from the relationship between $T_{1,i}$ and $T_{2,i}$.

$$T_{2,i} = T_{1,i} + d + X_i + [\theta_1 + (\theta_2 - 1) \bullet (T_{1,i} + d + X_i)] \quad (2)$$

Where $d + X_i$ represents the transmission delay from time $T_{1,i}$ to $T_{2,i}$, d is the definite part and X_i is the indefinite part; $\theta_1 + (\theta_2 - 1) \bullet (T_{1,i} + d + X_i)$ represents the sum of clock-phase offset θ_1 and clock-frequency offset during the time benchmark $T_{1,i}$ and $T_{2,i}$, which is the accumulative effects between node A and node B.

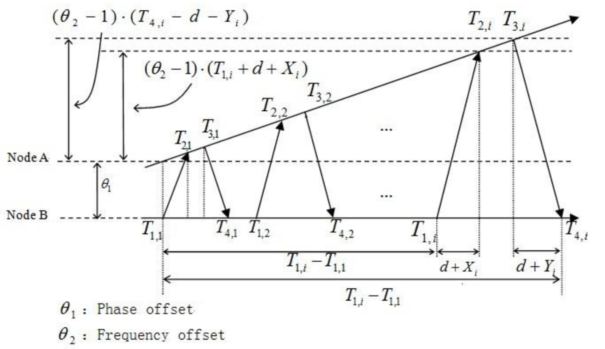


Figure.4. Time synchronization process between root node and first layer's leader node

Similarly, following equations can be respectively deduced

$$T_{3,i} = T_{4,i} - d - Y_i + [\theta_1 + (\theta_2 - 1) \bullet (T_{4,i} - d - Y_i)] \tag{3}$$

$$T_{2,1} = T_{1,1} + d + X_1 + [\theta_1 + (\theta_2 - 1) \bullet (T_{1,1} + d + X_1)] \tag{4}$$

$$T_{3,1} = T_{1,1} - d - Y_1 + [\theta_1 + (\theta_2 - 1) \bullet (T_{4,1} - d - Y_1)] \tag{5}$$

Next, following equations can be deduced from (2) - (4) and (3) - (5):

$$T_{2,i} - T_{2,1} = T_{1,i} - T_{1,1} + X_i - X_1 + (\theta_2 - 1)(T_{1,i} - T_{1,1} + X_i - X_1) \tag{6}$$

$$T_{4,i} - T_{4,1} = T_{3,i} - T_{3,1} + Y_i - Y_1 + (\theta_2 - 1)[T_{3,i} - T_{3,1} - (Y_i - Y_1)] \tag{7}$$

Then, let

$$D_k \Leftrightarrow T_{k,i} - T_{k,1} \tag{8}$$

$$P \Leftrightarrow X_i - X_1 \quad R \Leftrightarrow Y_i - Y_1 \tag{9}$$

The following equations can be deduced from equation (2) and (3), using equation (7), (8) and (9)

$$P = \frac{D_2}{\theta_2} - D_1 = \theta_2^{\bullet} \bullet D_2 - D_1 \tag{10}$$

$$R = D_4 - \frac{D_3}{\theta_2} = D_4 - \theta_2^{\bullet} \bullet D_3 \tag{11}$$

When $\theta_2^{\bullet} = \frac{1}{\theta_2}$

We previously assume that X and Y are the random portion of the transmission delay during the packet transfer process, then the complexity of WSN network data link allows us to see P and R as the accumulation of the independent and identically distributed random delay, and according to the central limit theorem, we can conclude that the probability density function of P and R is close to normal random variable.

Assuming random time sequence $\{X_i\}_{i=1}^N$ and $\{Y_i\}_{i=1}^N$ as normal distribution with mean μ and the variance of δ^2, then P and R is random variables with the mean of 0 and the variance of $2\delta^2$, we can conclude that as follows,

$$f_{P,R}(p,r) = \frac{1}{4\pi\delta^2} e^{-\frac{1}{4\delta^2}(p^2 + r^2)} \tag{12}$$

Using the maximum likelihood estimation method, we can get the clock phase offset and clock frequency offset as follows,

$$\hat{\theta}_1 = \frac{\overline{U_i^{\bullet}} - \overline{V_i^{\bullet}}}{2} \tag{13}$$

$$\hat{\theta}_2 = \frac{D_2^2 + D_3^2}{D_1 \bullet D_2 + D_3 \bullet D_4} \tag{14}$$

Where,

$$D_k = T_{k,N} - T_{k,1}, k = 1,2,3,4$$

$$\overline{U_i^{\bullet}} = \frac{1}{N-1}\sum_{i=2}^{N}(T_{2,i} - \hat{\theta}_2 T_{1,i}) \qquad \overline{V_i^{\bullet}} = \frac{1}{N-1}\sum_{i=2}^{N}(\hat{\theta}_2 T_{4,i} - T_{3,i})$$

N represents the number of package switching during the node synchronization phase, and the bigger the number is, the more accurate synchronization we can get, So after correction the time for Node B will be:

$$T_B^{\bullet} = T_B \bullet \theta_2 + \theta_1 \tag{15}$$

Leader offset solution stage can accurately calculate the clock phase and frequency of leader phase, but the frequent exchange of messages makes it relatively energy-intensive. However, as mentioned at the beginning of this chapter, this stage may cost lots of energy, but for the entire algorithm, the energy consumption actually save lots of trouble in other stages. The detailed reason will be described in the next section.

3. Node synchronization stage

After the leader node solution stage, because of the time offset having been obtained, the difficulty and energy consumption of the node synchronization stage, which includes the self-calibration of leader node and reference synchronization of node synchronization, are significantly reduced,

1) Self-calibration of leader node

In the previous section, clock-phase offset and clock-frequency offset are calculated through the modified bidirectional paired message synchronization mechanism. The clock-frequency offset of the node may be regarded as constant within a period. Thus, self-calibration may be implemented based on the "constant" nature of leader node frequency offset, and time may be calibrated without communication with other nodes.

At the moment $g(t_{N1,i})$, the leader node N1 carries out i^{th} synchronization and calibrates its local time to:

$$t_{N1,i} = \theta_2 \bullet g(t_{N1,i}) - (\theta_2 - 1) \bullet g(t_{N1,i-1}) \qquad (16)$$

Where $g(t_{N1,i})$ represents local time of last synchronization, and $g(t_{N1,i})$ represents local time of current synchronization.

2) Reference synchronization of member nodes

The leader node can maintain extremely high time accuracy in the remaining process of synchronization through self-calibration, a process that makes at least one node in each group of WSNs to have accurate local time and that significantly simplifies the time synchronization of member nodes. In this scenario, reference broadcast time synchronization mechanism, rather than existing time synchronization mechanisms, is more suitable. Taking the upper layer node as the reference node to synchronize the time between member nodes and leader node is simple, easy to implement, effective and energy-efficient.

As shown in Figure.5, where BS represents root node. N1 and N2 represent the leader nodes of the first layer and the member node of the first layer respectively. At first, BS broadcasts at the moment T_1, while the leader node N1 receives the broadcasted message at the moment T_2 and records the local time T_2. Similarly, the node N2 does at $T_2^{\bullet}$. The self-calibration ability of leader node N1 makes that local time of N1 regarded as the standard time. The time offset of N2 at this moment could be represented by its time offset with N1:

$$\theta = T_2^{\bullet} - T_2 \qquad (17)$$

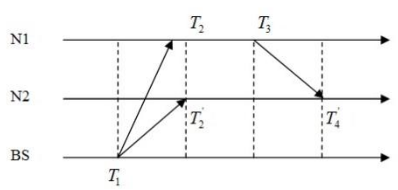


Figure.5. Reference Synchronization of Member Nodes

IV. Simulation and Comparison

1. Analysis of node error in first layer

This study adopts NS-2[12] network simulation software to compare and simulate HAEE and TPSN algorithms in analysis of node error in first layer, node error in different layers, and energy consumption. The simulation results are indicated below.

As shown by Figure.6, it is synchronized once every 500ms, and the time offset is reset every 500ms for the simulation. The figure gives rise to the conclusion that HAEE algorithm proposed by this paper improves the synchronization accuracy in comparison with TPSN algorithm. HAEE algorithm considers the clock-frequency offset and performs better in approximation of standard time, thus achieves more accurate synchronization time.

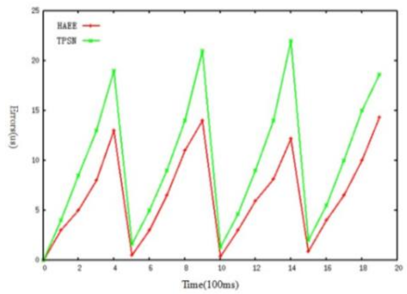


Figure.6. Comparison of Errors of First Layer Nodes at Different Moments in HAEE and TPSN

2. Analysis of node error in different layers

Figure.7 shows that the error is gradually amplified as the number of node layers increases. HAEE algorithm has relatively small error, with less divergence than TPSN algorithm.

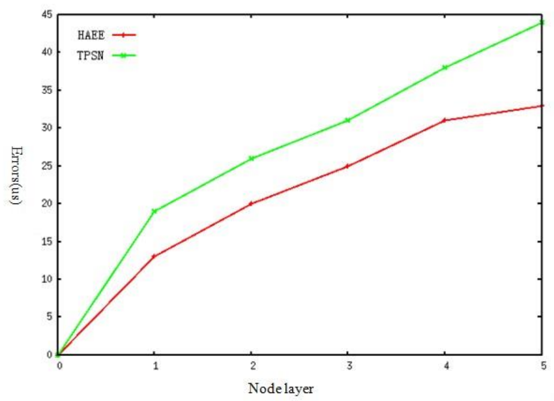


Figure.7. Comparison of error of nodes in different layers in HAEE and TPSN algorithms

3. Energy consumption analysis

As shown in Figure.8, HAEE algorithm does more message forwarding in early synchronizations, thereby causing high energy consumption. As the synchronization goes, HAEE algorithm gradually demonstrates the energy efficiency features in node synchronization in comparison with TPSN algorithm,

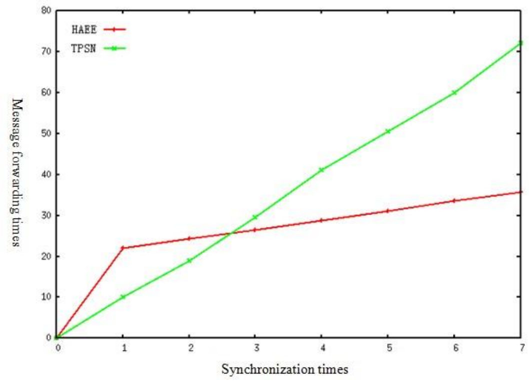


Figure.8. Comparison of energy consumption of HAEE and TPSN algorithms

V. Conclusions

Time synchronization of node is a critical issue to WSNs. Accurate time synchronization will play an active role in the conduct of applied studies on WSNs in the future. This paper proposes HAEE algorithm which refers to broadcast synchronization mechanism, is a modification of TPSN algorithm. HAEE algorithm considers both clock-phase offset and clock-frequency offset, and proves its ability to guarantee synchronization accuracy through simulation analysis. Self-calibration of leader node and reference synchronization of member nodes significantly reduces energy consumption and meets the requirements low energy-consumption. However, the leader offset solution stage entails high energy consumption due to frequent message exchange. Therefore, HAEE algorithm needs further improvement.

References

[1] J.Elson, D.Estrin (2001). Time Synchronization in Wireless Sensor Network [M]. In Proceedings of the 15th International Parallel & Distributed Processing Symposium (IPDPS-01)

[2] WANG Fuqiang, Zeng Peng, ZHANG Xiaoling, LIANG Wei, et al (2011). Survey on Time Synchronization in Wireless Sensor Network. Computer Engineering 37 (22): 70-73

[3] J.Elson, K.Romer (2002). Wireless Sensor Networks: A New Regime for Time Synchronization [J]. Proceeding of the First Workshop on Hot Topics in Networks (HotNets-I): 28-29

[4] David L. Mills (1992). Network Time Protocol (Version 3), Specification Implementation and Analysis. RFC 1305: 31-106

[5] Tao Zhiyong, Hu Ming (2012). Time synchronization algorithms on hierarchical structure in wireless sensor network. Journal of Computer Application 32 (6): 1513-1515, 1551

[6] Yang Peng, Yang Minghui (2011). Energy-Efficient Time Synchronization Algorithm based on Improved TPSN Algorithm. Industrial Control Computer 24 (2): 41-42

[7] Tao Zhiyong, Hu Ming, Fang Ning (2012). Cluster-Based Time Synchronization Algorithm for Wireless Sensor Network. Computer Measurement & Control 20 (7)

[8] Miao Yonggang (2009). Study on Application of Mathematical Method in Wireless Sensor Network. Zhejiang University of Technology

[9] Chaudhari, Q.M. E. Serpedin and K. Qaraqe (2008). On Maximum Likelihood Estimation of Clock Offset and Skew in Networks with Exponential Delays. IEEE Transaction on Signal Processing 56 (4)

[10] Leng, M (2010). On Clock Synchronization Algorithms for Wireless Sensor Networks Under Unknown Delay. Vehicular Technology 59 (1): 182-190

[11] H.S. Abdel-Ghaffar (2002). Analysis of synchronization algorithm with time-out control over networks with exponentially symmetric delays, IEEE Transactions on Communications 50 (10): 1652-1661

[12] Xu Leiming, Pang Bo, Zhao Yao (2003). NS and Analog Simulation. People's Post and Telecommunication Publishing Press

Development of phrase and music search engine by humming

Kiminori Sato and Eiji Hayashi

Department of Mechanical Information Science and Technology

Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology;
680-4, Kawazu, Iizuka-City, Fukuoka Prefecture, Japan;
Tel : 0948-29-7793;
e-mail address: sato@mmcs.mse.kyutech.ac.jp,haya@mse.kyutech.ac.jp;

Abstract: In recent years, the internet has become ubiquitous and is changing dramatically every aspect of how people live their lives. Users not only can purchase the CD at home but also can download the music files without leaving home using the internet, it is very easy and convenient for users who can find the lyrics. But the music's names are not able to be efficiently searched if users forget the music's names. If one could enter a sample of the voice, do a search using that voice's length and pitch, find the rhythm and a similarly pitched phrase, it may be possible to determine the name of the song. Therefore, we developed a phrase and music search engine. It searches for a tune from among various tunes, using one phrase of voice data. For the purpose of studying the relationship between tempo and pitch changes related to the frequency of the input voice data, we developed a method using dynamic programming (DP) matching as a way to search for similar phrases. This system converts notes into character strings, runs DP matching using the strings, and calculates the degree of disagreement between these strings. We then use these calculations as an index to determine whether the strings resemble each other.

***Keywords*:** automatic piano, knowledge database, computer music, DP matching

I. INTRODUCTION

The internet has significantly changed most people's environments. It is possible to search for and obtain information about nearly any subject, and it is possible to purchase many things on the internet.

Music purchases are no exception. Previously, CDs were purchased or rented from stores, but now they can be purchased or rented from home. This is convenient for consumers. However, even if the singer, a part of the musical phrase and the release date are known, a song cannot be accessed if one does not remember the name of the song.

On the other hand, related to music performance, we have developed pedal-powered equipment and typing equipment. The problem of fast repetitive typing with the same key and complaints about performance have been solved, and a system in which piano playing takes place automatically has been developed. That system can make use of data when you play piano with MIDI functions and edited data of loudness, the length of the sound, tempo, etc [1]. In addition, studies were conducted two years ago at time of the development of an interactive musical editing system. That system made it possible to perform efficient editing of large music data files. The automatic piano performs the music of the large musical data files. Finding the desired song requires a time-consuming search for the correct musical data. So, we thought it would be useful to have a music information search engine. The system makes an efficient search for the desired song names and musical data, which can reduce the burden on the user.

In this research we developed a system that uses DP matching to search the entire score to look for a match to a given phrase. This system converts notes into character strings and calculates the degree of disagreement between these strings during searching. We use these calculations as an index to determine whether the strings resemble each other.

In this paper, we describe the results of searching for similar phrases using DP matching.

Figure 1.1: View of the automatic piano

II. Musical Editing Support System

2.1 System Architecture

The structure of the system is shown in Figure 2.1. The user edits music via the user's interface on a

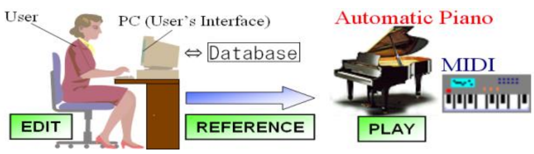


Figure 2.1: Structure of the editing

computer display. The user can also access a database that has musical grammar, the user's preferences, and so on. As a result, editorial work is reduced and efficient editing becomes possible.

2.2 Performance information

The score is read by the scanner. Scanned data is converted to MusicXML format by the KAWAI Score Maker FX5 program. A database of performance information containing notes, scales, and symbols of music data is created based on the MusicXML data. The created database will be subject to search.

2.2.1 Format of Performance Information

The parameters of performance information are shown in Tables 1 and 2. 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".

Table 2.1: Format of the note information

Parameters	Unit (numerical value)	Resolution	Setting
Key	21~108	Stage 88	Pitch
Velo	1~127	Stage 127	Volume
Gate	ms	1ms	Length
Step	ms	1ms	Interval
Time	ms	1ms	Pronunciation time
Bar	-	-	Bar numbers

Table 2.2: Format of the pedal information

Parameters	Unit (numerical value)	Resolution	Setting
Key	119 or 110	-	119=Damper,110=Shift
Velo	0~127	Stage 128	Position of the pedal
Time	ms	1ms	Pedal operation start time
Bar	-	-	Bar numbers

III. Phrase and music search engine

3.1 System summary

We developed a phrase and music search engine. It searches for one tune from among various tunes using one phrase of voice data. The voice is first inputted; then the interval between the sounds of the inputted voice data is converted back to notes according to the tempo of the music. The system then searches for sequence similarity to the phrase rhythm and notes.

In addition, the pitch of the voice input data was analyzed and pitch changes were calculated. Pitch changes are indicated by the ratio of the pitch of a sound to the pitch of the previous one. By using the change in pitch, we hoped to improve the retrieval accuracy of our system.

3.3 Processing that inputs voice

Voice is input in the shape of waves, as shown in Figure 3.1. The vertical axis of Figure 3.1 shows loudness, and the horizontal axis shows time. By setting a threshold on the vertical axis, the presence of voice can be recognized. The interval of this sound is measured as the time. In other words, as shown in Figure 3.1, the interval of sound is the period between the initial iterative peak and the next sound iterative peak.

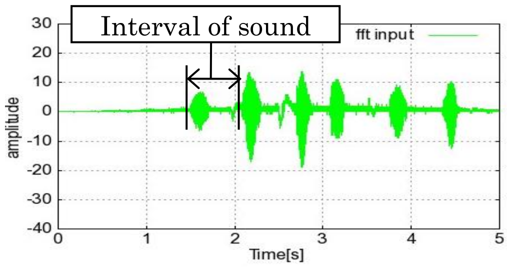


Figure 3.1: Result of sound input

3.4 Converting notes

Measurements of the interval of sound are converted to notes based on the velocity symbol, which appears at the top of the score and represents the number of beats per minute. For example, the indication in Figure 3.2 means 120 quarter note beats per minute. In other words, the length of a quarter note would be 0.5 seconds. The interval of sound is converted into note by using the length of a quarter note.

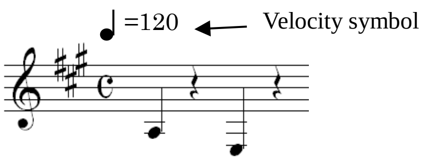


Figure 3.2: Velocity symbol of score

3.5 Frequency analysis of voice data

Pitch is determined by the frequency of the sound, and pitch changes can be determined by the transition of frequency. The voice data frequency analysis is performed using the FFT (fast Fourier transform) algorithm.

Frequency analysis of the voice waveform is conducted for each fixed intervals. And, in one interval, the highest power spectrum frequency is recorded as the main frequency.

IV. Searching for Similar Phrases

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. In the present study, we used DP matching to search for similar phrases.

4.1 DP matching

DP matching is a technique used widely in the field of speech recognition, bioinformatics and so on. It can calculate the similarity between two words that are different from each other in a number of characters.

In Figure 4.1, 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 uppermost level on the right, then it does not increase. Also, if the characters do not correspond in each point, then the cost is increased by 5.

$$g(i,j) = \min\begin{Bmatrix} g(i-1,j) \\ g(i,j) \\ g(i,j-1) \end{Bmatrix} + d(i,j)$$

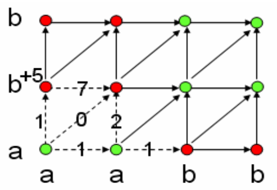


Figure 4.1: DP matching

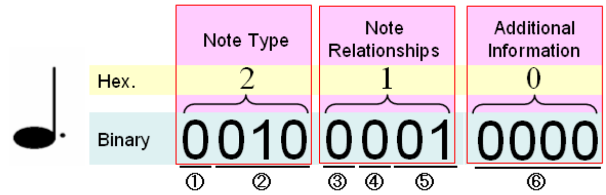
4.2 Searching with DP matching

In this section we describe a method of searching with DP matching. We had to convert a musical score into character strings (a Note Pattern) before searching for similar phrases. This process is explained below.

4.2.1 Patterning

Each note is expressed by three hexadecimal numbers and patterned using only the digits in the first two decimal places, as shown in Figure 4.2 However, the meanings of these two digits vary. Therefore, to avoid confusion, the second digit is distinguished by applying the letter of the alphabet (V from G).

In addition, the sound frequency obtained in Section 3.5 is made into a pattern. First, the system determines the frequency ratio of two successive sounds, and then the ratio of the frequency is patterned by size, as shown in Figure 4.3



1. Note or Rest (0: Note, 1: Rest)
2. Note Value (000: A whole note, 001: A half note …etc.)
3. Tie (0: No tie, 1~: The number of ties)
4. Ornament (1: This note has an ornament)
5. The number of dots
   (except that if it is "11" then the note is a tuplet.)
6. Additional Information
   The number of tuplets (Triplets: 0011)
   The type of ornament (Trill: 0101), etc.

Figure 4.2: Patterning of notes

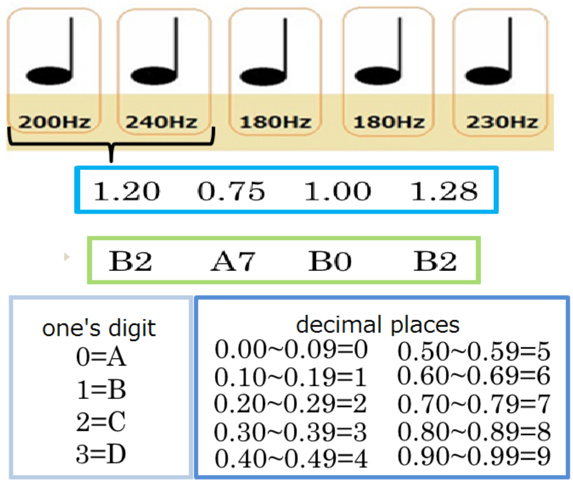


Figure 4.3: Patterning of frequency

4.2.2 The method of searching

A flow chart of a search for a music phrase similar to an input phrase is shown in Figure 4.4.

First, the data of patterned notes column is matched. The phrases with little distance between patterns are searched. Music containing those phrases becomes a search result candidate.

Next, the data of patterned sound frequency ratio are matched. Candidates of the search results are narrowed by matching of the frequency ratio.

The search is performed song by song through the database. When the search engine has processed one song, it then proceeds to the next song, continuing to look for the phrase. When the song changes, search engine is change the tempo of the song. And, search phrase is patterned again. After pattern matching in each of two, search result of the lowest distance between patterns is outputted.

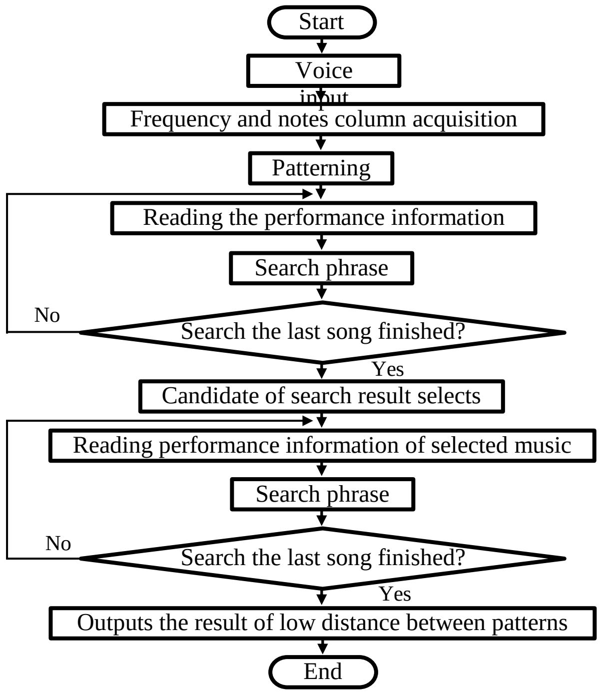


Figure 4.4: The flow of the similar phrase search

V. Retrieval experiment and consideration

We searched for music using the phrase and music search engine. The phrase beginning in bar 5 of "AGEHACYOU" is sung. The music is searched based on the inputted voice data.

First, the data of the patterned notes column is matched. The system output phrases with a distance of 30 or less. Next, the data of the patterned output phrase frequency ratio is matched. Table 5.1&5.2 summarizes the results of the matching.

A glance at the average distance between patterns (Total) of Table 5.1 will reveal that the phrase of "AGEHACYOU" is lower than the phrase of the other music. And a glance at the minimum distance between patterns (Total) of Table 5.2 will reveal that the phrase of "AGEHACYOU" is lower than the phrase of the other music.

Therefore, the input voice is most similar to the phrase of "AGEHA".

Table 5.1: Search result (average)

Name of Music	Distance between patterns (Note column)	Distance between patterns (Frequency)	Distance between patterns (Total)
AGEHACYOU	28.43	57.4	85.83
TENTAIKANSOKU	29.5	72	101.5

Table 5.2: Search result (minimum)

Name of Music	Distance between patterns (Note column)	Distance between patterns (Frequency)	Distance between patterns (Total)
AGEHACYOU	26	47	73
TENTAIKANSOKU	29	70	99

VI. CONCLUSION

We designed methods of searching for similar phrases using DP matching and combined these functions into a single system.

In the similar phrase search, the system was able to find similar phrases using DP matching in a short time, and it was even possible to find phrases whose resemblance might not be immediately apparent.

In the phrase and music search engine, we developed a system to search for a tune by inputting the data of a voice performing the tune. The interval between the sounds of the inputted voice data is converted to a note according to the tempo of the music. Using this technique, we can search for sequence similarity to a phrase’s rhythm and notes. In addition, the search using the patterns of frequency ratio became possible. The search using the patterns of frequency ratio could improve the retrieval accuracy.

In this study, we were able to perform similar phrase searches and searching for a tune by voice. In our future research we will perform evaluations with different pieces of music and will evaluate the existing system.

REFERENCES

[1] Hayashi, E. et al, “Behavior of piano-action in a grand piano.I”, Journal of acoustical Society of America, Vol.105, pp.3534-3544, 1999.

[2] Hayashi, E., et al, Interactive musical editing system for supporting human errors and offering personal preferences for an automatic piano, Proc. of the 7th International Symposium on Artificial Life and Robotics, Vol. 2, pp. 513-516, 2002.

[3] Hikisaka, Y., Hayashi, E., et al, Interactive musical editing system for supporting human error and offering personal preferences for an automatic piano –Method of searching for similar phrases with DP matching and inferring performance expression-, Proc. of the 12th International Symposium on Artificial Life and Robotics, GS4-3, 2007.

Human Recognition based on Gait Features and Genetic Programming

Dipak Gaire Sharma, Ivan Tanev and Kasunori Shimohara

Graduate School of Science and Engineering, Doshisha University
Kyotanabe City, Kyoto, 610-0394, Japan
Telephone : +81-774-65-6200, Fax : +81-774-65-6800
{sharma2013@sil.doshisha.ac.jp, itanev@mail.doshisha.ac.jp, kshimoha@mail.doshisha.ac.jp}

Abstract: Human walking has always been the curious field of research for different disciple of Social and Information Science. The study of human walk or human gait in association with different behaviors and emotions has not only fascinated social science researchers, but its uniqueness has also attracted many computer scientists to work in this arena for the quest of uncovering reliable mechanisms of biometric identification. In this research, we used a novel method for human identification based on inferring the relationship between the human gait features via genetic programming. Moreover, we focus on generating a unique numerical signature that is similar for different locomotion gaits of a particular individual but different across different individuals

Keywords: Human Identification, Biometrics, Genetic Programming, Human Gaits, Nature Inspired Computing

I. INTRODUCTION

Analysis of human behavior involves different methods to identify or recognize human. It is done by the visual analysis where different characteristics are studied to authenticate the people [1]. This approach is gaining substantial amount of interest because it is also driven by the task of automated human identification which is very crucial in different fields that are sensitive to the issues of security. There are different methods that are implemented for this task, and among them gait recognition is a new dimension unlike gait classification. Gait is a rhythmic physical movement of body parts. As analysis of gait does not require a person to face the sensors directly, it can be done without any notice by the subject, because of this feature, it has been the important dimension of research in computer vision among different security agencies throughout the world.

Human locomotion is said to be unique. It is unique in a sense that, human being holds distinctive ability to determine their close friend and family easily. It was proven by Johannson[2] in 1973 by the experiment where he attached light markers to the subject, and people were asked to identify the subject based on those attached markers. It is also unique because many animal behavior such as walking or swimming require rhythmic contraction of muscles which is generated by the signals from the neurons called Central Pattern Generator (CPG) [3] [4]. These CPG are responsible for generating rhythm and shapes the pattern of the motor neurons. Basically each and every individual have unique walking pattern, therefore, it can be considered as one of the most reliable source of human recognition.

This paper is the finding in a research of gait feature recognition where genetic programming (GP) is used to generate unique signature among different individuals. Although there are variety of genetic programming approaches available, in this case XML-based Genetic Programming [5] approach was deployed. The study on human gait analysis using a genetic programming approach is relatively an understudied problem. Therefore, our study will also be the milestone towards the analysis of efficiency and reliability of GP in case of human gait analysis and human recognition.

II. PREVIOUS WORK

Human gaits recognition and human recognition by motion is one of the active topics in the field of Computer Research. Many existing approaches focus on analyzing of motion using the video frame and using the different processing techniques [6] [7] [8]. We can find vast research done in this field using different dimensions. However, some are focused on hand limb gait features while other are based on merging the features and most use classifiers from the gait classification record such as Back Propagation Neural Network Algorithm, Fisher Distance, Support Vector Machine and K-Nearest Neighbor. If we analyze different research that has been done so far, we could find that most of the research are focused on the classification of gait features but none of these research focuses of finding the relationship between the several gaits of individuals. Therefore, this research is concerned on generating a unique numerical signature which is similar for different motion gaits of a particular individual but different across different individuals which further can be used as a mechanism of human recognition.

III. COMPUTATIONAL APPROACH

This study involves different computational steps and methodologies that were implemented to achieve the experimental result of this research, these steps includes;

1. Data Acquisition

The initial step involves the technique of identifying a person while in motion and generates the skeleton frame from the captured data. For this task, a device from Microsoft called Kinect™ for Windows® is used. This device has the ability to detect human and generate skeleton with 20 different joints. To interface this device, we developed an application that could track a person. A sample snapshot of this application is shown in Figure 1.

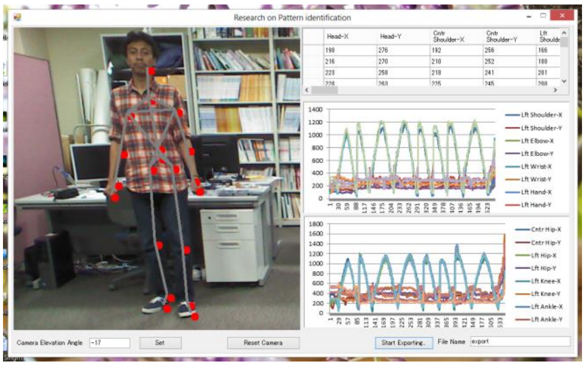


Fig.1: Application Interface for analyzing the gaits to identify the major joints that are used to determine the features

2. Smoothing Parameters

The next step comprises of method to generate the numerical dataset for the skeleton previously generated where joint coordinates are accessed from the device and are stored in a data file. Since data from the Kinect™ are not consistent for our work, the task of smoothing those dataset are required, which is done using the mathematical model known as sliding window average method, as shown in Equations 1, 2, 3 and 4. This method is important because it help to exclude the extreme parameters and helps to increase the probability of minimizing noise level in our dataset.

$$y_1 = \frac{1}{k}(x_1 + x_2 + \cdots + x_k) \qquad \text{(Eq.1)}$$

$$y_2 = \frac{1}{k}(x_2 + x_3 + \cdots + x_{k+1}) \qquad \text{(Eq.2)}$$

$$y_3 = \frac{1}{k}(x_3 + x_4 + \cdots + x_{k+2}) \qquad \text{(Eq.3)}$$

$$\cdots$$

$$y_{n-k+1} = \frac{1}{k}(x_{n-k+1} + x_{n-k+2} + \cdots + x_n) \qquad \text{(Eq.4)}$$

3. Feature Extraction

The third step is basically concerned with feature finding. These features are calculated from the dataset which are previously obtained after performing mathematical model. For this task we made some basic calculation and extracted the values of major features of human gaits. These include the angle between joints, the angular displacement, velocity, distance between joints, rate of change in distance, etc. These features are calculated for few selected joints from arms, legs, hip and shoulder. As the features are calculated frame by frame, they were reduced to a single value by considering the average of each feature. As a whole, total of 30 features of dataset were evaluated and stored in database to use as an offline source. We used these features as the terminal set in genetic programming.

4. Genetic Programming

The fourth step involves implementation of genetic programming framework. Since our problem is concerned with finding relationship between human gaits which is not known in advance as well as does not possess any predefined method of performing the operation, the use of genetic programming is the best technique to implement for this task. For the purpose of evolving the association between gaits feature, we used our XML-based Genetic Programming framework [5] (XGP). Beside this, the communication with different sub-systems and calculation of fitness value from fitness evaluator is also performed in this phase. A sample snapshot of the XGP applied for the recognition of human gaits is shown in Figure 2. The main parameters of XGP are show in Table 1.

IV. EXPERIMENTAL RESULTS

We implemented the system with Intel Core 2 Quad processors each with 2.5GHz speed, and 2GByte of physical memory. We acquired the data about the feature of human gaits in the Socio Informatics Laboratory of Doshisha University. Data were collected for one individual at the initial experimental test, where the individual was asked to make three different gait movements starting from (i) normal walk, (ii) slow walk and (iii) fast walk, respectively. The motion detection sensor was placed perpendicular to the walking direction at the distance of 3 to 4 meters.

The experimental results of the fitness convergence of 20 independent runs of XGP are shown in Figure 3. Table 2 shows the values of the signatures of the three human gaits, obtained from a sample best-of-run genetic program. Notice that despite the fact that these three gaits have different values of their respective features; the values of the tree signatures are very close. The obtained fitness value is very small too, and it is equal to 2.

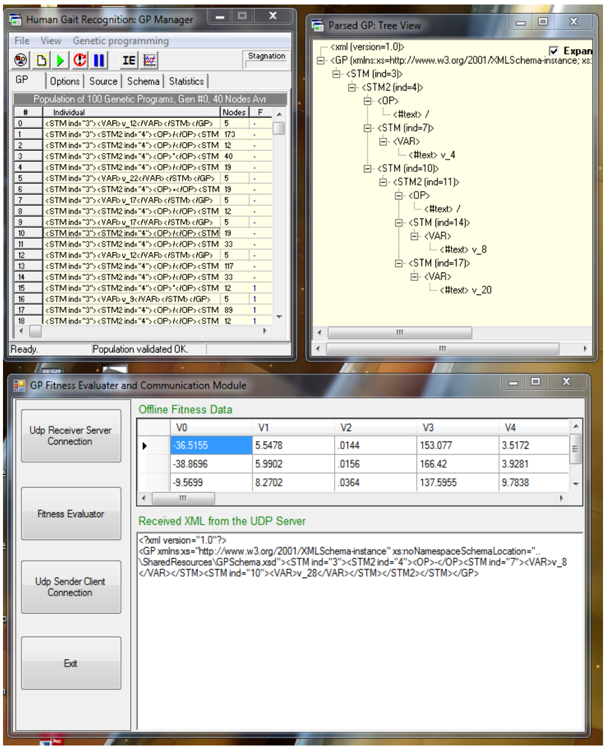


Fig.2: XGP applied for recognition of human gaits: GP Manager (top) manages the population of genetic programs, and the fitness evaluation subsystem (bottom), which calculates the signatures for three gaits of the same person, and measures the quadratic deviation of these three signatures.

Table 1: Main Parameters of XGP

Parameter	Value
Terminal Set	(i) Variables v_0,v_1,v_2,......,v_30 representing the features of the human gait, and (ii) a random integer constant within the range [0..10]
Function Set	{+, -, *, /}
Population Size	100 genetic programs
Selection Ratio	10%
Mutation Rate	2%
Selection Method	Binary Tournament Selection
Fitness Value	Quadratic deviation of three different gaits scaled up 10000 times. Lower fitness values correspond to better solutions
Bloat control	Parsimony pressure
Termination Criteria	(Fitness value<=2) or (# Generations = 100)

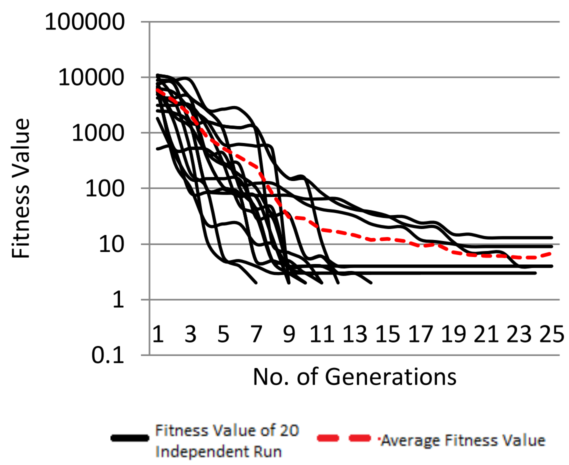


Fig.3: Fitness convergence characteristics of 20 independent runs of XGP. The dashed line represents the convergence of the average fitness value.

Table 2: The values of signatures of three human gaits, obtained from sample best of run genetic program (fitness value=2).

Gait	Value of the Signature
Normal Walk	17.0000869346814
Slow Walk	17.000062104833
Fast Walk	16.9998314813823

V. CONCLUSION

The characteristics shown in Fig.3 illustrate how the fitness improves as the evolution progresses; the improvement is dramatic and the average fitness decreases from the initial value of around 7000 to around 8, which is almost 1000 times. The result is obtained relatively quickly within about 20 generations, which correspond to about 4 hours of runtime. Such an impressive performance of XGP could be explained by the fact that so far we tried to evolve very close signatures of three different gaits of a *single person*, and we have yet to evolve similarly close to each other three signatures of the gait of *another person*. Moreover, the signatures of the gait of both persons should be distinguishably different. We anticipate that the addition of additional person(s) in the gait identification and classification task would significantly increase the computational effort of XGP.

The future work includes experiments with multiple individuals and researching on the methods to evolve the distinguishable signatures of their respective gaits.

REFERENCES

[1]L. Wang, W.M. Hu, and T.N. Tan, “Recent Developments in Human Motion Analysis,” Pattern Recognition, vol. 36

[2]Johannson, G.: Visual Perception of Biological Motion and A Model For Its Analysis. Perception and Psychophysics.

[3]Jacques Duysens, Henry W.A.A. Van de Crommert, Neural control of locomotion; Part 1: The Central Pattern Generator from cats to humans.

[4]Levitan, Kaczmarek, The Neuron Cell and Molecular Biology,Third Edition, Oxford

[5]I.Tanev and K.Shimohara, XML-based Genetic Programming Framework: Design Philosophy, Implementation, and Applications, Artificial Life and Robotics, Vol.15, No.4.

[6]Hu Ng, Hau-Lee Tong, Wooi-Haw Tan, Timothy-Van Yap, Pei-Fen Chong, Junaida Abduallah. Human Identification Based on Extracted Gait Feature

[7]Liang Wang, Tieniu Tan, Huazhong Ning, and Weiming Hu. Silhoutte Analysis-Based Gait Recognition for Human Identification.

[8]Nikolaos Gkalelis, Anastasios Tefas, Ioannis Pital. Human Identification from Human Movements.

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

An Action Selection Method Using Degree of Cooperation in a Multi-agent System

Masanori Kawamura, Kunikazu Kobayashi

Aichi Prefectural University
1522-3 Ibaragabasama, Nagakute, Aichi 480-1198, Japan.
Tel: :+81-561-64-1111; Fax: +81-561-64-1108
kobayashi@ist.aichi-pu.ac.jp

Abstract: In recent years, a concept of a dividual is proposed to interact properly with another person. To construct a model of the dividual, the degree of cooperation is assigned to the corresponding dividual. By introducing the degree of cooperation into multi-agent systems, we evaluate what kind of changes appears in the agent behavior. In addition, we propose an action selection method by introducing the degree of cooperation into the soft-max method in multi-agent systems. Using the proposed method, we confirm whether the cooperative action is promoted or suppressed through computer simulations.

Keywords: Multi-agent system, Reinforcement leaning, Cooperative action, Dividual, Degree of cooperation,

I. INTRODUCTION

In the human society, a person act cooperatively by taking some kinds of communication such as gesture, language, and eye contact. In recent years, a concept of a dividual is proposed by Hirano to interact properly with another person [1]. At present, by introducing the above concept into a multi-agent system [2]–[4], we construct a model of the dividual to realize cooperative behavior.

In this paper, we treat a difference of how to interact with another person, which characterizes the dividual model. When a self-agent recognized the other agent, a dividual is formed in the self-agent. At the same time, the degree of cooperation proposed in this paper is assigned to the corresponding dividual.

By introducing the measure into multi-agent systems, we evaluate what kind of changes appears in the agent behavior. In addition, in this paper, we propose an action selection method by introducing the degree of cooperation into the soft-max method in multi-agent systems. Using the proposed method, we confirm whether the cooperative action is promoted or suppressed through computer simulations.

In the simulations, we prepare for two agents and one goal, two agents act in a discrete grid world. The agents acquire a route to the goal using the Q-learning method [5], which is one of the representative reinforcement learning methods [6] [7]. In the proposed method, the agent identifies the direction of the other agent such as the right, left, top, or bottom direction. Then the agent reflects the degree of cooperation in its action probability. To evaluate the degree of cooperation, we prepare two levels of the degree. The performance of the proposed method is evaluated by the distance between agents in every step at the last episode. From the simulation results, when the degree of cooperation of two agents is both high, the agents became close at few steps from the beginning and arrived at the goal with keeping the distance. On the other hand, when the degree of the cooperation is low, the agents did not become close in all the steps and did not reach the goal. As a result, it is revealed that the cooperative action is promoted or suppressed by the degree of cooperation. In addition, it is clarified that the agent gave priority to cooperation over arriving to a goal.

In section II, an action selection method using degree of cooperation is proposed. In section III, the performance of the proposed method is evaluated through computer simulations. In section IV, we give a summary of this paper and describe future problems.

II. PROPOSED METHOD

First of all, we define a problem. Then, we explain a concept of a dividual and the degree of cooperation. After that, we proposed an action selection method using the degree of cooperation.

1. PROBLEM DEFINITION

We treat a problem that plural agents arrive the same goal in a grid world in Fig.1.

2. DIVIDUAL

A person usually communicate by using how to contact other person properly. For example, you face your boss, I bet you communicate politely with him/her. But you face a good friend, I bet you communicate friendly with him/her. Although we give a simple example, a person may change a way to interact with other person according to sex, nationality, relationship with him/her. This concept is named as a dividual by Hirano [1].

A different dividual interacts with the other person, i.e. families, friends, and acquaintances. The number of dividual therefore corresponds to that of other persons who interact with. When you communicate with person *A* and become *A*'s acquaintance, *A*'s dividual is constructed in yourself. Similary, when you communicate with person *B* and become *B*'s acquaintance, *B*'s dividual is also constructed

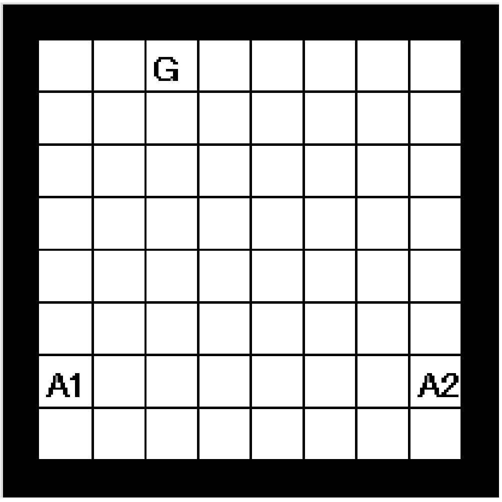


Fig. 1. An example of field.

in yourself. It is suggested that a set of your dividuals may characterize a human personality.

In this paper, when a dividual is created, its degree of cooperation is defined so as to cooperate with other person.

3. DEGREE OF COOPERATION

In this paper, the degree of cooperation c which corresponds to its dividual is defined. The c takes a scalar value and fall within the range of $0 \leq c \leq 1$. When a dividual is firstly created, a social dividual is defined as a default dividual and grown by interacting each other. The degree of cooperation for the social dividual is defined as 0.5. If $c > 0.5$ and $c < 0.5$, the degree of cooperation is regarded as high and low cooperation, respectively. The high and low degrees of cooperation promotes and suppress the cooperation.

4. ACTION SELECTION METHOD USING DEGREE OF COOPERATION

In this section, we propose an action selection method using the degree of cooperation in order to cooperate with others. In the proposed method, we identify the direction of the other agent and realize the cooperation by reflecting the degree of cooperation toward its direction. This is explained using Fig.2.

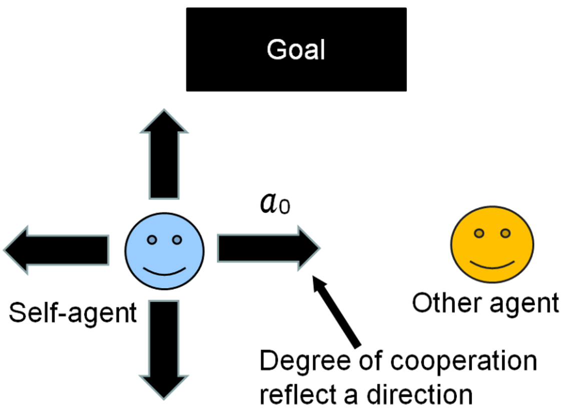


Fig. 2. Relationship between two agents.

In Fig.2, there is a goal in the top direction and a cooperative partner in the right direction. Let us assume that the self-agent can select an action out of moving the right, left, top, or bottom. The proposed method realizes that the self-agent tends to select an action toward the other agent as a_0 out of available actions by reflecting the degree of cooperation on the probability of selecting a_0.

In the proposed mehod, we use the soft-max action selection method. But, our approach apply to ϵ-greedy method. The proposed action selection method is defined as Eq.(1).

$$\pi(a|s) = \frac{\exp(Q(s,a)\,g(c))}{\sum_{b \in A} \exp(Q(s,b)\,g(c))}, \cdots\cdots\cdots\cdots\cdots (1)$$

where A is a set of available actions. The degree of cooperation is introduced as a function $g(c)$. In proposed method, $g(c)$ is defined as Eq.(2).

$$g(c) = \begin{cases} 2c & (a = a_o) \\ 1 & (otherwise) \end{cases} \cdots\cdots\cdots\cdots\cdots\cdots\cdots (2)$$

We assume that the degree of cooperation for the social dividual is set to 0.5. The value of function $g(c)$ is calculated as $g(0.5) = 1$ so that the probability for the social dividual does not change at all. This corresponds to not considering cooperation with a stranger. On the other hand, when the degree of cooperation is larger than 0.5, it becomes $g(c) > 1$ and the cooperation is promoted. When the degree of cooperation is smaller than 0.5, it becomes $g(c) < 1$ and the cooperation is suppressed. Therefore, function $g(c)$ is set as Eq.(2).

III. COMPUTER SIMULATIONS

The proposed method is evaluate using the 10×10 field in Fig.1. In the field, the black surrounding is wall, there are two agents A_1 and A_2 and one goal G. The agents can select an action out of moving the right, left, top, or bottom and their aim is moving towards the goal.

The number of episodes is set as $10,000$ and the maximum number of steps is limited to 100. The performance is evaluated by the distance between two agents at every step. In case of a cooperative agent, the degree of cooperation is set as 0.8 and in case of a non-cooperative agent, it is set as 0.2. In this simulation, agents can only move four directions, i.e. up, down, right, and left. The minimum numbers of steps toward the goal for agents A_1 and A_2 are 8 and 11 steps, respectively.

Firstly, the transition of the distance between two agents whose degrees of cooperation are both high is shown in Fig.3. The image of an action sequence in the field is also illustrated in Fig.4. In these figures, the degree of cooperation from A_1 to A_2 is denoted by $C(A_1, A_2)$.

As a result, two agents approach each other and move toward the goal while keeping close distance. Therefore, it followed that they cooperated when each other's degree of cooperation was high. In addition, considered minimum step of agent A_2 was 11 steps, it is thought that they arrived at a goal after they drop in a little by they give priority to cooperate.

Secondly, we show the transition of the distance between two agents when the degree of cooperation from A_1 to A_2 is high but that from A_2 to A_1 is low in Fig.5. The image of

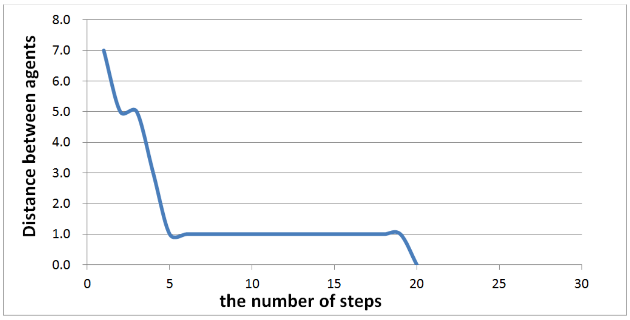


Fig. 3. Transition of distance between agents when $C(A_1, A_2) = 0.8$ and $C(A_2, A_1) = 0.8$.

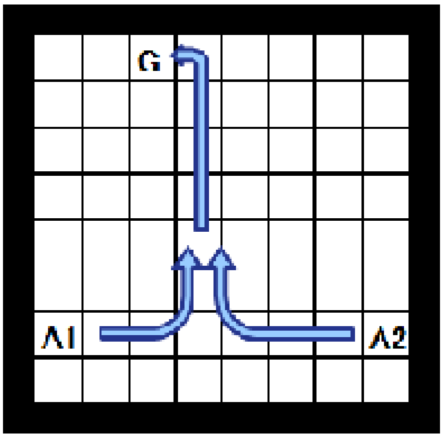


Fig. 4. An action image when both $C(A_1, A_2)$ and $C(A_2, A_1)$ are high.

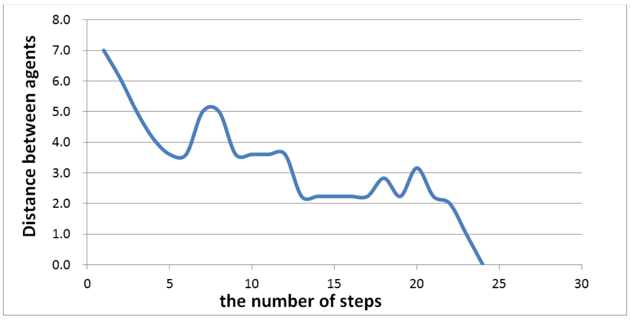


Fig. 5. Transition of distance between agents when $C(A_1, A_2) = 0.8$ and $C(A_2, A_1) = 0.2$.

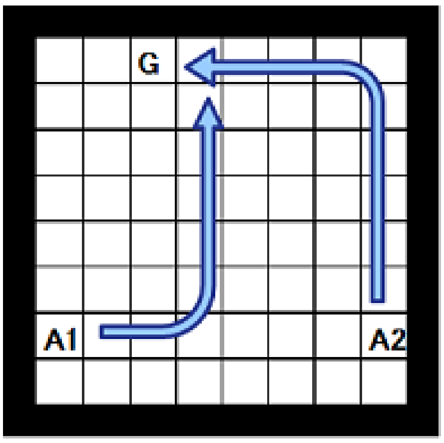


Fig. 6. An action image when $C(A_1, A_2)$ is high and $C(A_2, A_1)$ is low.

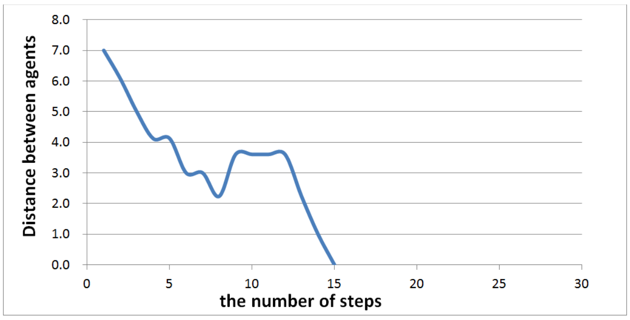


Fig. 7. Transition of distance between agents when $C(A_1, A_2) = 0.2$ and $C(A_2, A_1) = 0.8$.

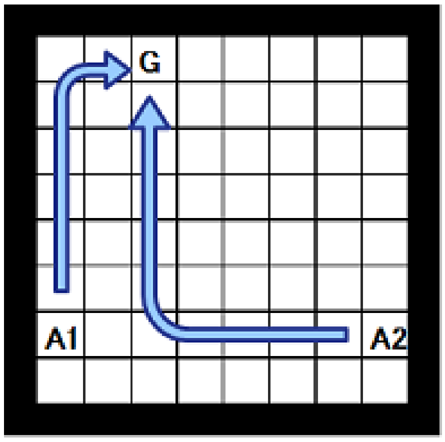


Fig. 8. An action image when $C(A_1, A_2)$ is low and $C(A_2, A_1)$ is high.

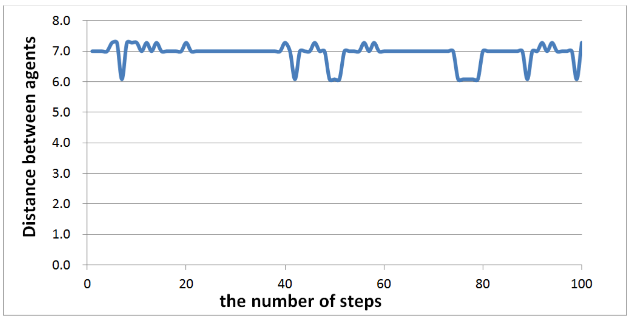


Fig. 9. Transition of distance between agents when $C(A_1, A_2) = 0.2$ and $C(A_2, A_1) = 0.2$.

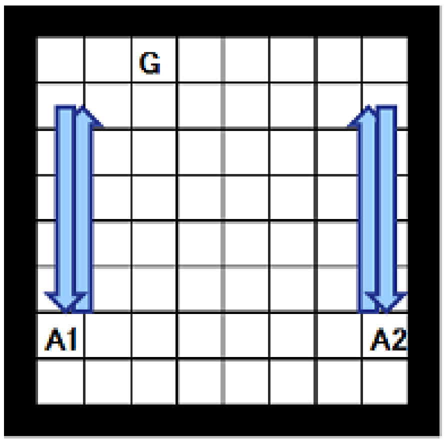


Fig. 10. An action image when both $C(A_1, A_2)$ and $C(A_2, A_1)$ are low.

an action sequence in the field is also illustrated in Fig.6.

As a result, the distance between agents does not decrease rapidly but become gradually close toward the goal. Therefore, when $C(A_1, A_2)$ is high and $C(A_2, A_1)$ is low, cooperative action is slightly suppressed. In addition, as considering the number of steps for A_2, it takes 24 steps until it arrives at the goal. This shows that A_1 gives first priority to cooperate with A_2 than to arrive at the goal and A_2 gives first priority not to cooperate with A_1 than to arrives at the goal.

Thirdly, we show the transition of the distance between two agents when the degree of cooperation from A_1 to A_2 is low but that from A_2 to A_1 is high in Fig.7. The image of an action sequence in the field is also illustrated in Fig.8.

As a result, the distance between agents does not also decrease rapidly but become gradually close toward the goal. Therefore, when $C(A_1, A_2)$ is low and $C(A_2, A_1)$ is high, cooperative action is slightly suppressed. In addition, as considering the number of steps for A_2, it takes 15 steps until it arrives at the goal. This number of steps is a little bit smaller than the above case, i.e. $C(A_1, A_2) = 0.8$ and $C(A_2, A_1) = 0.2$ because the goal position is closer to A_1 than A_2. This shows that A_2 gives first priority not to cooperate with A_1 than to arrives at the goal and A_1 gives first priority to cooperate with A_2 than to arrive at the goal.

Finally, we show that the transition of the distance between two agents whose degrees of cooperation are both low is shown in Fig.9. The image of an action sequence in the field is also illustrated in Fig.10.

As a result, the distance between agents does not decrease at all. This shows that both agents give first priority not to cooperate with each other than to arrives at the goal.

IV. CONCLUSION

In this paper, we focused on the concept of the dividual and introduced the degree of cooperation. Then, we proposed the action selection method based on the soft-max method using the degree of cooperation. Through computer simulations, it was verified that two agents with high degree of cooperation approached each other and arrived at the goal. On the other hand, it was clear that two agents with low degree of cooperation did not approach each other and arrive at the goal neither. In addition, when one agent with high degree of cooperation and the other agent with low degree of cooperation, it was shown that the distance between agents became gradually low and finally agents arrived at the goal.

ACKNOWLEDGMENTS

This work was partly supported by Grant-in-Aid for Scientific Research (No.23500181) from JSPS, Japan.

REFERENCES

[1] Hirano K (2012), Who am I?: From individual to dividual. Kodansha shinsho (in Japanese)

[2] Stone P and Veloso M (2000), Multiagent Systems: A Survey from a Machine Learning Perspective. Autonomous Robots, 8:345–383

[3] Ohuchi A, Yamamoto M, and Kawamura H (2003), Basics and Applications of Multi-agent Systems. Corona Publishing (in Japanese)

[4] Takadama K (2003), Multi-agent Learning. Corona Publishing (in Japanese)

[5] Watkins CJCH and Dayan P (1992), Q-learning. Machine Learning, 8:279-292

[6] Sutton RS and Barto AG (1998), Reinforcement Learning: An introduction. MIT press

[7] Kaelbling LP, Littman ML, and Moore AP (1996), Reinforcement Learning: A Survey. Journal of Artificial Intelligence Research, 4:237–285

Cooperative Action Acquisition Based on Intention Estimation Method in a Multi-agent Reinforcement Learning System

Tatsuya Tsubakimoto, Kunikazu Kobayashi

Aichi Prefectural University
1522-3 Ibaragabasama, Nagakute, Aichi 480-1198, Japan.
Tel: :+81-561-64-1111; Fax: +81-561-64-1108
kobayashi@ist.aichi-pu.ac.jp

Abstract: In this paper, we propose a method that can acquire cooperative action to reach an appropriate goal without controlling reward to realize cooperative action by designers. We assume every action of other agents as a process to achieve an ultimate goal and then treat it unimportant. In order to confirm the effectiveness of the proposed method, we carried out computer simulations. The simulation results show that the proposed method is superior to a standard Q-learning method and a Q-learning method with cooperation in terms of the number of successful cooperation.

Keywords: Multi-agent system, Reinforcement learning, Cooperative action, Reward interpretation, Intention estimation

I. INTRODUCTION

Current machines such as robots and computers can only be operated according to predesigned actions by human. In other words, it is impossible for machines to think and act by itself like human. Machines have therefore low flexibility to cope with little environmental changes. To solve the above problem, reinforcement learning has been proposed as a method for getting autonomous actions [1] [2]. In reinforcement learning, a learning agent like robots can perceive environmental information by itself and select an appropriate action. As a result, it can autonomously acquire an optimal action sequence through given reward or punishment from an environment. Q-learning is one of representative reinforcement learning methods. Q-learning is originally designed for a single-agent system but is recently introduced to a multi-agent system [3]–[5].

In this paper, our goal is the acquisition of cooperative action by Q-learning in multi-agent environment. Nagayuki et al. presented a policy estimation method which can estimate the other's action to be taken based on the observed information about the other's action sequence [6]. The policy estimation method however is required to store all the state and action sequences and then has state explosion problem.

In this paper, we assume every action of other agents as a process to achieve an ultimate goal and then treat it unimportant. The important thing is the intention of other agents, namely their goals. Then, we try to acquire cooperative action by intention estimation. Consequently, we use a small amount of storage capacity. We propose a method that can acquire cooperative action to reach an appropriate goal without controlling reward to realize cooperative action by designers. To accomplish this, we introduce the concept of "reward interpretation". The reward interpretation is a method that can increase or decrease the reward given by an environment through the interpretation of agent itself and reflect it to Q-values. Then, we use two important variables in this method. One is a goal priority. This corresponds to an estimate of priority which goal the other agent reaches. The other is a goal value. This means a value of each goal that the agent estimates. The system flows is summarized as follows.

The agent calculates the goal priority at each step and then updates goal value by using goal priority. When it reaches a goal, it estimates the other's intention using the goal value corresponding to the goal. By repeating this process, it gradually acquires cooperative action sequence. In addition, we try to acquire an effective policy adapted to the environment in minimal learning steps. In order to confirm the effectiveness of the proposed method, we carried out computer simulations. In the simulations, we prepares plural agents and plural goals, and aim to acquire the cooperative action. In the simulations, we compare three methods, namely the standard Q-learning method, a Q-learning method designed to give reward only when all agent act cooperatively, and the proposed method. The simulation results show that the proposed method is superior to the other two methods in terms of the number of successful cooperation.

In section II, a reinforcement learning method based on intention estimation is proposed. In section III, the performance of the proposed method is evaluated through computer simulations. In section IV, we give a summary of this paper and describe future problems.

II. PROPOSED METHOD

In this section, we propose an intention estimation method. An important concept in the proposed method is *reward interpretation*. In the standard reinforcement learning such as Q-learning, an agent directly receives rewards from an environment. An agent incorporating intention estimation, however, receives rewards through this reward interpretation.

By interpreting rewards, the agent can easily acquire cooperative actions. The reward interpretation requires two variables, i.e. a goal priority and a goal value. The goal priority denotes an estimate that the other agent prefers which goal. The goal value represents an estimated value that the self-agent considers how much the goal has value.

1. Estimation of other agent's Intention

In this paper, we propose two measures, i.e. a goal priority and a goal value. Firstly, a self-agent estimates a goal priority. The goal priority $GP_{t,g,k}$ represents a measure that the self-agent predict for other agent k to arrive to goal g at time t by observing its action. Note that there are plural goals. As calculating $GP_{t,g,k}$, the distance d and the angular θ between directions of agent a and goal g as shown in Fig.1.

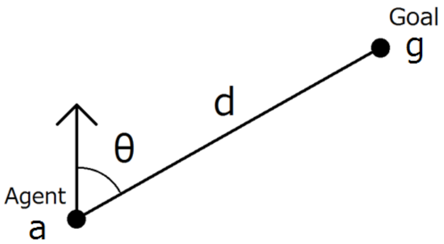


Fig. 1. Positional relation between an agent and a goal.

$GP_{t,g,k}$ is calculated by Eq.(1).

$$GP_{t,g,k} = \left(\alpha\, Dir_{t,g,k} + \beta\, Dis_{t,g,k} + Dir_{t,g,k}\, Dis_{t,g,k}\right) / (\alpha + \beta + 1)\,, \quad (1)$$

where α and β are weighting parameters and generally set to 1.0. $GP_{t,g,k}$ has a normalized value as $0 \le GP_{t,g,k} \le 1$. $Dir_{t,g,k}$ and $Dis_{t,g,k}$ are defined by Eq.(2) and Eq.(3), respectively.

$$Dir_{t,g,k} = 0.5\,(\cos\theta + 1), \quad (2)$$
$$Dis_{t,g,k} = d^{-1}, \quad (3)$$

where $Dir_{t,g,k}$ and $Dis_{t,g,k}$ have a normalized value as $0 \le Dir_{t,g,k} \le 1$ and $0 \le Dis_{t,g,k} \le 1$, respectively.

Secondly, the self-agent calculates a goal value. The goal value $GV_{t,g}$ represents a measure that the self-agent considers how much the goal has value. $GV_{t,g}$ is updated by Eq.(4).

$$GV_{t,g} = (1 - \eta)\, GV_{t-1,g} + \eta\, e(\overrightarrow{GP}_{t,g}), \quad (4)$$
$$\overrightarrow{GP}_{t,g} = (GP_{t,g,1},\, GP_{t,g,2},\, \cdots,\, GP_{t,g,n}), \quad (5)$$

where η is a learning rate, $e(\cdot)$ is an evaluation function defined by a designer and n is the number of other agents. $\overrightarrow{GP}_{t,g}$ is a vector which contains all of goal values $GP_{t,g,k}$ except for the self-agent. $GV_{t,g}$ has a normalized value as $0 \le GV_{t,g} \le 1$. Note that $GV_{t,g}$ reflects intention which estimated by other agent's action.

2. Reward Interpretation

As introducing the reward interpretation, the self-agent can adjust a reward value given from an environment. When the self-agent reaches goal g at time t, the reward interpretation is conducted by the goal value defined in Eq.(6) as soon as the self-agent receives a reward r_g from an environment.

$$r_t \leftarrow r_t \times 2\, GV_{t,g}. \quad (6)$$

The agent updates the Q-values by interpreting the reward and acquires the cooperative action. $GV_{t,g}$ denotes a goal value for goal g at time t. If the self-agent assumes that the goal value is worth at a maximum, i.e. $GV_{t,g} = 1$, the agent interprets the reward as double, i.e. $2r_t$. On the other hand, If the self-agent assumes that the goal value is not worth at all, i.e. $GV_{t,g} = 0$, the agent interprets the reward as 0.0. If $GV_{t,g} = 0.5$, the self-agent interprets the reward as standard.

3. Proposed Algorithm

The proposed algorithm is described as follows.

(1) Let the present time be t. The self-agent calculates $GP_{t,g,k}$ for all the combination of goal g and other agent k. Then it calculates $GV_{t,g}$ for all the combination of g.

(2) The self-agent determines action a according to it's policy π and executes a.

(3) The state s_t transits to the next state s_{t+1} after executing a. If the self-agent reaches the goal, it gets a reward and then interprets it.

(4) The self-agent updates $Q(s, a)$ by the Q-learning method. Return to the first step after $t \leftarrow t + 1$ until a terminal condition is satisfied.

III. SIMULATION

1. Problem Setting

- We prepare a two-dimensional grid field (20×20) with two agents and two goals shown in Fig:2. The field is surrounded by the black wall. Two agents and two goals place randomly.

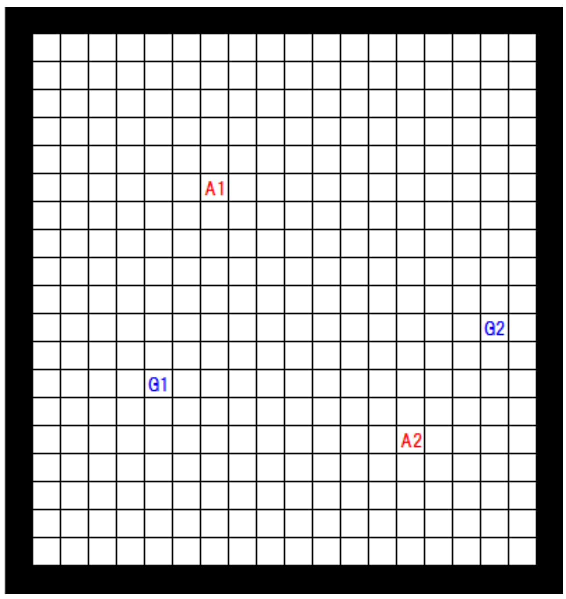


Fig. 2. A simulation field

- The agents have to haul a heavy box, which only one agent cannot carry. In this simulation, the positions of two boxes correspond to those of two goals. The aim is for two agents to reach the same goal by acquiring cooperative action.
- An agent can move up, down, left or right to an adjacent cell.
- One episode is defined that two agents are placed at the initial position and reach the same goal.
- If two agents cannot reach the same goal until 300 steps, terminate the episode and start a new episode.
- If an agent reaches one of two goals, the agent stops moving and learning despite whether cooperation succeeds or not. If both two agents reach the goal, an episode terminates and the agents replace to their initial positions.
- One trial is defined for 10,000 episodes.
- An agent gets 20 when it reaches a goal, −0.1 when it bumps into a wall, and −0.01 otherwise as a reward from an environment.
- An agent can completely observe the other agent's position and action.

2. Parameter Setting

For Q-learning, a learning rate is 0.1, a discount rate is 0.9. A temperature parameter T is scheduled by the following equation.

$$T = \frac{T_{max}}{1 - \exp\{a(t+\phi)\}},$$

where T_{max} is the maximum temperature, a is a gain, t is a time (episode), and ϕ is an initial phase. In this simulation, we set that $T = 20$, $a = 0.01$, $\phi = 500$.

An agent needs information on goal positions for intention estimation method. An agent can get the information only when it reach a goal and store it. At an initial state, an agent cannot calculate a goal value because of no information on goal positions. An agent start calculating a goal value once it find a goal. An agent with the proposed method therefore behaves an agent with a normal Q-learning until it find a goal.

A function $e(\cdot)$ set as an average function $E[\cdot]$ because two agents reach the same goal. We set $\alpha = 1$ and $\beta = 1$ for calculating a goal priority because distance and direction has same importance. Then we set learning rate η as 0.3.

3. Evaluation Method

The standard Q-learning and a Q-learning with cooperation are used for performance evaluation. The Q-learning with cooperation is assumed that two agents get reward if and only if they arrive at the same goal. In other words, it is assumed that two agents do not get reward if they arrive at the different goal.

The aim of this simulation is for two agents to reach the same goal. Successful episode is assumed that two agents reach same goal, we count it every 10 episodes. The maximum number of success is 10. We evaluate the successful number of cooperation.

4. Simulation Results

We show simulation results for three methods, i.e. the standard Q-learning, Q-learning with cooperation, and the

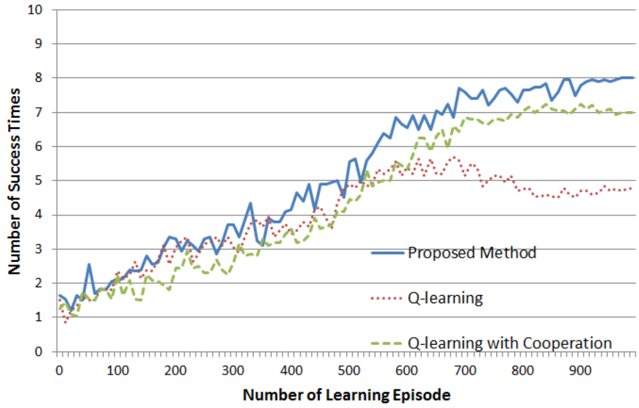


Fig. 3. Simulation results

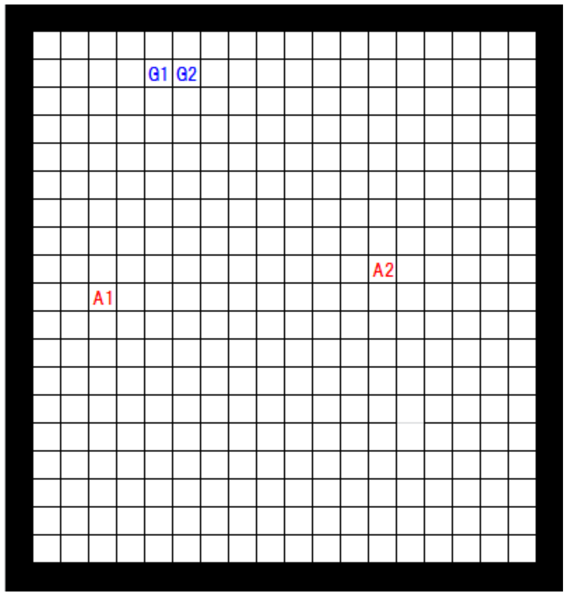


Fig. 4. A placement that two agents are hard to cooperate(case1)

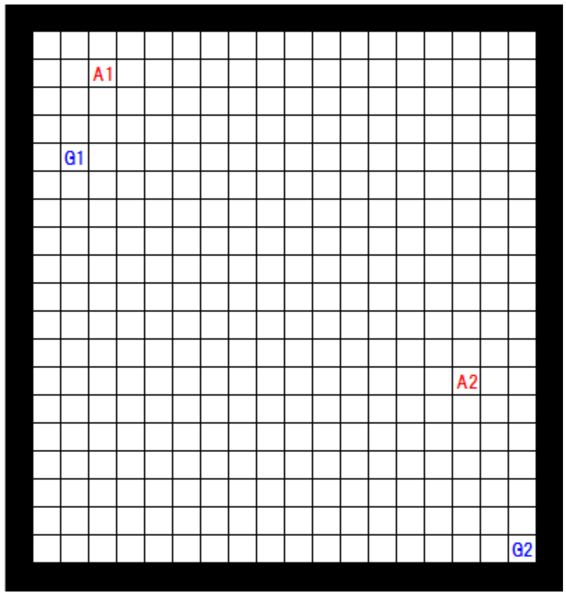


Fig. 5. A placement that two agents are hard to cooperate(case2)

proposed method in Fig.3. This simulation aim to acquire cooperative action in a multi-agent system. We conducted the simulation for 20 times. Figure 3 shows averaged results for 20 trial.

5. Discussion

As seen in Fig.3, there are no significant difference among three methods at the beginning. The proposed method and Q-learning with cooperation show a good performance after 600 episodes. This is because the standard Q-learning has no cooperative function. The proposed method is slightly better than Q-learning with cooperation.

We found two types of initial placements for evoking uncooperative action in the proposed method. One is an initial placement which two goals are very close as shown in Fig.4. Because an agent judge two goal values as almost same, it tends to select the closest goal. The other is an initial placement which two agents and two goals are far away as shown in Fig.5. An agent tends to reach the closest goal because another goal is far away from it. An agent therefore cannot update Q-value adequately. It is shown that an agent acquire cooperative action for an initial placement which two distances between agents and between goals are almost same as depicted in Fig.2.

In Q-learning with cooperation, a designer has to decide to give a reward whether cooperation succeeds or not at every step. In the proposed method, we do not need the above design.

Furthermore, memory capacity required by the proposed method is much smaller than that for the policy estimation method. In this simulation, required memory is $(400\times 4) + (1 \times 2) = 1,602$ for the intention estimation method (proposed method) and $(400 \times 4 \times 4^1) + (400 \times 4) = 8,000$ for the policy estimation method. The difference becomes much significant as increasing the number of agents.

IV. CONCLUSION

We have proposed the new method for autonomously acquiring cooperative action in a multi-agent system. At the same time, the proposed method solved a memory problem in the policy estimation method.

Besides, we compared three methods, i.e. the proposed method, Q-learning, and Q-learning with cooperation in a multi-agent system in order to evaluate the proposed method. The aim of this simulation is to acquiring cooperative action. We have shown that the proposed method can acquire cooperative action without changing reward by a designer. Finally, we confirmed that the proposed method is superior to the standard Q-learning and the Q-learning method with cooperation.

In this paper, we assume that an environment can be fully perceived by other agents. As future developments, we will extend it to partially observable environment and the agent cooperative even if the agent has limited sight. Furthermore, we try to treat more complicated environments.

ACKNOWLEDGMENTS

This work was partly supported by Grant-in-Aid for Scientific Research (No.23500181) from JSPS, Japan.

REFERENCES

[1] Sutton RS and Barto AG (1998), Reinforcement learning: An introduction. MIT press

[2] Kaelbling LP, Littman ML, and Moore AP (1996), Reinforcement Learning: A Survey. Journal of Artificial Intelligence Research, 4:237–285

[3] Stone P and Veloso M (2000), Multiagent Systems: A Survey from a Machine Learning Perspective. Autonomous Robots, 8:345–383

[4] Ohuchi A, Yamamoto M, and Kawamura H (2003), Basics and Applications of Multi-agent Systems. Corona Publishing (in Japanese)

[5] Takadama K (2003), Multi-agent Learning. Corona Publishing (in Japanese)

[6] Nagayuki Y and Ito M (2003), A Reinforcement Learning Method with the Inference of the Other Agent's Policy for 2-Player Stochastic Games. IEICE Transactions on Information and Systems, J86-D-I:821–829 (in Japanese)

Integrating the Event Generation Mechanism in the Propp-based Story Generation Mechanism into the Integrated Narrative Generation System

Shohei Imabuchi, and Takashi Ogata

Iwate Prefectural University, Takizawa, Iwate, Japan
Tel : 019-694-2000
g231k005@s.iwate-pu.ac.jp, t-ogata@iwate-pu.ac.jp

Abstract: In the Propp-based story generation mechanisms and the integrated narrative generation system which we have been developing, the former functions as a module in the latter. These systems have used respective event generation mechanisms to generated events which are the most important units in a narrative structural representation. A common event generation mechanism needs to be used in these two systems towards the complete blending. This paper will present the first tentative attempt as a prototype to be revised in the future.

Keywords: narrative generation; story generation; integrated narrative generation system; Propp-based story gene ration mechanism; event generation mechanism.

I. INTRODUCTION

One of the current major goals in our study is developing an integrated narrative generation system based on artificial intelligence, cognitive science and litery theories. The system aims to organically combine a variety of mechanisms or modules grounded on our previous fundamental and individual research results into one organic mechanism. We have described its ultimate philosophy and overall design [1], and introduced the current status of its actual implementation [2]. A significant characteristic of this system's research is the point that it adopts an interdisciplinary principle with such literary theories and narratology as Genette [3], Jauss [3] and inter-textuality. Propp's theory on "morphology of the folktales" [4] to be described in this paper is also equal to one of the literary theories relevant to some significant functions for the narrative generation. We have been considering the following two types of approaches for the literary theory by Propp: the first is developing a story generation mechasnism using Propp's theory as comprehensively as possible ("Propp-based story generation mechanism") and the second approach is its introduction as one of the modules into the integrated narrative generation system [5].

As described above, the Propp-based story generation mechanism can function independently to generate stories with mainly the structural style of a Russian fairy tale which is the material of Propp's theory. On the other hand, we have been addressing the issue of its use as one module of the integrated narrative generation system [6]. For example, we have introduced the Propp-based mechanism into the integrated system as a set of techniques to generate macro structures of stories, and defined a story content grammar relating to the conceptual dictionaries based on our original interpretation and formalization of an essential part of Propp's theory. The event generation mechanism in the Propp-based mechanism, however, has not been connected fully organically to the integrated narrative generation system. An event means one of the most fundamental elements in each of the narrative structures to be generated by the Propp-based mechanism and the integrated narrative generation system.

The event description forms in the two systems are basically same, but mechanisms for the respective event generation mechanisms are not unified formally since the Propp-based event generation mechanism has some special techniques to generate stories with the style of a Russian folk tales and is comparatively simplified. The Propp-based event generation mechanism, for example, has a function to set seven types of characters including hero, villain, victim, helper, dispatcher, donor and false hero in the Russian fairy tales into an agent in an event, while the event generation mechanism in the integrated system does not have the function. This paper considers the two event generation mechanisms in detail to synthesize them.

Gervás [7] describes that many of story generation systems using Propp's theory aim at general story generation mechanisms beyond the description of Propp even though the most important and essential part in story generation study using Propp's theory should be conducted on the basis of the valuable insight in accordance with the corpus of Propp. We have also been insisting that most of the systems using Propp' theory focus on only the element of the "function" to be mentioned below, but the theory includes many productive and theoretical ideas to contribute to story generation in various sides. One of our goals of the Propp-based story generation mechanism is implementing a story generation system introducing a variety of elements of Propp's theory as comprehensively as possible, while another goal is integrating this special approach into the framework of a more generalized architecture of the integrated narrative generation system. We simultaneously pursue specialty and generality.

II. AN OVERVIEW OF THE INTEGRATED NARRATIVE GENERATION SYSTEM

The integrated narrative generation system consists of the following three types of generation phases: story, discourse and expression. The system architecture, on the other hand, has the following parts from the type of processing: narrative techniques, knowledge bases, and conceptual dictionaries. For the conceptual dictionaries [8], we have been developing systems for verb concepts, noun concepts, adjective concepts and adjective verb concepts with the respective hierarchical structures from higher concepts to lower concepts according to a "is-a" relation. We differentiate a word's meaning or concept from the word itself as the fundamental policy of developing the conceptual dictionaries. In particular, one or more meanings related to a noun are described in the noun conceptual dictionaries as one or more noun concepts, and the word to be used normally and the notations for these concepts are described in another linguistic dictionary. The noun conceptual dictionary contains 115765 terminal concepts and 5808 intermediate concepts in the current version. The verb conceptual dictionary similarly has 12174 terminal concepts and 36 intermediate concepts. A terminal verb concept in the hierarchical structure has the following information: a basic sentence pattern for natural language generation, one or more case structures and constraints for each of the cases. A case structure defines several cases for nouns required in the verb. In addition, knowledge for managing relations among events and mutual relationships between events & states are composed on the basis of the conceptual dictionaries for verb and noun concepts. The input information by a user to the first story generation phase is several parameters including length, unreality and repetiton. The system selects one or more adequate story techniques according to the parameters in order to expand a story tree structure.

Although the Propp-based story generation mechanism can be used as an independent story generation system, its integration into the story generation phase in the integrated narrative generation system contributes to increase the diversity and flexibility in story generation. A story is described, in the story generation phase, as a conceptual representation form including a temporal sequence of events at the bottom level in a hierarchical tree structure with several types of relations as the intermediate nodes. Each of the events is described as a specific case structure which has a verb concept and the necessary instantiated noun concepts linked to the conceptual dictionaries. Fig. 1 shows a story structure and a case structure of an event within the story structure. The story generation mechanism generates story structures having a sequence of events according to structural operation techniques using the dictionaries and the other narrative knowledge. We call a procedure for expanding the tree structure of a story using the story knowledge base, which stores concrete narrative fragments or structures corresponded to the material for narrative generation for such various types of narrative relations as scripts and causal relations, a narrative technique for story generation or directly a story technique. The input to a story technique is an event, the whole of a story structure or a sub-structure in a story structure and the output is chiefly the expanded story structure that is one of the possible various structures. They can treat from micro story structures to macro ones.

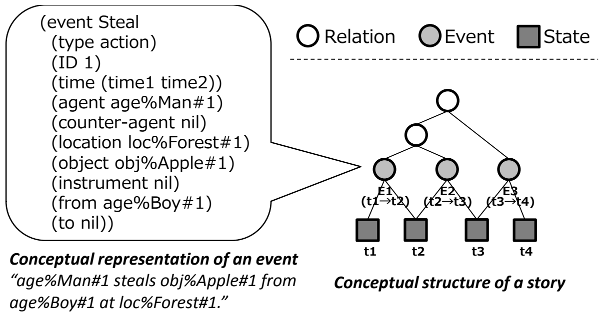


Fig.1. Structures of a story and an event

III. PROPP-BASED STORY GENERATION SYSTEM

1. The overview

Propp's theory proposed the concept of a "function", which means an action seen from the result and the principle of the sequential arrangement of thirty one "functions" as the most important theoretical idea. The "function" of "villainy", for example, means that a character causes the damage of a different character through a concrete action. However, there is a variety of specific actions to actualize it since the "function" means a collective definition to abstractly define many concrete and real actions. We call more specified actions for actualizing each of the "functions" "sub-functions" on the basic of hierarchical concept. Propp described a variety of examples corresponded to the level of sub-functions. The "Propp-based story content grammar" based on the framework of Propp's theory with "functions" and sub-functions is the central part in our Propp-based story generation mechanism.

We describe the Propp-based story generation mechanism. A story as the output information is equivalent to a sequence of events bound hierarchically using "story relations" including "causal relation", "continuation relation", etc. The form of an event as one of the basic components in a story takes a frame representation, which is formed by a verb concept and the corresponding case elements for such noun concepts as agent, object and location. The verb concepts and noun concepts are associated with each of the corresponding elements in the conceptual dictionaries. The structure generation mechanism and the event concept generation mechanism are respectively corresponded to generative mechanisms, while the Propp-based story content grammar and the conceptual dictionaries are equivalent to the knowledge parts to be referred by the generative mechanisms. Since these generative parts and the parts to be referred by them are mutually independent, the redundancy of processing lessens.

2. Propp-based Story Content Grammar

The story content grammar [5, 6] that is a new style of the reorganization of Propp's theory, which gradually generates a macro story with the structure of the style of a Russian fairy tale from the highest level to the lowest level, consists of the following five hierarchical levels. The actual description by Common Lisp is shown in Fig. 2. The elements including "reserve portion", "beginning", etc. in the level 1 and 2 are defined based on the description of structures of Russian fairy tales in [4]. The levels of "functions" and sub-functions were faithfully defined from the description of the analysis of 100 tales. The role of these hierarchical levels is to make the structural part of a story which is a tree structure with relations as the intermediate nodes. Events in the level 5, on the other hand, are generated by the event generation mechanism associating with the above structure generation mechanism. An event particularly forms a case structure including a verb concept and the corresponding noun concepts as noted above. We have defined by hand patterns of these case structures based on examples corresponded to sub-functions in [4]. For example, we have formally defined a case structure of "(interdict 1 (agent human) (object outgo@out) (to !hero))" based on an example in the sub-functions

```
(setq *Propp-level1-list*
 ' ((Russian folktale
    (00_Preliminary-part P-Problem P-Trial P-Solution))))
(setq *Propp-level2-list*
 ' ((P-Problem (Reserve-portion Beginning))
    (P-Trial (OR (Reserve-trial Battle-and-victory)
                 (Reserve-trial Task-and-solution))) …))
(setq *Propp-level3-list*
 ' ((Reserve-portion (OR (01_Absentation 02_Interdiction
                           03_Violation) …))
(setq *Propp-level4-list*
 ' ((01_Absentation (OR bet-1_Going-out1 bet-2_Death
                        bet-3_Going-out2))
    (02_Interdiction (OR gam-1_Interdiction
                         gam-2_Command-or-proposal)) …))
(setq *Propp-level5-list*
 ' ((alp-1_Preliminary-part ((Preliminary-part (1))))
    (bet-1_Going-out1 ((Go-out (agent Parents)
                               (object Business))))
    (bet-2_Death ((Dead (agent Parents)))) …))
```

Fig.2. A part of the description of the Propp-based story content grammar

under "interdiction (γl)" as "If Baba Jaga comes, don't you say anything, and be silent." There is a asset of "Pair of sub-functions" as another significant elements. This was defined according to the description on mutual relations between two sub-functions in Propp's theory.

IV. COMBINING THE EVENT GENERATION MECHANISM WITH THE INTEGRATED NARRATIVE GENERATION SYSTEM

In the Propp-based story generation mechanism, when the structure generation mechanism reaches to the lowest level in the story content grammar, the event generation mechanism is executed to make the event structure based on the format of the case structure described in the story content grammar. On the other hand, in a story generation mechanism in the integrated narrative generation system, when a story technique refers to the corresponding unit in the story content knowledge base, the generation mechanism is executed to make the event structure based on the format of the case structure described in the story content knowledge base. The story knowledge base in the integrated narrative generation system stores concrete knowledge for story generation, in particular a variety of sequences of events described in the form of case structure corresponded to the types of relations such as "cause-effect" and "script". The story content grammar in the Propp-based story generation mechanism will be also a part of the story knowledge base in the future. Two types of event generation mechanism have been independently prepared for the Propp-based mechanism and the integrated system since both of the description forms are different each other. To develop a common event generation mechanism is the goal for incorporating the Propp-based mechanism into the integrated system completely. We will describe previous four types of procedures according to the different types of case structure's description in the Propp-based mechanism and the tentative revision in the integrated narrative generation system.

When the generation of a story structure according to the story content grammar in the Propp-based story generation mechanism reaches to the level of event generation, the event generation mechanism is executed to make the case structure. The mechanism inserts specified instances into each of the cases based on the difference of the second term in each element of the case structure. Fig.3 shows an original case structure and a transformed instance by the following techniques in the event generation mechanism.

(1) When one of the seven types of roles of actors is described in the second term of a case such as "!hero", the event generation mechanism sets the instance of the actor, we call "agent" in the system, based on the predefined or inputted correspondence knowledge such

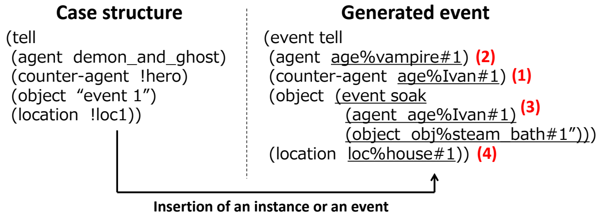


Fig.3. The correspondence between a case structure and the actgually generated event

as "Ivan: hero", "Baba-Yaga: donar", etc. For incorporating it into the integrated narrative generation system, when an agent as a Propp-based agent including hero, heroin etc. appears in a story generation process, the system describes the label such as "hero" in the attribute frame. An attribute frame is a data set associated with a noun concept and the noun concept is embodied by inserting instances or the other various labels into it.

(2) When the value of a case in an event is a name of a noun concept in which a specified instance of an agent, an object or a location should be inserted, the event generation mechanism sets an instance made from a noun concept in the noun conceptual dictionary selected within the range of the constraint. The name of the noun concept specifies the constraint in the current design. For example, if the name is "demon_and_ghost", the values to be inserted is selected in the range under the name in the noun conceptual dictionary. On the other hand, when the name is described in the style such as "horse@beast", it will be used as the value directly.

(3) When an event concept itself should be the value of a case (object case in particular) in an event recursively, the event generation mechanism expands the description such as "event 1" to a case structure different from the case structure in which the expanded event is included. The previous mechanism has used predefined sets of case structures to be inserted in the position based on examples described by Propp's theory. The examples have not three or more nests. The integrated narrative generation system uses a recursive function for dealing with this type of processing based on a predefined list of nested events. To achieve more general recursive processing will be a future issue.

(4) When one or more events have a mutual-dependence relation, the mechanism processes a special operation. In particular, when the value of the agent, object or location in a case structure must be equal to the value of the other case structure for the consistency among events, the mechanism uses a same value for the one or more elements and designates the range which the values can take in the noun conceptual dictionary. This is related to the "pairs of sub-functions". The integrated narrative generation system prepares a procedure for treat this processing with reference to the story knowledge base in section II. Each of many data stored in the story knowledge base is a set with two or more events and the elements in the events need to have a mutual-dependence relation. One of the issues is that the number in the Propp-based mechanism is not many, while the number in the story knowledge base will be very many.

V. CONCLUSION

This paper described both of the overviews of our Propp-based story generation system and an integrated narrative generation system, and presented ways for incorporating the Propp-based system in the level of event generation mechanism into the integrated system towards the complete blending. The basic policy was that we incorporate the special mechanisms in the Propp-based system into more generalized mechanisms in the integrated system, however a part of the mechanisms such as actors' roles in the Propp-based system will be considered to be arranged in the integrated system too.

REFERENCES

[1] Ogata T and Kanai A (2010), An Introduction to Informatics of Narratology (in Japanese). Gakubunsha.
[2] Akimoto T and Ogata T (2012), Macro Structure and Basic Methods in the Integrated Narrative Generation System by Introducing Narratological Knowledge. Proc. of the 11th IEEE International Conference on Cognitive Informatics & Cognitive Computing: 253-262.
[3] Akimoto T and Ogata T (2012), A Narratological Approach for Narrative Discourse: Implementation and Evaluation of the System based on Genette and Jauss. Proc. of the 34th Annual Confer-ence of the Cognitive Science Society: 1272-1277.
[4] Propp V (1968), Morphology of the Folktale. University of Texas Press.
[5] Imabuchi S and Ogata T (2012), A Story Generation System based on Propp Theory: As a Mechanism in an Integrated Narrative Generation System. Lecture Notes in Artificial Intelligence, vol.7614: 312-321.
[6] Imabuchi S and Ogata T (2013), Methods for Generalizing the Propp-based Story Generation Mechanism. Lecture Notes in Computer Science/Lecture Notes in Information Systems and Applications, incl. Internet/Web, and HCI, vol.8210: 333-344.
[7] Gervás P (2013), Propp's Morphology of the Folk Tale as a Grammar for Generation. Workshop on Computational Models of Narrative, a satellite workshop of CogSci 2013: The 35th meeting of the Cognitive Science Society: 106-122.
[8] Oishi K, Kurisawa Y, Kamada M, Fukuda I, Akimoto T and Ogata T (2012), Building Conceptual Dictionary for Providing Common Knowledge in the Integrated Narrative Generation System. Proc. of the 34th Annual Meeting of the Cognitive Science Society: 2126-2131.

Relationship input object position and optimal error diffusion coefficients for Kinoform using error diffusion method

Daisuke Kashima. Dept. of Electronics and Bioinformatics, Meiji Univ.
E-mail: ce11021@meiji.ac.jp
Ken-ichi Tanaka. Dept. of Network Design, Meiji Univ.
E-mail: tanaken@meiji.ac.jp

Abstract: In order to improve quality of reconstructed image from Kinoform, we investigated about separation object and noise.

It's considered that taking error diffusion method is good for separation object and noise from Kinoform. And it's considered that optimal error diffusion coefficients by Genetic Algorithm (GA) exist. Now, first we searched relationship input object position and optimal error diffusion coefficients. Second we derive relational expression with multiple linear regression analysis. Third we perform estimating optimal error diffusion coefficients for various input object position by derived relational expression. Therefore we report about efficacy of deriving relational expression of optimal error diffusion coefficients and input object position for Kinoform using error diffusion method.

Keywords: Kinoform, Computer-Generated-Hologram (CGH), error diffusion method, GA, multiple linear regression analysis.

I. INTRODUCTION

CGH is made in order to reconstruct three-dimensional image of object or virtual abject having difficult in irradiating a laser directly. It's known that the reconstructed image from Kinoform is brighter than the reconstructed image from CGH [1-3]. In order to practical use Kinoform, it's necessary to develop simple calculation method that high-quality reconstructed image is obtained.

In process of making Kinoform, factor to be negative impact most in reconstructed image is quantization of wave surface calculated for Hologram. In order to mitigating this impact, Kinoform was performed to compose Kinoform with repetition algorithm [4] or combination optimization problem [5]. However, the composition took a long time. In contrast, error diffusion method can separate object and noise from reconstructed image by one quantization. Therefore, error diffusion method is regarded as a good method to obtain high-quality reconstructed image by low cost processing.

In error diffusion method, it's considered that selection of error diffusion coefficients can be optimized by searching such as GA [6]. However, GA has a problem to take a long time for processing.

In this paper, we derived relational expression by searching relationship input object position and optimal error diffusion coefficients with multiple linear regression analysis. This analysis used data that is calculated optimal error diffusion coefficients by GA. We report efficacy of deriving relational expression of optimal error diffusion coefficients and input object position for Kinoform using error diffusion method.

II. ERROR DIFFUSION METHOD FOR KINOFORM

In this paper, we perform error diffusion method expanded using complex to quantize complex value into complex value.

(1) Scaling complex value

Before quantization of complex value, scaling is performed as follows.

$$F_1(u,v) = \frac{F(u,v)}{\max|F(u,v)|} \tag{1}$$

(2) Quantization of complex value

Complex value is quantized as follows.

$$F_2(u,v) = \exp\{j\arg\{F_1'(u,v) - F_1(u,v)\}\} \tag{2}$$

Quantization is performed by raster scanning.

(3) Calculation and diffusion of quantization error

Quantization error as $s(u,v)$ is generated.

$$s(u,v) = F_2(u,v) - \left(F_1'(u,v) - F_1(u,v)\right) \tag{3}$$

This $s(u,v)$ is diffused to adjacent and not quantized pixels with weight. The $s(u,v)$ is diffused to 4 directions such as Fig.1 Then, weight a, b, c, d are regarded as complex and relationship as follows.

$$|a| + |b| + |c| + |d| = 1 \tag{4}$$

We make a direction of "a" of Fig.1 agree with a direction of scanning.

The $F_1'(u,v)$ is quantized by diffusion of quantization error as follows.

$$F_1'(u+1,v) = F_1'(u+1,v) + a \cdot s(u,v) \quad (5)$$
$$F_1'(u+1,v+1) = F_1'(u+1,v+1) + b \cdot s(u,v) \quad (6)$$
$$F_1'(u,v+1) = F_1'(u,v+1) + c \cdot s(u,v) \quad (7)$$
$$F_1'(u-1,v+1) = F_1'(u-1,v+1) + d \cdot s(u,v) \quad (8)$$

In these 4 expressions, they are forms by the procedure language.

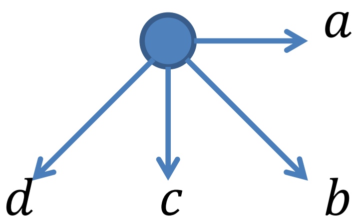


Fig.1 Diffusion coefficients

III. DERIVING RELATIONAL EXPRESSION

We derive relational expressions of input object position (x, y) and diffusion coefficients a, b, c, d. It's considered that relational expressions are composed of n-ordered approximation as follows.

$$a = \sum_{k=0}^{n}\sum_{r=0}^{k} a_{kr} x^{k-r} y^{r} \quad (9)$$
$$b = \sum_{k=0}^{n}\sum_{r=0}^{k} b_{kr} x^{k-r} y^{r} \quad (10)$$
$$c = \sum_{k=0}^{n}\sum_{r=0}^{k} c_{kr} x^{k-r} y^{r} \quad (11)$$
$$d = \sum_{k=0}^{n}\sum_{r=0}^{k} d_{kr} x^{k-r} y^{r} \quad (12)$$

However, input object central position x and y take values from -1/2 to 1/2. This is to retain generality when a number of pixels of Kinoform change. On a $N-\text{pixel/side}$ Kinoform, if input object central position by pixels is regarded as (i, j), x and y is derived as follows.

$$x = \frac{i - N/2}{N} \quad (13)$$
$$y = \frac{j - N/2}{N} \quad (14)$$

Now, we derive coefficients of relational expressions with multiple linear regression analysis. For that, response variable and predictor variable is defined from Eq. (9)-(12) as follows.

Response variable: Optimal error diffusion coefficients which was calculated beforehand by GA.

Predictor variable: Variables with combination from input object position(x, y).

IV. SIMULATION RESULT

In this paper, we simulated reconstructing image from Kinoform and evaluating quality of reconstructing image. Now, quality is regarded as better if evaluation value is lower. Then, we set following:

Kinoform length: $N = 512$.
Input object form: A white square.
Input object length: 96.
Input object central position: (i, j)
Evaluation form: Square
Evaluation length: 128
Evaluation central position: (i, j)
Approximate order: n
Number of samples: p

I. Comparison by approximate order

We simulated reconstructing image with relational expression from calculated optimal error diffusion coefficients. Fig.2 (a)-(g) are shown as results for GA and changing n from 1 to 7.

As shown in Fig.2, the reconstructed image using the greater approximate order is nearly equal to the reconstructed image using GA. And as shown in Table1, quality of reconstructed image using the greater approximate order tends to improve.

II Comparison by number of samples

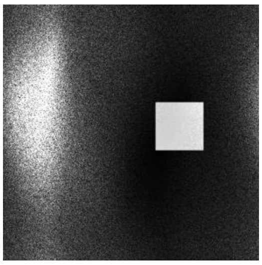
We simulated reconstructing image with relational expression from 49 or 98 of samples as p. Samples are used to derive relational expression. Fig.3 (a),(b) are shown as results of reconstructed image.

As shown in Fig.3, these don't have many changes. However, as shown in Table2, p has better 49 than 98. Large number of samples isn't better at all. It's considered that quality of reconstructed image is better when relational expression is used for considered input object position.

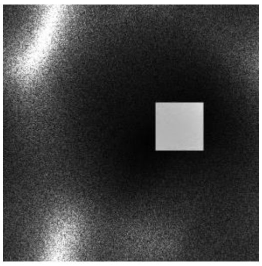
III Estimating for various input position

We simulated reconstructing image for various input object position with relational expression from 49 or 98 of samples as p and 7 of approximate order as n. Fig.4 (a)-(c) are shown as results of reconstructed image.

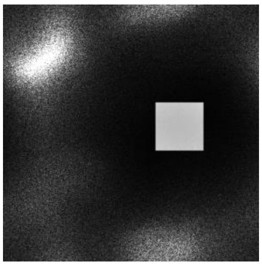
As shown in Fig.4, (b) is shown not to separate input object and noise very well. On the other hand, (c) is shown to separate input object and noise a little. And position of noise of (c) is similar to (a) than (b).



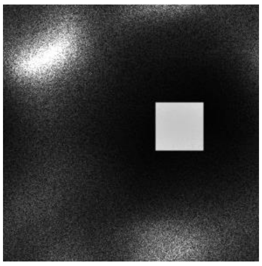
(a) n=1



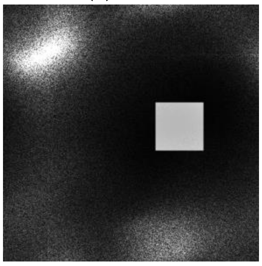
(b) n=2



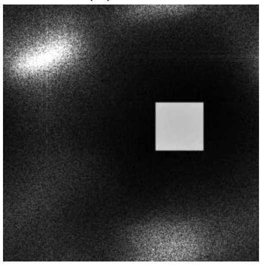
(c) n=3



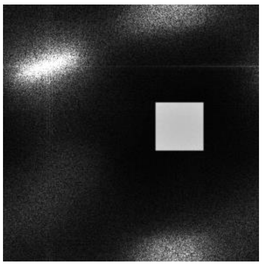
(d) n=4



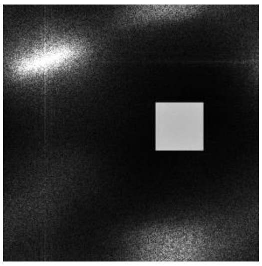
(e) n=5



(f) n=6



(g) n=7

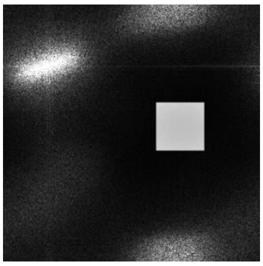


(h) GA

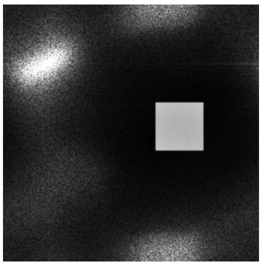
Fig.2 Comparison of reconstructed image by changing approximate order (i = 353, j = 243)

Table1 Comparison of evaluation value by changing approximate order (i = 353, j = 243)

Deriving method	Evaluation value
n=1	0.005529698
n=2	0.004612562
n=3	0.001690110
n=4	0.001600949
n=5	0.001808101
n=6	0.001615804
n=7	0.001576671
GA	0.001471379



(a) p=49

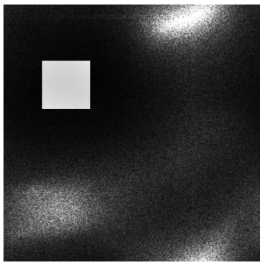


(b) p=98

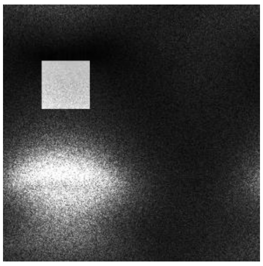
Fig.3 Comparison of reconstructed image by changing number of samples (i = 353, j = 243)

Table2 Comparison of evaluation value by changing number of samples (i = 353, j = 243)

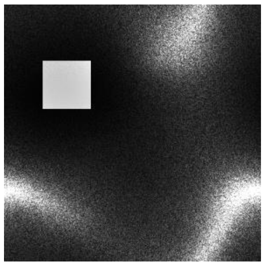
Number of samples	Evaluation value
p=49	0.001576671
p=98	0.001767945



(a) GA



(b) p=49



(c) p=98

Fig.4 Comparison of reconstructed image by using estimated optimal error diffusion coefficients (i = 125, j = 160)

Table2 Comparison of evaluation value by using estimated optimal error diffusion coefficients (i = 125, j = 160)

Deriving method	Evaluation value
GA	0.001573104
p=49	0.033828547
p=98	0.002430123

V. CONCLUSION

In this paper, we performed to improve reconstructing image by deriving relational expression of input object position and optimal error diffusion coefficients from Kinoform. As a result, we suggest follows:

1. Better quality of reconstructed image is obtained by relational expression with the greater approximate order.
2. It's suitable for estimating by relational expression with larger number of samples.

In order to improve quality of reconstructed image from Kinoform, these have obviousness about efficacy of deriving relational expression of input object position and optimal error diffusion coefficients.

In this paper, only one input object is used. So we have a future task for various input object to confirm efficacy.

REFERENCES

[1] S. Weissbach, F. Wyrowski and O. Bryngdahl :"Digital phase holograms: coding and quantization with an error diffusion concept", Opt. Comunn., 72, pp.37-41 (1989)

[2] L. B. Lesen. P. M. Hirxch and Jr J. A. Jordan. :"The Kinoform: A new wavefront binary phase only filters", Appl. Opt. 27, 18, pp.3785-3790 (1988)

[3] F. Mok, H. K. Disp Liu and D. Psaltis :"Real-time computer generated hologram by means of liquid crystal television", Opt. Lett., 11,11,pp748-750 (1986)

[4] Hiroshi Akahori :"Detection of the Complex Amplitude Distribution of a Kinoform Using an Iterative Fourier-Transform Algorithm", Journal of Trans., IEICE.,J78-C-1, 9, pp.400-408 (1995)

[5] Ken-ichi Tanaka and Yuta Shimizu :"Search of Optimal Error Diffusion Coefficient for Kinoform", Journal of IIEEJ., Vol.53, No.3, pp.1-6 (1999)

Optimization of dither matrix by hybrid of Genetic Algorithm and Simulated Annealing (P)

Kohei Kato
Department of Electronics and Bioinformatics, Meiji University
ce21019@meiji.ac.jp
Ken-ichi Tanaka
Department of Network Design, Meiji University
tanaken@isc.meiji.ac.jp

Abstract: Halftoning is used by ink-jet printer, color copier and so on. Dither method is the method of halftoning. Bayer method that is one of the dither methods is able to obtain a superior evaluation value compare with other dither methods, though it is the decided matrix. Several investigations of Optimization of the dither matrix are reported. However, there is no investigation from point of view of the combinational optimization problems. In this paper, we search the optimum dither matrix by Genetic Algorithm (GA) and Simulated Annealing (SA) and hybrid of GA and SA(GA/SA).We weigh conventional method against these methods.

Keywords: Dither matrix, Genetic Algorithm, Simulated annealing, Hybrid**.**

I. INTRODUCTION

Halftoning is used by ink-jet printer, color copier and so on. Dither method is the method of halftoning. Bayer method that is one of the dither methods is able to obtain a superior evaluation value compare with other dither methods, though it is the decided matrix. Several investigations of Optimization of the dither matrix are reported. However, there is no investigation from point of view of the combinational optimization problems. In this paper, we search the optimum dither matrix by Genetic Algorithm (GA) and Simulated Annealing (SA). GA is superior to search for the state of global search. SA is superior to search for the state of local search. So we introduce hybrid of GA and SA (GA/SA) for halftoning. By GA/SA, the optimum dither matrix is obtained without converging in a local solution. In this paper, we research dither matrix (4−4) and (8-8).

II. Genetic Algorithm（GA）

GA is the method of combinatorial optimization. GA is the method of global search by individual group of solution candidate. Each individual is given fitness that calculated by cost function. The individual which has high fitness is selected. The individual which does not have high fitness is went out. The selected individual is performed crossover and mutation. This process is continued by decided generation. In the end optimum solution is the individual which has highest fitness.

In this paper, the individual are dither matrix（4-4）and（8-8）. Elements of individual are threshold level of gray level from 0 to 255. The threshold level is as below.

I[x][y]=(2*i+1)*8, I[x][y]=(2*i+1)*2

I: element of the matrix

i: numerical number from 0 to 16 or from 0 to 64.

The same number is not set in the same individual.

In this study, the algorithm is as below

[step1] Initialization sets up the parameter such as number of individual the number of generation, crossover rate and mutation rate.

[step2] Initial individual is created by the number set up. Initial individual of element is putted in random

[step3] Evaluation value of fitness to all individual is calculated.

[step4] Selection: Selection is to decide the individuals of the next generation by the method of Roulette Tournament.

[step5] Crossover: Crossover is performed if uniform random number is lower than crossover rate

[step6] Mutation: Mutation is performed if uniform random number is lower than mutation rate. Two elements of the individual which is selected at random are exchanged.

[step7] It is repeat from step[3] to step[6] until the number of generation.

[step8] In the final generation, the individual which has the highest fitness is the optimal solution. and output the evaluation value ,and dither matrix.

Ⅱ.Ⅰ Cost function

Cost function E is the addition of cost E_m and cost E_c. Cost E_m is evaluated value of gray level. Cost E_c is Evaluated value of contrast. Equation of cost function E is as below.

$$E = 0.2 \times Em + 0.8 \times Ec$$

Fitness is given by

$$f = 1/E$$

Ⅱ.Ⅱ selection

In this paper, the method of Roulette tournament selection is used. The method of Roulette selection is to select the individual in probability P. P is defined by

$$P = \frac{f(I_i)}{\sum_{j=1}^{N} f(f_i)}$$

The method of Tournament selection is as below.

1) select individual at random as many as the number of tournament size
2) select the individual that has the highest fitness.

In this paper, tournament size is 2.Roulette tournament selection is the method of combining the two methods.

Ⅱ.Ⅲ Crossover

Perform crossover for the two individual which is selected at random. In this paper, cycle crossover is used.

Cycle crossover: Select position X at random. The element of individual A which of position is X is just inherited. Element of individual A same as the element of individual B is inherited. This process repeat until the position which has already been selected is selected.

The elements that are not selected is inherited in the order of appearance of elements of individual B.

Ⅱ.Ⅳ Mutation

The elements of the individual which is selected at random are exchanged in order to converge in a local solution.

III. Simulated Annealing（SA）

This section, we describe SA. The algorithm of SA is below

[step1] initialization sets up parameter such as an initial temperature T_0, and temperature T_{min}, temperature reduction rate α and initial dither matrix which is decided at random.

[step2] A new matrix is generated how to exchange two elements of the dither matrix.

[step3] solution of cost function E to all individual is calculated.

[step4] According to acceptance probability P ,whether a new matrix is accepted is decided. If $\Delta E \le 0$ is less than 0, a new matrix is accepted. If $\Delta E \ge 0$ is higher than 0, P is given as below

$$P = \frac{\Delta E}{T}$$

If P is less than uniform random number, a new matrix is accepted.

[step5] repeat from[2] to [4] enough and decrease temperature. T is given by

[step6] until $T \le T_{min}$,repeat from step[2] to step[5].

Ⅲ.Ⅰ Set up parameters

Parameters are as below

T_0=0.15904, $T_{min} = 4 * 10^{-6}$ (4-4)

T_0=0.114719, $T_{min} = 4 * 10^{-6}$ (8-8)

This value are given by

$$T_0 = \Delta f_{min} + (\Delta f_{max} - f_{min})/10$$

f_{min} and f_{max} are determined by multiple experiment preliminarily. In this preliminary experiment, the temperature is divided temperature 1 degree into 32 parts from 1000 degree. Maintaining the constant temperature, SA is performed. the initial temperature is obtained by the result. A temperature reduction rate α is 0.99.The termination condition is that the temperature is lower than T_{min}, or a new dither matrix is hardly accepted.

IV. GA/SA

The problem of GA and SA is as below

GA: non-good at local search. Around the optimal solution, calculation cost for the improvement of the solution rise so much.

SA: to set up initial temperature is very difficult. If initial temperature is low, the optimum solution converges in a local solution. If initial temperature is high, it takes a long time.

In this study, at first, the dither matrix is searched by GA. In the next, the dither matrix that gotten by GA is set to initial dither matrix of SA. The optimum combination is searched by SA using the matrix. In this way, the problem is expected to improve by hybrid of GA/SA. In GA process, if a solution becomes the equilibrium, break off operation of GA. The problem of GA is improved by this way. In SA process, the initial temperature can be set low temperature, because it is begun with optimum solution neighborhood. The optimum solution is obtained without converging in a local solution by using hybrid.

Table 1 .the result by GA

4*4							
crossover rate		0.7	0.75	0.8	0.85	0.9	0.95
mutation rate	0.2	101.721	101.7632	101.8071	101.6875	101.7337	101.7603
	0.15	101.7392	101.8115	101.7428	101.681	101.7034	101.6948
	0.1	101.7433	101.7301	101.7555	101.7229	101.7392	101.7136
	0.05	101.7775	101.7708	101.7575	101.7548	101.7326	101.7898
	0.01	101.8276	101.7686	101.8071	101.8321	101.7387	101.776
8*8							
crossover rate		0.7	0.75	0.8	0.85	0.9	0.95
mutation rate	0.2	101.7457	101.6455	101.6949	101.6934	101.6934	101.7245
	0.15	101.7648	101.7865	101.7078	101.797	101.797	101.7226
	0.1	101.7624	101.7496	101.7684	101.6936	101.6936	101.6878
	0.05	101.7262	101.7087	101.7228	101.7403	101.7403	101.7577
	0.01	101.9554	102.2744	101.7651	101.9025	101.9025	102.128

V. Simulation

Ⅴ. Ⅰ Conventional method

In this section, the result that is conventional method, Bayer method, is shown. Fig.1 shows original image. Fig.2,3 shows Bayer matrix and the image that is obtained by Bayer method (4-4) and (8-8). Evaluation value E is 101.8021(4-4) and 101.9177(8-8)

Ⅴ. Ⅱ GA

Table.1. shows the results obtained by GA.

The results are the average obtained by simulating several times. According to parameters, the results is greatly varies.

Ⅴ. Ⅲ SA

Table.2 shows the results obtained by SA. In (4-4), the result is same as by GA. In (8-8), the result obtained by SA is better value than GA. The result is improved by increasing temperature reduction and the number of annealing. In this paper, the optimum solution is obtained in parameter that α is 0.99 and the number of annealing is 2000.

Ⅴ. Ⅳ GA/SA

We set the parameters of GA that the best value is obtained by only GA. The initial temperature sets it that acceptance probability of a minimum change for the worse solution becomes 90%. The temperature reduction α is 0.9 and the number of annealing is 3000. These parameters are obtained from a result of various simulation. Table.3 shows the result. Optimum solution is obtained even if α is lowed until 0.9.

Ⅴ. Ⅴ Comparison of 3 methods

Table.4 shows the result that is obtained by Bayer, GA, SA and GA/SA. The result by GA/SA is better evaluation value than bayer method, GA and SA （4-4） and (8-8). Fig.4,5 show the image and dither matrix obtained by GA/SA. By GA/SA, the processing time is much earlier because beginning with optimum solution neighborhood is able to set low parameters, T and α. This is because the problem of SA and GA are able to be improved.

VI. CONCLUSION

The result of GA, SA and GA/SA is better evaluation than Bayer method. The result obtained by GA/SA is the best evaluation value. By using GA/SA, the optimum solution is obtained in the parameters that are not able to obtain it by only GA. So it is not necessary to set up detailed parameters. The processing time by using GA/SA is much earlier than GA and SA. These show the effectiveness of GA/SA

Table 2. The result by SA

4*4			
T_0	α	an	E
0.15904	0.9	2000	101.681
0.15904	0.99	2000	101.681
8*8			
T_0	α	an	E
0.114719	0.9	2000	101.6352
0.114719	0.99	2000	101.5754

Table 3. The result by GA/SA

crossover rate	mutation rate	T_0	α	an	E
4*4					
0.85	0.15	0.026775	0.9	2000	101.6585
		0.026775	0.99	2000	101.6585
0.85	0.1	0.007508	0.9	2000	101.6585
		0.007508	0.99	2000	101.6585
8*8					
0.85	0.15	0.026775	0.9	3000	101.5678
		0.026775	0.99	2000	101.5781
0.85	0.1	0.007508	0.9	3000	101.6071
		0.007508	0.99	2000	101.5881

Table 4. Compares the methods

	bayer	GA	SA	GA/SA
E(4*4)	101.8021	101.681	101.681	101.6585
E(8*8)	101.9177	101.6552	101.5754	101.5678
CPUTIM(4*4)	31	476755	4424220	1498078
CPUTIM(8*8)	37	1563145	16473800	2892149

Fig. 1 Original image

Fig. 2 Bayer method(4-4)

Fig. 3 Bayer method(8-8)

Fig. 4GA/SA(4-4)

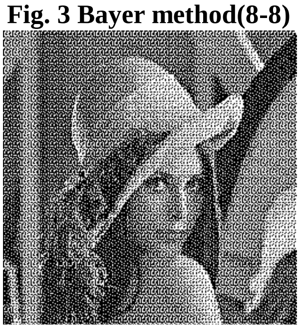


Fig. 5GA/SA(8-8)

Table 5 Optimum dither matrix (4-4)

7	12	4	13
11	0	8	2
5	14	6	15
9	3	10	1

Table 6 Optimum dither matrix (8-8)

56	36	9	43	35	50	10	30
40	17	24	53	12	19	37	2
54	6	57	31	4	61	46	22
63	27	41	47	23	58	29	7
38	11	18	3	34	16	39	13
51	48	32	62	49	8	60	33
26	1	42	15	25	44	20	5
14	52	21	55	0	28	59	45

REFERENCES

[1] Toshiharu Nagao"evaluation image processing " publication shoukoudou Inc.(2002).

[2] Ken-ichi Tanaka."Color Image Halftoning Using Simulated Annealing"publication the journal of the Institute of Image Information and Television Engineers. VOL.61,pp. 828-837(2007)1

[3] Yoshito Abe "A New Method of Designing a Dither Matrix"IEICE Trans.Fundamentals,VOL. E85-A ,NO.7, pp.1702-1709 (2002).

[4] Subtours in Complete Subtour Exchange Crossover and the Behavior Property"Institute of Electronics,Information,and Comunication Engineers,Vol. J81-D-I,No.2,pp213--218(1998)

[5] Yutaka Shirai,Naofumi Matumoto"A Hybrid Method for Block Layout Design based on Metaheuristics" Japan Society of Mechanical Engineers,Vol. 63, No. 611, pp.2337--2344 (1997).

Inverse Halftoning using Multi-Layer Feed-Forward Neural Network (P)

Hiroki Hamashoji

Department of Electronics and Bioinformatics, Meiji University
ce21063@meiji.ac.jp
Ken-ichi Tanaka
Department of Network Design, Meiji University

Abstract: In this paper, we proposed novel inverse halftoning technique, which use Multi-Layer Feed-Forward Neural Network (MLFFNN) and Gaussian filtering. This method reconstruct continuous tone image from any halftoning methods. As a result, a high quality image can reconstructed from binary image compared with conventional method. Especially the proposed method image quality is good both visually and Peak Signal-to-Noise Ratio (PSNR) value, compared with reconstructs inverse halftone images by Look-up table (LUT) based inverse halftoning.

Keywords: inverse halftoning, neural network, image processing, halftoning

I. INTRODUCTION

Halftoning is the techniques that convert continuous tone image into a binary image. It is used in printing device and image display, which cannot display continuous tone image. Human visual system acts as low-pass filter, and halftoning binary image have high frequency information. Binary images are inconvenient for image processing such as the scaling, and image analysis. Therefore the inverse halftoning are required. Inverse halftoning is a technique that reconstructing continuous tone images from binary images.

As the conventional halftoning methods, various algorithms have been proposed ordered dither[1], the error diffusion[2], Direct Binary Search (DBS), Genetic Algorithm (GA)[3] and so on. Besides, various algorithms have been proposed to inverse halftoning. However some of these algorithms are only for halftone images produced by error diffusion method, because the error diffusion method is widely used. Above all deconvolution linear model of error diffusion halftoning (inverse error diffusion method) are known produces very good results.

Look-up table (LUT)[4] based inverse halftoning method can be processed regardless half-tone processing method. However LUT method is not better than inverse error diffusion method.

In this paper, we proposed novel inverse halftoning technique, which use Multi-Layer Feed-Forward Neural Network (MLFFNN) and Gaussian filtering. This method reconstruct continuous tone image from any halftoning methods.

We consider inverse halftoning as a nonlinear problem. MLFFNN has function approximation ability. Therefore we applied MLFFNN to inverse halftoning. We used a low-path filter to remove a noise, because an image provided by MLFFNN has some noise.

The proposed MLFFNN inverse halftoning method input 7x7 halftone image pixels to reconstruct center pixels. We use Gaussian filtering to the image processed by MLFFNN. We use back propagation algorithm for training MLFFNN. When the neural network learn, the neural network is made to learn the original continuous tone image as a teacher signal. Thus the neural network can be made to function as reconstruct continuous tone image from binary image.

More over this paper examine some parameters such as the number of hidden nodes.

II. MLFFNN INVERSE HALFTONING

1. CONFIGURATION OF NEURAL NETWORK

We regard inverse halftoning as reconstruction processing from the neighborhood of the target pixel of the halftoning image. As neural network has function approximation ability, neural network can estimate the suitable function for inverse halftoning. In order to output center pixel, 7x7 pixels are input into MLFFNN. MLFFNN consist an input layer, two hidden layer, and an output layer. This network is shown in Fig 1. Liner function is used to the output layer activation function, and sigmoid function is used to the hidden layer activation function.

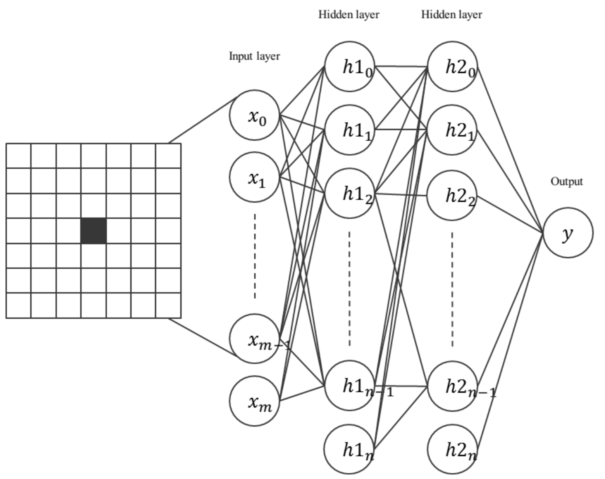


Fig.1. Structure of MLFFNN

Output y is expressed as follows:

$$y = \sum_{i=0}^{n} h2_i \times w3_i \qquad (1)$$

where.

$$h2_j = f(\sum_{i=0}^{n} h1_i \times w2_{ji}) \qquad (2)$$

$$h1_j = f(\sum_{i=0}^{m} x_i \times w1_{ji}) \qquad (3)$$

$$f(x) = \frac{1}{1+\exp(-x)} \qquad (4)$$

Where n is the number of hidden elements, x_m, $h1_n$, $h2_n$ is the bias, and $w1_{ji}$, $w2_{ji}$, $w3_{ji}$ are the weight between the elements.

2. LEARNING OF THE NEURAL NETWORK

MLFFNN trained by back propagation, and the square error between output signal and the teacher signal was minimized.

Neural network has a problem called the over-fitting. This is a problem that generalization ability became worse by the excessive fit to the training data. To avoid over-fitting, we use α function[5] for a learning rate. Furthermore, we mixed noises into a teacher signal.

Online training is used. The online training biased to training data of near the end. Therefore we took the average of the weights when the neural network learned all the training data. Furthermore, neural network is input the training data in random order in order to avoid bias in the learning. When the learning rate was under 0.1, we stopped the training. The training rules are written as:

$$\Delta w_t^{(i)} = -\eta^t \nabla f\left(w_t^{(i)}\right) \qquad (5)$$

$$w_t^{(i)} = w_t^{(i-1)} + \Delta w_t^{(i)} \qquad (6)$$

$$w_t = \frac{1}{N}\sum_{i=1}^{N} w_t^{(i)} \qquad (7)$$

where.

$$\eta^t = \frac{3t}{100}\exp(-\frac{3t}{100}) \qquad (8)$$

and teacher signal of the t epoch

$$tsignal^t = tsignal + 0.05 \times rand * \exp(-\frac{1}{120}) \qquad (9)$$

Where N is the number of training data. Weights were initialized with uniform random numbers of -0.2 to 0.2, and rand is uniform random number of -0.5 to 0.5. These parameters were decided by a preliminary experiment.

III. EXPERIMENTAL RESULT

1. EXAMINATION OF THE PARAMETER

We used 8bit, 768x512 images for training of the

Fig.2 The image that was used to training of the MLFFNN

MLFFNN. The image which used for learning is shown in Fig 2. These images are Kodak Lossless True Color Image Suite. We tried the reconstruction of the halftone image which processed by Floyd-Steinberg error diffusion method. Lenna and pepper image are used for evaluation.

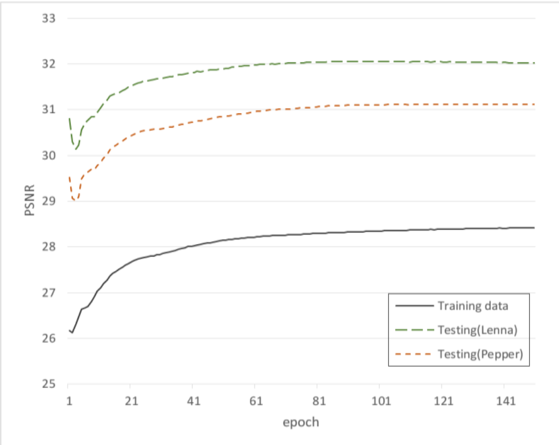


Fig.3. Variation of PSNR for the test image and the training image.

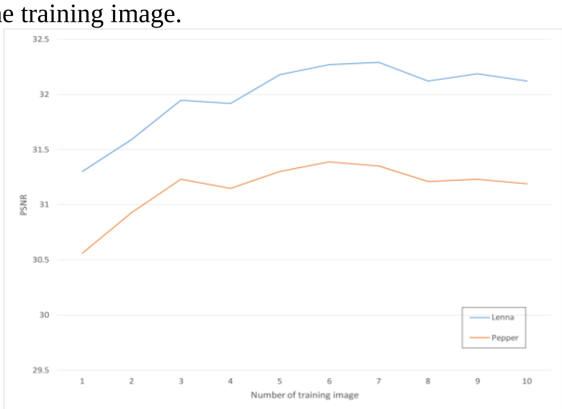


Fig.4. PSNR for the number of the training i mage.

Table 1. PSNR (dB) values for the number of hidd en layer elements.

Number of elements	Lenna	Pepper
5	30.97	29.98
10	31.53	30.63
15	31.79	30.84
20	31.93	31.02
25	32.00	31.07
30	32.05	31.16
35	32.10	31.14
40	32.11	31.17

To confirm over-fitting whether can be suppressed, the graph of PSNR for the test data and training data is shown in Fig 3. There is not the aggravation of the

Fig.4 Inverse halftone image processed by M LFFNN

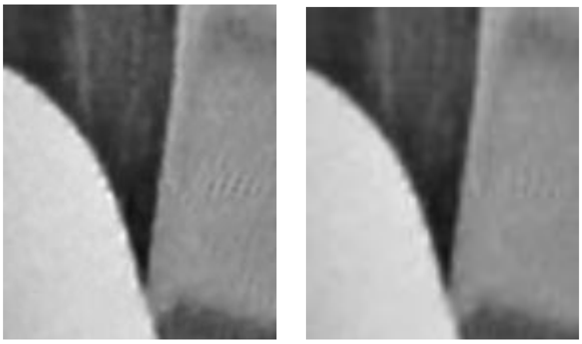


Fig.5 Enlarged inverse halftone image processe d by MLFFNN(left), MLFFNN + Gaussian(right)

solution for the test data, and the over-fitting is suppressed.

Table 1 is the PSNR for the number of hidden elements. From Table 1, a number of hidden layer elements improve PSNR. However, computational cost also increases. Thus it is necessary to set appropriately the number of hidden layer.

Fig 4 is the PSNR for the number of training data. Even less training data, MLFFNN have some performance. Besides when MLFFNN used a lot of training data, performance is improved as much as when a number of hidden layer elements used. When using more 7 images for the training data, performance is deteriorating. This can be improved because this arise from discontinued learning when the learning rate is cut to 0.1. A lot of training data increase the time for training. However the training process required only once, and the time processed inverse halftoning is not change.

From the above, elements of hidden layer set the 25 and 7 images are used to train the MLFFNN. Fig 4, 5 is inverese halftone image processed by proposed MLFFNN. This image has some noises in flat area. Gaussian filter is used to remove these noises. Standard deviations σ is set so that a square error minimized between training image and image processed by MLFFNN.

2. COMPARISON

The comparison between the proposed method and conventional method is showed in Table 2. The proposed method obtained better results than GA and LUT. Furthermore, equivalent results are obtained in comparison with corresponding to only the error diffusion method. From this result, it is thought that MLFFNN can estimate the suitable function for inverse halftoning.

Then, we tried to reconstruct the image that has been processed by the ordered dither method and GA. MLFFNN is trained by the image shown in Fig 2 and the image processed by each halftoning method. In the halftoning using GA, we use the MSE and the Gaussian filter to the evaluation. However the proposed method is becoming better results than the Gaussian filter. From table 3, the proposed method is good results in inverse halftoning for ordered dithering.

Table 2. PSNR (dB) values for restoration from Floyd-Steinberg error diffusion.

Inverse halftoning method	Lenna	Pepper
Proposed	**32.5**	**31.5**
Wavelet (Xiong et al.)[6]	31.7	31.0
POCS-Wavelet(Bozkurt and Ceting)[7]	32.2	30.9
Anisortropic LPA-ICI(Foi et al.)[8]	32.4	31.6
LUT(Mese and Vaidyanathan)[4]	31.0	
GA(Furuya and Mori)[9]	31.4	28.8
LMS-MMSE(Change et al.)[10]	31.4	31.2

Table 3. PSNR (dB) values for restoration from GA.

Inverse halftoning method	Lenna	Pepper
Proposed	**33.3**	**31.4**
Only Gaussian filter(σ=1.5)	32.1	31.2

Table 4. PSNR (dB) values for restoration from ordered dither image(4x4 bayer dither Matrix).

Inverse halftoning method	Lenna	Pepper
Proposed	**29.0**	**28.7**
GA(Furuya and Mori)[9]	26.6	25.9
LMS-MMSE(Change et al.)[10]	28.2	
Only Gaussian filter(σ=1.5)	28.2	27.7

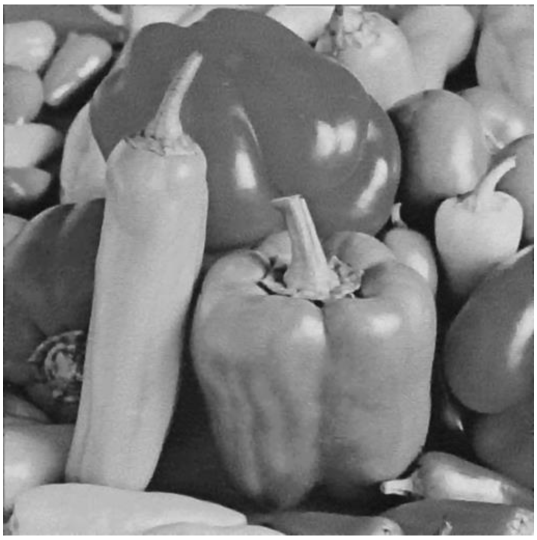


Fig.6 Floyd-Steinberg error diffusion image inverse halftoning by Proposed method. Lenna(upper, PSNR=32.54), Pepper(under, PSNR=31.50)

IV. CONCLUSION

In this paper the novel inverse halftoning method using combination of MLFFNN and a Gaussian filter is proposed.

Fig.7 Halftone image processed by GA(upper), and inverse halftone image processed by proposed method(under, PSNR=33.28).

Fig.8 Bayer ordered dither image inverse halftoning by proposed method(PSNR=28.71).

The proposed method, neural network calculates the suitable weights for the inverse halftoning from a pair of halftone image and the original image. Therefore, it is possible to perform inverse halftoning without considering the computational model of the halftoning. Thus proposed method can optimal inverse halftoning even the halftone image that result changed by an initial value like DBS method and GA. In addition, it was able to reconstruct image from processed by the error diffusion that was equivalent to the method using a linear model of the error diffusion.

Multiple parameters such as the initial weight width and learning rate with in the neural network, those parameters influence it mutually and affect a result. As future problem, there is the adjustment of parameters for a neural network further performance.

REFERENCES

[1]B. E. Bayer, An optimum method for two-level rendition of continuous-tone pictures. ICC Conf. Rec.26 pp11-15 (1973)

[2]R. W. Floyd and L. Steinberg, An Adaptive Algorithm for Spatial Greyscale. Proc Society for Information Display vol. 17-2:75-77 (1976)

[3] E Myodo, H Aguirre and Kiyoshi Tanaka, Improved Image Halftoning Technique Using GAs with Concurrent Inter-block Evaluation. GECCO 2003 Lecture Notes in Computer Science Vol 2724, pp2264-2276 (2003)

[4]M Mese and P.P. Vaidyanathan, Look Up Table(LUT) Method for Inverse Halftoning. IEEE Trans. Signal Processing 10-10:1566-1578 (2001)

[5]T Takahashi and M Adachi, Nonlinear Scheduling of the Learning Rate in an Actor-Critic Reinforcement Learning (in Japanese). Technical Report of IEICE 13-18 (2004)

[6]Z Xiong, M.T. Orchard, and K. Ramchandran(1999), Inverse halftoning using wavelets. IEEE Trans. Signal Processing 8:1479-1482

[7]Bozkurt Unal G and A.E. Çetin, Restoration of ErrorDiffused images using Projection onto Convex Sets. IEEE Trans. Image Processing (2001)

[8]A. Foi et al, Inverse Halftoning Based on The Anisotropic LPA-ICI Deconvolution, TICSP Workshop on SMMSP2004 (2004)

[9]T Furuya and K Mori, Restoration of Halftone Image Using Genetic Algorithms (in Japanese). T.SICE 45-7:361-368 (2009)

[10]P Chang, C. Yu and T Lee, Hybrid LMS-MMSE Inverse Halftoning Technique
IEEE Transaction on image processing VOL10-1(2001)

Embedding used binary number conversion into a Color image (P)

Hiroaki Oguma
Department of Electronics and Bioinformatics, Meiji University
1-1-1, Higashi-Mita, Tama, Kawasaki, Japan
ce31016@meiji.ac.jp
Ken-ichi Tanaka
Department of Network Design, Meiji University
4-21-1, Nakano, Nakano-ku, Tokyo, Japan
tanaken@isc.meiji.ac.jp

Abstract: Digital watermarking is widely used as security of a binary format image. The purpose of this paper is extension to a color image from grayscale image. Then, it thought as important giving many embedding information to a color image. A color image is divided into RGB, the concentration value of a pixel is made into a binary number, and last 3 digits are changed to arbitrary numbers. It embeds by carrying out by repeating this, and secure many information. 10x19660 times of embedding are able to be performed to the picture of 256x256 pixels. Moreover, the signal to noise ratio is about 40dB. Therefore, it seems that there is little influence which it has on a picture.
Keywords: Digital watermarking, Binary number, RGB,

I. Introduction

In recent years, in addition to rapid osmosis of the Internet, it is becoming the environment where everyone can circulate easily the digital information which led the network by the spread of personal digital assistants.

In connection with this, practical use we which gathers information according to its purpose is performed out of the information which various people built.

On the other hand, as for the digital contents which have value as social wealth, the kind of media also spreads from numerical text to a still picture, a sound, and video by progress of an information technology, and a sale of the movie and musical piece using a network also came to be made.

However, since it has the feature that a duplicate is easy and does not deteriorate, digital information is, when circulating worthy information, the structure which prevents unjust use and protects intellectual property rights is indispensable [1].

After generating of the illegal use of a work, detection, discernment, and proof are enabled and digital-watermarking technology is mentioned as a representative case of the technology in which this expects deterrence of an illegal use indirectly [2].

Digital watermarking is technology which embeds watermark information, as not perceived by man at digital contents, such as a picture, an image, and a sound.

Even if ID information and copyright information are embedded at the contents themselves if it is digital watermarking, and partial cutoff etc. are moreover edited, it is also possible to take out the embedded information.

As for a binary format image, many security methods exist focusing on a document picture. Although how to change between characters, and the method of putting a design into a background are famous, they also have many problems, there being little amount of information of a watermark, or the use range being restrictive [3] [4].

The previous techniques is performed the digital-watermarking method for a color image, therefore still now [5]. In this paper, the technique of embedding with emphasis on giving many information to a color image is considered.

II. Experimental

A color image is divided into a RGB ingredient and the concentration value of a pixel is acquired. The acquired concentration value is changed into a binary number, and the beforehand fixed embedding information bit is put into last 3 digits portion one after another. Many embedding information is secured by performing the above processing to all the pixels of RGB each color.

The bit value sequence of 0 and 1 used as the contents embedded as a sample and picture figure 1 of 256x256 were prepared. Here, the bit value sequence of 10 digits of (0,1,0,1,0,1,0,1,0,1) was used.

Fig.1 Original picture image

Fig.2 After embedding

Table.1 MSE and the signal to noise ratio in embedding picture

Name	Embed	Extranction	Chenge	MSE	S/N[dB]
Parrots	19660	19660	65360	7.784	39.2188
Girl	19660	19660	65434	7.437	39.4168
Pepper	19660	19660	65432	7.574	39.3376
Earth	19660	19660	65415	7.416	39.4291
Aerial	19660	19660	65412	7.559	39.3462
Average	19660	19660	65410.6	7.554	39.3497

It separates into each color of RGB, Fig. 1 is shown, and a concentration value is acquired per pixel from the upper left to the lower right.

The acquired numerical value is changed into a binary number, and it replaces in the bit value sequence which was preparing the last 3 digits. The numerical value after replacing is re-changed into a decimal number, and it stores in the pixel of a basis, and moves to the following pixel. In the next replacement, the 4th to 6th three of the prepared bit value sequence is replaced.

This processing is performed one after another and embedding is performed to all the pixels. After embedding is completed to the last, the separated RGB ingredient is compounded and it returns to a color image. The picture after embedding is shown in Fig. 2. The signal to noise ratio of Fig. 2 was investigated.

About Fig. 2, it separates into each color of RGB and reads, after acquiring a concentration value, binary number conversion is carried out, and lower 3 figure is recorded. It is checked whether it is in agreement with the bit value sequence which connected and embedded the bit value sequence which is read one by one and recorded.

Moreover, in the pictures I and K of picture size m×n, the maximum pixel value which a picture can take was defined as MAX, and the mean squared error (MSE) and the peak signal-to-noise ratio (S/N) were calculated from the formula (1) and (2).

$$\mathrm{MSE} = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2 \qquad (1)$$

$$\frac{\mathrm{S}}{\mathrm{N}} = 10\log_{10}\frac{MAX^2}{MSE}\ [\mathrm{dB}] \qquad (2)$$

The calculation result to some pictures is shown in Table 1.

III. CONCLUSION

It succeeded in repeating the bit value sequence of 10 digits of (0,1,0,1,0,1,0,1,0,1) 19660 times, and embedding it to the picture of 256x256. Moreover, extraction was also successful same. Since the signal to noise ratio of the embedded picture was about 40 dB, it can be said to be that there are little change and sense of incongruity before and behind embedding. In the viewpoint of the quantity of the embedded bit value sequence, the technique of this paper has secured sufficient amount of information to use for digital watermarking.

REFERENCES

[1] Norihiko Sakurai, Yutaka Kidawara, Yoichi Takashima, Noburou Taniguchi, and Koji Namba, A Survey on Copyright Protection Technologies for Network Distribution of Digital Information, Information Processing Society of Japan, Vol.42, No. SIG 15(TOD 12), p-p63-76, (2001)

[2]Ryoichi Sasaki, Hiroshi Yoshida, Consideration on Copyright and Illegal Copy Countermeasures under IT Revolution, Information Processing Society of Japan, CSEC, Vol.13, No.7, p-p37-42, (2001).

[3] Cox IJ, Kilian J, Leighton FT, and Shamoon T, Secure spread spectrum watermarking for multimedia, IEEE Transactions on Image Processing, Vol.6, No.12, p-p1673-1687, (1997).

[4] Jack T.Brassil, Steven Low, Nicholas F.Maxemchuk, Lawrence O'Gorman, Electronic marking and identification techniques to discourage document copying, IEEE Journal on Selected Areas in Communications, Vol.13, No.8, p-p1495-1504, (1995).

[5] Go Tanaka, Noriaki Suetake, and Eiji Uchino, High Capacity Data Embedding Method for Palette-Based Images, Technical Report of the Institute of Electronics, Information and Communication Engineers, SIS, a smart information media system, Vol.108, No.461, p-p13-16, (2009).

[6] Hiroyuki Kitayama, C++Builder6 component practical use guide & practice programming Vol.6, a cut system, pp203-352, (2004).

[7] Yasushi Yamanaka, Hiroshi Fujii, Digital Image Scrambling Method for Information Distribution, Information Processing Society of Japan, Vol.38, No.10, p-p1945-1955, (1997).

[8] Yasushi Abe, Koichi Inoue, Digital Watermarking for Bi-Level image, Ricoh Technical Report, No.26, pp8-16, (2000).

[9] Naohisa Komatsu, Ken-ichi Tanaka, digital-watermarking technology : the security of digital contents, Tokyo Denki University Press, (2004).

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Three Dimensional Images Reconstruction on Computer Generated Holograms the Multiple Regression Analysis (Poster)

Kenta Ayabe
Dept. of Electronics and Bioinformatics Meiji University
4-21-1, Nakano, Nakano-ku, Tokyo,Japan
E-mail; ce21002@meiji.ac.jp
Ken-ichi Tanaka
Dept. of Network Design Meiji University
E-mail;tanaken@meiji.ac.jp

Abstract: **In this paper, for a Computer Generated Hologram (CGH) reconstructing a three-Dimensional (3D) object, we present the improvement of the reconstruction image, reducing of the calculation ti me using error diffusion method and a multiple regression analysis. The multiple regression analysis predicts on e response variable with plural explanatory variables. The purposes of this paper are reducing calculation time a nd improving reconstruction images. Therefore we move 26 pieces of images from A to Z at various positions. W e prepare for 650 pieces of images in total and perform GA in each.We expect diffusivity for a multiple regressio n analysis from the result and regenerate it.**

Keywords:**Three Dimensional Image Reconstruction ,Computer Generated Hologram Multiple Regression Analysis, Genetic Algorithm, Diffusion Coefficient**

I. INTRODUCTION

There is the Three-Dimensional(3D) reconstruction of an object having difficulty in irradiating a virtual object and laser beam as one of the making purposes of the computer generated Hologram.It is one of the crucial issue to develop a method of the hologram pattern indication that can get high quality reconstruction image with the number of few pixels to be of practical application a CGH.

In the CGH had as one's objective for image reconstruction, the improvement of the reconstruction image will have an important meaning.

Although it is made with the use of repetition algorithm for reconstruction image improvement conventionally, this is improvement of the reconstruction image by optimizing phase to give an input object and is only the indirect improvement method. Therefore a superior reconstruction image can search for the CGH that it is directly when we consider the reconstruction image improvement of the CGH to be a kind of combination optimization problem.We apply the genetic algorithm (GA) that is one of the method of solution to a CGH and perform reconstruction image improvement.GA is algorithm for optimization and a search.

GA can demand the most suitable error diffusion coefficient.It is relations with fitness, individual, diffusion coefficient in GA. And the process of choice, crossing-over, the mutation is important.

In this conventional method, the reconstruction images are improved than we reconstruct only in the error diffu sion method, however length of the calculation time is a fault. In this paper, we played a Computer Generated Hologram (CGH) reconstructing a three-Dimensional (3D) object. We present the improvement of the reconstruction image, re ducing of the calculation time using error diffusion meth od and a multiple regression analysis,
provided CGH is used in this paper and the fault type.

II. To the operation of CGH of GA

Here, We describe the computer-generated hologram. When creating on a computer two-dimensional amplitude hologram, CGH assumes the reference light suitable for calculating the hologram pattern. Since all processing is performed in the computer, the following advantages can be mentioned.

- It is not required for the imaging laser.
- It is not required, such as vibration damping real-time holography.
- It's less noise due to the light-receiving element.
- It can be done by computer generated simulation regeneration process is possible.

A. To the operation of a 3D-CGH reconstruction

Here, using GA, where it describes a method for the synthesis of 3D- CGH of fault type. In this paper, we applied GA to a search of the diffusion coefficient.

The following shows the flow of reconstruction

[Step1]Initialization: Bit string A - D is encoded as an individual in GA to show it in Fig.1. We shall take an integer of -128~+127 each, and bit string A~D has following Eq.(1) between diffusion coefficient A~D and bit string A~D relation having 8

[Step2] Evaluation: We evaluate the reconstruction image from CGH by diffusion coefficient provided from each individual.

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \frac{1}{|A|+|B|+|C|+|D|}\begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} \qquad (1)$$

[Step3]Selection: We evaluate the reconstruction image from CGH by diffusion coefficient provided from each individual.

[Step4] Crossover: Fig.2 shows the operation of the crossover. First of all, We pick out at random under the condition that does not allow duplicates the two individuals. Then, We select the operator having 32 bits at random. About the each individual, and to replace the bit position where the bit of the operator is equal to 1 , as shown in Fig .2.

[Step5] Mutation: We take out one individual optionally and decide an operator at random. Fig.3 is shown mutatetion, it is assumed that obtained by the exclusive OR of the operator and the individual.

[Step6] It is repeat from [Step2] to [Step5] until the number of generation.[3]

A	B	C	D
0101****	0101****	0101****	0101****

$A,B,C,D : -128 \sim +127$ (8bit)

Fig.1 Indivisual.

individual #0	0	0	1	1	0	1	···	1
individual #1	0	1	0	0	1	0	···	0
operator	0	1	1	0	0	1	···	0

↓

individual #0'	0	1	0	1	0	0	···	1
individual #1'	0	0	1	0	1	1	···	0

Fig.2 Crossover.

individual #0	0	0	1	1	0	1	···	1
operator	0	1	1	0	0	1	···	0

↓

individual #0'	0	1	0	1	0	0	···	1

Fig.3 Mutation.

he experiment was carried out by computer generated si mulation.The computer generated simulation assume to use an optical system to show in Fig.4

and carries out it by a diffraction calculation the wave surface recorded to the CGH.[1][2]

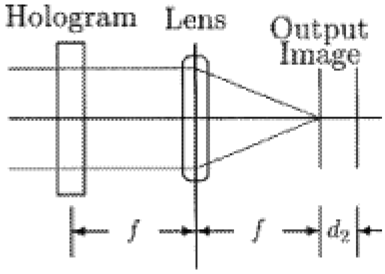


Fig.4 Optical system

And r(u,v) the complex amplitude in the image pla ne f-d_2 from the focal plane away and H(u,v) the complex amplitude in the phase plane that is recor ded as a hologram.The diffraction method uses Fres nel diffraction integral calculus and calculates as fol lows, provided, N is the sample point, L is size of the hologram pattern and k,l,m,s is an integer to ta ke a value of N-1 out of 0. ω is defined by the fo llowing equation.

$$r(m,n) = \sum_{k=0}^{N-1}\sum_{l=o}^{N-1} H'(k,l)$$
$$\cdot \exp\left\{\frac{\pi(d_1-\omega)L^2}{j\lambda d_1^2 N^2}(k^2+l^2\right\} \qquad (2)$$
$$\cdot \exp\left\{\frac{2\pi}{jN}(km+kn)\right\}$$

B. Application of multiple regression analysis

The multiple regression analysis predicts one response variable with plural explanatory variables. The response variable is the quantity made the object of the forecast. It is necessary to collect a large number of sets of data of the result and the corresponding data and activities operation.The explanatory variable is a variable to explain response variable.The experiments were carried out using a linear regression model, such as Eq.(3).

$$y=a_0+a_1 \cdot x_1+\dots+a_p \cdot x_p \qquad (3)$$

y: response variable $x_1 \sim x_p$,:explanatory variables

a_0 :constant term $a_1 \sim a_p$:regression coefficient

p:the number of explanation variables.

$$Q = \sum_{i=1}^{n}\{Y_i - (a_0 + a_1 X_{i1} + a_2 X_{i2})\}^2 \qquad (4)$$

To determine the coefficients which minimize the Q shown by the Eq.(4).

This method called a least square method, coefficients obtained is called the least squares estimate.

We set 0 in partially differentiated by a_0, a_1, a_2 Eq.(4).

$$\begin{cases} \dfrac{\partial Q}{\partial a_0} = -2\sum_{i=1}^{n}\{Y_i - (a_0 + a_1 X_{i1} + a_2 X_{i2})\} \\ \dfrac{\partial Q}{\partial a_1} = -2\sum_{i=1}^{n} X_{i1}\{Y_i - (a_0 + a_1 X_{i1} + a_2 X_{i2})\} \\ \dfrac{\partial Q}{\partial a_2} = -2\sum_{i=1}^{n} X_{i2}\{Y_i - (a_0 + a_1 X_{i1} + a_2 X_{i2})\} \end{cases} \tag{5}$$

Relation obtained when the $\overline{Y}, \overline{X}_1, \overline{X}_2$ the average value of the variable Y, X_1, X_2 to Eq.(5).

$$a_0 = \overline{Y} - a_1\overline{X}_1 - a_2\overline{X}_2 \tag{6}$$

And, change, covariation between the independent variable

And, covariation of independent variable and dependent variable

Organize by substituting.

This is called a normal equation

a_1, a_2 partial regression coefficient is determined by solving the equation.

In this paper, response variables are error diffusion coeff icients and predictor variables are input object position and optimal error diffusion coefficients calculated by GA.

Ⅲ EXPERIMENTAL RESULT

The original images which we used for an experiment 26 pieces of images from A to Z.

Create images total of 25 sheets is displayed at the position of the (32, 32), (0,0), (0,8), ..., "A" and "a.

We prepare for 650 pieces of images in total and perform GA in each.GA is reproduced number 200 individuals, selection probability 0.2, mutation probability 0.01.

The setting in the optical system assumes wavelength λ=632.8nm of the light source (He-Ne laser), focus distance f=300(mm) of the lens, it size L=10(mm) of the hologram pattern.When we use in this paper, it determines a cost function like Eq.(7)and decides to make all a cost function the same, and to reproduce each image.

$$E = E_1 + E_2 = \left\langle \left| |f|^2 - \alpha|r|^2 \right| \right\rangle + \left\langle \left| \frac{|f| - \langle |f| \rangle}{\omega |f|} - \frac{|r| - \langle |r| \rangle}{\omega |r|} \right|^2 \right\rangle \tag{7}$$

We are shown in Fig.4 part of the original image using the experiment.

Fig.5 is result when using the iterative method. Raster scan is used. It is spreading a little result. However, it is not completely diffusion

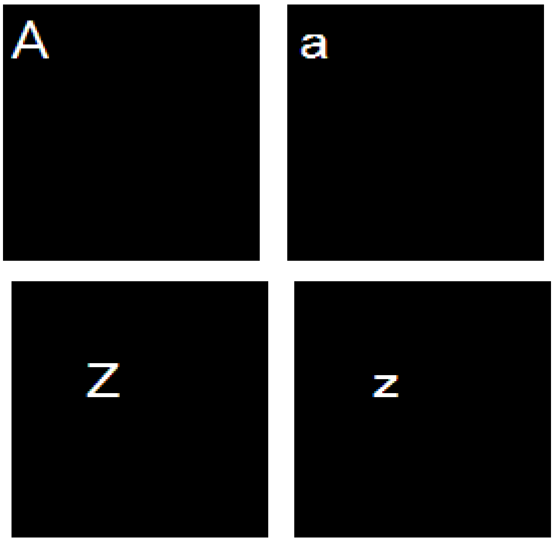
Fig.6 is result when using the GA method. This is rec onstructed cleanly compared with Fig.5.

However, Reproducing time of one image it takes three and a half hours. Fig.7 is one that has been reproduced by predicting the diffusion coefficient by using multiple regression analysis. R egression coefficients are as follows.

Table.1.Diffusion coefficient that led in the multiple regression analysis

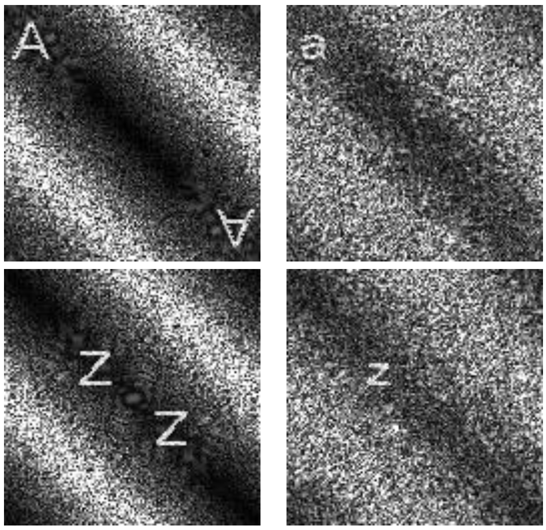
$$\begin{aligned} B_1 &= -0.053755132 + 0.000320769 * a_1 \\ &\quad -0.1591234 * a_2 + 0.18783775 * a_3 \\ B_2 &= -0.50258853 - 0.001005151 * a_1 \\ &\quad -0.139871449 * a_2 + 0.292502106 * a_3 \\ B_3 &= 0.128416966 + 0.002375966 * a_1 \\ &\quad +0.090660218 * a_2 - 0.305896088 * a_3 \\ B_4 &= 0.551553118 + 0.000311596 * a_1 \\ &\quad +0.219211536 * a_2 - 0.407921349 * a_3 \end{aligned}$$

It was an image similar in comparison with the results Fig.6. Reconstruction time took less than one second by one image.



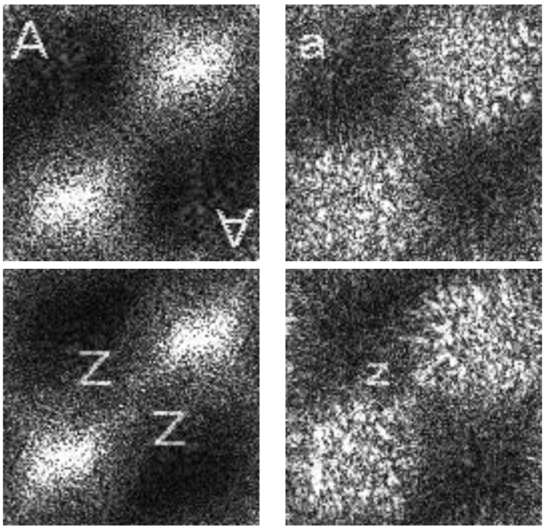
(a)d$_2$=300mm (b)d$_2$=330mm

Fig.4 Original image



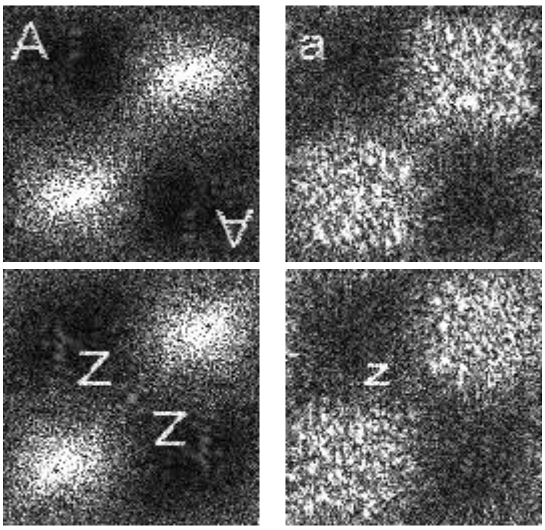
(a)d$_2$=300mm (b)d$_2$=330mm

Fig.5 Reconstructded images(iterative method)



(a)d_2=300mm (b)d_2=330mm

Fig.6 Reconstructded images(GA)



(a)d_2=300mm (b)d_2=330mm

Fig.7 Reconstructded images (multiple regression analysis)

Ⅳ CONCLUSION

In this paper, we applied GA to the CGH of the fault type and the effect and time reduction of reconstruction image improved diffraction calculation was examined from the standpoint of lowering the temperature setting. When applying a multiple regression analysis, it was possible to greatly improve the reconstruction time image is not changed substantially. Also, when you set the diffusion coefficient, in this paper, Again and again conducted an experiment, we have to grasp the trend of the diffusion coefficient.

Ⅴ REFERRNCES

[1] Ken-ichi Tanaka,Yoshimasa Sakamoto,and Teruo Shimomura,"Effect of Cost Function on Computer Generated Hologram Using Simulated Annerling",IEICE,C-I,Electronics, J80-C-1(2),100-104,1997.

[2] Miki Mitsunori,Hiroyasu Tomoyuki,and Ono Keiko,"Simulated Annealing with Advanced Adaptive Neighborhood"IPS,44(1),1-6,2003

[3] Ken-ichi Tanaka," Genetic Algorithm Based Method of Estimating Optimal Error-Diffusion for Computer-Generated Holograms(Special Section Three-Dimensional Image Information Media) ITE , 54(3), 394-401, 2000

Embedding used binary number conversion into a Color image (P)

Hiroaki Oguma
Department of Electronics and Bioinformatics, Meiji University
1-1-1, Higashi-Mita, Tama, Kawasaki, Japan
ce31016@meiji.ac.jp
Ken-ichi Tanaka
Department of Network Design, Meiji University
4-21-1, Nakano, Nakano-ku, Tokyo, Japan
tanaken@isc.meiji.ac.jp

Abstract: Digital watermarking is widely used as security of a binary format image. The purpose of this paper is extension to a color image from grayscale image. Then, it thought as important giving many embedding information to a color image. A color image is divided into RGB, the concentration value of a pixel is made into a binary number, and last 3 digits are changed to arbitrary numbers. It embeds by carrying out by repeating this, and secure many information. 10x19660 times of embedding are able to be performed to the picture of 256x256 pixels. Moreover, the signal to noise ratio is about 40dB. Therefore, it seems that there is little influence which it has on a picture.
Keywords: Digital watermarking, Binary number, RGB,

I. Introduction

In recent years, in addition to rapid osmosis of the Internet, it is becoming the environment where everyone can circulate easily the digital information which led the network by the spread of personal digital assistants.

In connection with this, practical use we which gathers information according to its purpose is performed out of the information which various people built.

On the other hand, as for the digital contents which have value as social wealth, the kind of media also spreads from numerical text to a still picture, a sound, and video by progress of an information technology, and a sale of the movie and musical piece using a network also came to be made.

However, since it has the feature that a duplicate is easy and does not deteriorate, digital information is, when circulating worthy information, the structure which prevents unjust use and protects intellectual property rights is indispensable [1].

After generating of the illegal use of a work, detection, discernment, and proof are enabled and digital-watermarking technology is mentioned as a representative case of the technology in which this expects deterrence of an illegal use indirectly [2].

Digital watermarking is technology which embeds watermark information, as not perceived by man at digital contents, such as a picture, an image, and a sound.

Even if ID information and copyright information are embedded at the contents themselves if it is digital watermarking, and partial cutoff etc. are moreover edited, it is also possible to take out the embedded information.

As for a binary format image, many security methods exist focusing on a document picture. Although how to change between characters, and the method of putting a design into a background are famous, they also have many problems, there being little amount of information of a watermark, or the use range being restrictive [3] [4].

The previous techniques is performed the digital-watermarking method for a color image, therefore still now [5]. In this paper, the technique of embedding with emphasis on giving many information to a color image is considered.

II. Experimental

A color image is divided into a RGB ingredient and the concentration value of a pixel is acquired. The acquired concentration value is changed into a binary number, and the beforehand fixed embedding information bit is put into last 3 digits portion one after another. Many embedding information is secured by performing the above processing to all the pixels of RGB each color.

The bit value sequence of 0 and 1 used as the contents embedded as a sample and picture figure 1 of 256x256 were prepared. Here, the bit value sequence of 10 digits of (0,1,0,1,0,1,0,1,0,1) was used.

Fig.1 Original picture image

Fig.2 After embedding

Table.1 MSE and the signal to noise ratio in embedding picture

Name	Embed	Extranction	Chenge	MSE	S/N[dB]
Parrots	19660	19660	65360	7.784	39.2188
Girl	19660	19660	65434	7.437	39.4168
Pepper	19660	19660	65432	7.574	39.3376
Earth	19660	19660	65415	7.416	39.4291
Aerial	19660	19660	65412	7.559	39.3462
Average	19660	19660	65410.6	7.554	39.3497

It separates into each color of RGB, Fig. 1 is shown, and a concentration value is acquired per pixel from the upper left to the lower right.

The acquired numerical value is changed into a binary number, and it replaces in the bit value sequence which was preparing the last 3 digits. The numerical value after replacing is re-changed into a decimal number, and it stores in the pixel of a basis, and moves to the following pixel. In the next replacement, the 4th to 6th three of the prepared bit value sequence is replaced.

This processing is performed one after another and embedding is performed to all the pixels. After embedding is completed to the last, the separated RGB ingredient is compounded and it returns to a color image. The picture after embedding is shown in Fig. 2. The signal to noise ratio of Fig. 2 was investigated.

About Fig. 2, it separates into each color of RGB and reads, after acquiring a concentration value, binary number conversion is carried out, and lower 3 figure is recorded. It is checked whether it is in agreement with the bit value sequence which connected and embedded the bit value sequence which is read one by one and recorded.

Moreover, in the pictures I and K of picture size m×n, the maximum pixel value which a picture can take was defined as MAX, and the mean squared error (MSE) and the peak signal-to-noise ratio (S/N) were calculated from the formula (1) and (2).

$$\mathrm{MSE} = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2 \qquad (1)$$

$$\frac{\mathrm{S}}{\mathrm{N}} = 10\log_{10}\frac{MAX^2}{MSE}\ [\mathrm{dB}] \qquad (2)$$

The calculation result to some pictures is shown in Table 1.

III. CONCLUSION

It succeeded in repeating the bit value sequence of 10 digits of (0,1,0,1,0,1,0,1,0,1) 19660 times, and embedding it to the picture of 256x256. Moreover, extraction was also successful same. Since the signal to noise ratio of the embedded picture was about 40 dB, it can be said to be that there are little change and sense of incongruity before and behind embedding. In the viewpoint of the quantity of the embedded bit value sequence, the technique of this paper has secured sufficient amount of information to use for digital watermarking.

REFERENCES

[1] Norihiko Sakurai, Yutaka Kidawara, Yoichi Takashima, Noburou Taniguchi, and Koji Namba, A Survey on Copyright Protection Technologies for Network Distribution of Digital Information, Information Processing Society of Japan, Vol.42, No. SIG 15(TOD 12), p-p63-76, (2001)

[2]Ryoichi Sasaki, Hiroshi Yoshida, Consideration on Copyright and Illegal Copy Countermeasures under IT Revolution, Information Processing Society of Japan, CSEC, Vol.13, No.7, p-p37-42, (2001).

[3] Cox IJ, Kilian J, Leighton FT, and Shamoon T, Secure spread spectrum watermarking for multimedia, IEEE Transactions on Image Processing, Vol.6, No.12, p-p1673-1687, (1997).

[4] Jack T.Brassil, Steven Low, Nicholas F.Maxemchuk, Lawrence O'Gorman, Electronic marking and identification techniques to discourage document copying, IEEE Journal on Selected Areas in Communications, Vol.13, No.8, p-p1495-1504, (1995).

[5] Go Tanaka, Noriaki Suetake, and Eiji Uchino, High Capacity Data Embedding Method for Palette-Based Images, Technical Report of the Institute of Electronics, Information and Communication Engineers, SIS, a smart information media system, Vol.108, No.461, p-p13-16, (2009).

[6] Hiroyuki Kitayama, C++Builder6 component practical use guide & practice programming Vol.6, a cut system, pp203-352, (2004).

[7] Yasushi Yamanaka, Hiroshi Fujii, Digital Image Scrambling Method for Information Distribution, Information Processing Society of Japan, Vol.38, No.10, p-p1945-1955, (1997).

[8] Yasushi Abe, Koichi Inoue, Digital Watermarking for Bi-Level image, Ricoh Technical Report, No.26, pp8-16, (2000).

[9] Naohisa Komatsu, Ken-ichi Tanaka, digital-watermarking technology : the security of digital contents, Tokyo Denki University Press, (2004).

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Markov Chain Analyses of Random Local Search and Evolutionary Algorithm

Hiroshi Furutani[1], Hiroki Tagami[1], Ichihi To[2], Makoto Sakamoto[3]

[1] Faculty of Engineering, University of Miyazaki, Japan
[2] Graduate School of Engineering, University of Miyazaki, Japan

e-mail: furutani@cs.miyazaki-u.ac.jp

Abstract: Theoretical studies of evolutionary algorithms (EAs) have been developed by researchers whose main interests are convergence properties of algorithms. In this paper, we report the computational complexity of an algorithm that is a variant of (1+1) EA, called Random Local Search (RLS). While a standard EA uses a mutation of flipping each bit in a parent string, RLS flips exactly one bit at each step. It has been noted the close resemblance of RLS with the coupon collector problem (CCP). CCP has a long history of probabilistic research, and many interesting results are obtained. This study makes use of such results with some modifications. We also show some useful results representing the evolution process of $(1 + 1)$ EA.

Keywords: Evolutionary algorithm, Random Local Search, Coupon collector's problem , OneMax function, Markov chain

1 INTRODUCTION

Theoretical studies of evolutionary algorithms (EAs) have been developed by researchers whose main interests are convergence properties of algorithms [1, 2]. Especially, the time complexity of algorithm is a most attracting topic for them. For the runtime analysis of EAs, one usually chooses (1+1) EA as the first trial. A detailed overview of these studies are presented in [3].

In this paper, we report the computational complexity of an algorithm called Random Local Search (RLS), which is a variation of (1+1) EA. While a standard EA uses a mutation of flipping each bit in a parent string, RLS flips exactly one bit at each step. It has been noted the close resemblance of RLS with the coupon collector's problem (CCP). The CCP has a long history of probabilistic research, and many interesting results are obtained [4]. This study makes use of such results with some modifications. We show some useful theorems representing the evolutional process of RLS.

A general framework for analyzing the average hitting times of EAs by applying Markov chain model is studied by many researchers. The state space S of the chain is given by the number of bit ones, $\{0, 1, \ldots, \ell\}$, where ℓ is the length of bit string.

The behavior of this Markov chain is completely determined by the transition matrix $\boldsymbol{P}$ of an absorbing Markov chain. The explicit form of absorbing Markov chain can be obtained. Using this expression of $\boldsymbol{P}$, we show that the expected runtime of RLS algorithm is ℓH_1 steps, where H_1 is the first Harmonic number. The runtime of this chain is approximately given by $\ell \log(\ell)$. Furthermore, the distribution of runtime T is also obtained, and can be expressed in the closed form by using Stirling number of the second kind. There are many such interesting results for the RLS algorithm. The paper reports some of them.

2 EVOLUTIONARY ALGORITHMS

Evolutionary Algorithm $(\mu + \lambda)$ came from Evolutionary Strategy developed by Rechenberg and Schwefel, where μ and λ are numbers of parent solutions and offspring solutions, respectively. We choose the parameters $\mu = 1$ and $\lambda = 1$ for simplicity. Though (1+1) EA seems too simple for an analysis, it has attracting properties. As Wegener states that [2],

(1) it is efficient for many problems,

(2) it can not get stuck in a local optimum, and

(3) the analysis of it reveals many tools that can be used in more practical algorithms.

Algorithm 1 (1+1) EA

1 Initialize $x \in \{0, 1\}^\ell$ uniformly at random.
2 Create x' by flipping each bit in x with probability p_m.
3: Select if $f(x') \geq f(x)$ then $x := x'$.
4: Go to 2 until a termination condition is fulfilled.

As a test function, we apply OneMax function $f(x)$

$$f(x) = \sum_{i=1}^{\ell} x_i, \quad x_i \in \{0, 1\},$$

where x is a binary string of length ℓ. We consider the maximization problem of OneMax function. The optimum solution is $x_{opt} = \{1\}^\ell$, and $f(x_{opt}) = \ell$. Since many studies suggested that the mutation probability of $p_m = 1/\ell$ may be the best choice, we carried out our analysis using this value.

The second choice of Evolutionary Algorithm is the Random Local Search (RLS). We define RLS as

Algorithm 2 Random Local Search

1 Initialize $x \in \{0,1\}^\ell$ uniformly at random.
2 Create x' by flipping one bit in x which is selected at random.
3: Select if $f(x') \geq f(x)$ then $x := x'$.
4: Go to 2 until a termination condition is fulfilled.

The difference of two algorithms are in the step 2.

3 COUPON COLLECTOR'S PROBLEM

Consider the case that there are c types of coupons, and each day a collector randomly gets one coupon with the equal probability $1/c$. How many days has the collector to wait for collecting all types of coupons.

As an elementary problem, we consider the number of days n until both of the 2 types of coupons, A and B, are collected. First, we calculate the expectation value $E[n]$. If two types of coupons are collected at the nth day, there are two cases,

$$\underbrace{\mathrm{AAA}\cdots\mathrm{A}}_{\mathrm{n}-1}\mathrm{B}, \qquad \underbrace{\mathrm{BBB}\cdots\mathrm{B}}_{\mathrm{n}-1}\mathrm{A}.$$

Since there are n^2 cases, the probability of collecting one set of coupons is

$$p^{(2)}(n) = \frac{2}{2^n} = \frac{1}{2^{n-1}}.$$

The expectation of n is given by

$$E[n] = \sum_{n=2}^{\infty} n \, (\frac{1}{2})^{n-1} = \frac{1}{(1-1/2)^2} - 1 = 3.$$

Next, we try to solve this problem in another approach. First, we calculate the probability $q^{(2)}(n)$ of collecting one set of coupons until n days. Among all events in n days, there are two events which do not collect at least one of two coupons,

$$\underbrace{\mathrm{AAA}\cdots\mathrm{A}}_{n}, \qquad \underbrace{\mathrm{BBB}\cdots\mathrm{B}}_{n}.$$

Hence, we have

$$q^{(2)}(n) = 1 - \frac{2}{2^n} = 1 - \frac{1}{2^{n-1}}.$$

Using this result, we have

$$p^{(2)}(n) = q^{(2)}(n) - q^{(2)}(n-1) = \frac{1}{2^{n-1}}.$$

Then, we consider the case of c coupons. Let A_i be the event that the ith coupon is still not collected at the nth day. Probabilities $P(A_i)$ and $P(A_i \cup A_j)$ with $(i \neq j)$ are given by

$$P(A_i) = \left(1 - \frac{1}{c}\right)^n, \quad P(A_i \cup A_j) = \left(1 - \frac{2}{c}\right)^n,$$

We define events F and G. Event F is the case of failure that at least one coupon is still not collected at the nth day, while G is the case of success that a collector gets all coupons until the nth day. Since the relation

$$P(G) = 1 - P(F),$$

holds, we calculate the failure probability $P(F)$ instead of of the success probability $P(G)$. Event F is given by

$$F = A_1 \cup A_2 \cup \ldots \cup A_c = \bigcup_{i=1}^{c} A_i.$$

Using the principle of inclusion-exclusion

$$\begin{aligned} P(\bigcup_{i=1}^{c} A_i) &= \sum_{i=1}^{c} P(A_i) - \sum_{i1<i2} P(A_{i1} \cap A_{i2}) \\ &+ \cdots + (-1)^{c+1} P(A_1 \cap \cdots \cap A_c). \end{aligned}$$

Thus we have the probability of collecting all types of coupons until nth day

$$q^{(c)}(n) = 1 - P(F) = \sum_{i=0}^{c} \binom{c}{i} (-1)^i \left(1 - \frac{i}{c}\right)^n. \quad (1)$$

The final result of the probability of collecting c types of coupons at the nth day is

$$\begin{aligned} p^{(c)}(n) &= q^{(c)}(n) - q^{(c)}(n-) \\ &= \sum_{i=0}^{c-1} \binom{c-1}{i} (-1)^i \left(1 - \frac{i}{c}\right)^{n-1}. \quad (2) \end{aligned}$$

4 MARKOV CHAIN MODEL

This section presents the Markov chain approaches to the EAs.

4.1 (1+1) Evolutionary Algorithm

The search space of OneMax function is $\Omega = \{0,1\}^\ell$, and we divide Ω into $(\ell+1)$ subsets $\Omega = S_0 \cup S_1 \cup \cdots \cup S_\ell$, where $f(S_i) = i$. The transition matrix $P_{i,j} = P(j|i)$ represents the evolution of $(1+1)$ EA,

(1) for $i > j$

$$P_{i,j} = 0.$$

(2) for $i < j$

$$\begin{aligned} P_{i,j} &= \sum_{r=0}^{k} \binom{i}{r} {p_m}^r (1-p_m)^{i-r} \binom{\ell - i}{j-i+r} \\ &\times {p_m}^{j-i+r} (1-p_m)^{\ell-i-(j-i+r)} \\ &= \sum_{r=0}^{k} \binom{i}{r} \binom{\ell - i}{s} {p_m}^{r+s} (1-p_m)^{\ell-(r+s)} \quad (3) \end{aligned}$$

where

$$s = j - i + r, \quad k = \min(\ell - j, i).$$

(3) for $i = j$

$$P_{i,i} = 1 - \sum_{j>i}^{\ell} P_{i,j}.$$

In a matrix form, these equations are given by

$$\boldsymbol{P} = \begin{pmatrix} P_{0,0} & P_{0,1} & \cdots & \cdots & P_{0,\ell} \\ 0 & P_{1,1} & \cdots & \cdots & P_{1,\ell} \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \tag{4}$$

The form of $\boldsymbol{P}$ suggests that this Markov chain is absorbing, and there are ℓ transient and one absorbing states, respectively. The canonical form of the transition matrix is given by

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{Q} & \boldsymbol{R} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \tag{5}$$

The $(\ell \times \ell)$ submatrix $\boldsymbol{Q}$ represents the transition between $S_0, S_1, \ldots, S_{\ell-1}$. The unit matrix $\boldsymbol{I}$ in this case is a scalar 1.

For the calculation of the hitting time of the optimum solution, we use the fundamental matrix

$$\boldsymbol{N} = (\boldsymbol{I} - \boldsymbol{Q})^{-1}, \tag{6}$$

where the size of $\boldsymbol{I}$ is the same as $\boldsymbol{Q}$. The Markov chain theory tells that the expected steps to enter into the absorbing state(s) are given by [5]

$$\boldsymbol{m} = \boldsymbol{N}\,\boldsymbol{1},$$

where m_i is the expected absorbing steps started from S_i, and $\boldsymbol{1}$ is a colummn vector whose all entries are 1. By solving a set of linear equations

$$(\boldsymbol{I} - \boldsymbol{Q})\,\boldsymbol{m} = \boldsymbol{1}, \tag{7}$$

we obtain the solutions

$$m_{\ell-1} = 1/p_{\ell-1,\ell},$$

and for $(0 \leq i \leq \ell - 1)$

$$m_i = (1 + \sum_{j=i+1}^{\ell-1} P_{i,j}\, m_j)/(\sum_{j=i+1}^{\ell} P_{i,j}).$$

4.2 RLS algorithm

The transition matrix for the RLS process is given by

$$\boldsymbol{P} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1/\ell & (\ell-1)/\ell & 0 & \cdots & 0 \\ 0 & 0 & 2/\ell & (\ell-2)/\ell & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & (\ell-1)/\ell & 1/\ell \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \tag{8}$$

From eq.(7)

$$\begin{aligned} \frac{1-i}{\ell}(m_i - m_{i+1}) &= 1, \quad (0 \leq i \leq \ell - 2) \\ \frac{1}{\ell} m_{\ell-1} &= 1, \end{aligned}$$

we have a set of solutions

$$m_i = \ell\left(1 + \frac{1}{2} + \cdots + \frac{1}{\ell - i}\right). \tag{9}$$

Thus the expected first hitting time of the optimum solution in RLS is given by eq.(9) for i, the number of bit ones in the initial state. The distribution of the first hitting time n is given by $p^{(\ell)}(n)$ for $i = 0$ using eq.(2).

5 NUMERICAL EXPERIMENT

In this section, we compare results of the theoretical predictions with experiments of RLS and $(1 + 1)$ EA. We used the mutation rate $p_m = 1/\ell$ for $(1 + 1)$ EA. The length of string is $L = 100$, and We performed 10000 runs for each parameter set, and averaged over them.

Figure 1 shows the time dependence of probability $p^{(\ell)}(n)$ of the first hitting time of optimum solution in RLS. The solid line is the result of numerical calculation. The result is moving averaged with window size of 21. The dotted line is the theoretical prediction obtained by eq.(2), which well reproduces the result of the numerical calculation.

Figure 2 shows the time dependence of probability $p^{(\ell)}(n)$ of the first hitting time of optimum solution in $(1 + 1)$ EA. The solid line is the result of numerical EA calculation. The dotted line is the theoretical prediction obtained by using Markov chain transition matrix.

6 SUMMARY

In this paper, we demonstrate that the probabilistic methods including Markov chain model can reproduce the first hitting time of optimum solution in RLS and $(1 + 1)$ EA for OneMax function. These studies can help to understand the working mechanism of Evolutionary Algorithms, and give some suggestions to design algorithms for other problems. We cannot show results for linear functions obtained by present algorithms and methods of population genetics [6]

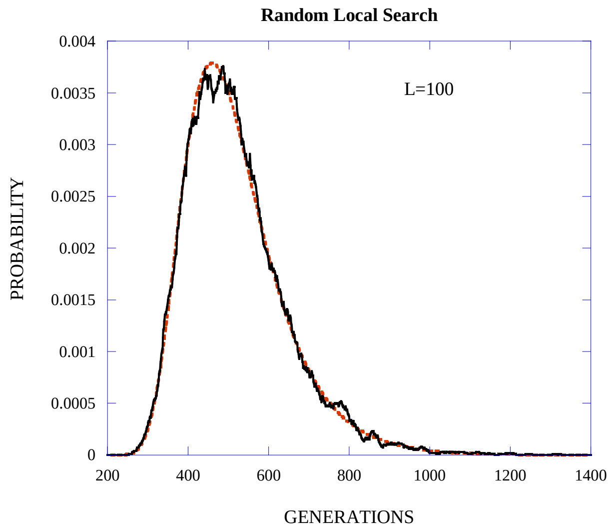


Fig. 1. Distribution of the first hitting time of optimum solution in RLS with $\ell = 100$. The initial state is a bit string of all zeros. The solid line is the result of RLS calculation. The dotted line is the theoretical prediction.

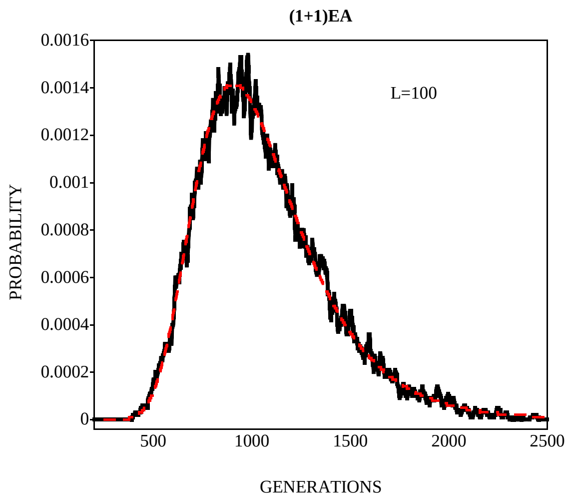


Fig. 2. Distribution of the first hitting time of optimum solution in $(1+1)$ EA with $\ell = 100$. The initial state is a bit string of $\ell/2$ zeros and $\ell/2$ ones. The solid line is the result of $(1+1)$ EA calculation. The dotted line is the theoretical prediction using Parkov chain model.

due to limitation of space. We will report them in other occasions.

Our next aim is to apply drift analysis to optimization problems. Drift analysis was a class of tools developed by Yao and other researchers [3, 7], and can give upper and lower bounds to running time of a given algorithm. We are now trying to combine drift analysis and classical probabilistic methods.

REFERENCES

[1] A. Auger and B. Doerr (eds.) (2011), *Theory of Randomized Search Heuristics*, World Scientific, Singapore.

[2] I. Wegener (2001), Methods for the analysis of evolutionary algorithms on pseudo-Boolean functions, in *Evolutionary Optimization*, Kluwer Academic Publishers, Dordrecht, pp.349–369.

[3] P. S. Oliveto and X. Yao (2011), Runtime analysis of evolutionary algorithms for discrete optimization, in [1], pp.21–52.

[4] T. Nakata and I. Kubo (2006), A coupon collector's problem with bonuses, in *Proceedings of Discrete Mathematics and Theoretical Computer Science*, **AG**, pp.215–224.

[5] M. Iosifescu (1980), *Finite Markov chain processes and and their applications*, John Wiley & Sons, New York.

[6] J. W. J. Ewens (2004), *Mathematical Population Genetics. I. Theoretical Introduction*, Second Edition. Springer-Verlag, New York.

[7] B. Hajek (1982), Hitting-time and occupation-time bounds implied by drift analysis with applications, *Advances in Applied Probability* **13**, pp.502–525.

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Runtime Analysis of OneMax Problem in Genetic Algorithm

Ichihi To[1], QuinLian Ma[2], Makoto Sakamoto[3], Hiroshi Furutani[3], and Yu-an Zhang[4]

[1]Graduate School of Engineering, University of Miyazaki, Japan
[2]Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki, Japan
[3]Faculty of Engineering, University of Miyazaki, Japan
[4]Department of Computer Science and Technology, Qinghai University, China

e-mail: furutani@cs.miyazaki-u.ac.jp

Abstract: Genetic algorithms (GAs) are stochastic optimization techniques that simulate the biological evolution. Theoretical study of the evolution of GA is very important for the application of GA. We have studied the effects of stochastic fluctuation in the process of GA evolution. A mathematical study is carried out for GA on OneMax function within the framework of Markov chain model. We obtain the steady state solution, which represents a distribution of the first order schema frequency. This paper treats the task of estimating hitting time of the optimum solution for OneMax problem, and study the effect of mutation probability on the hitting time.

Keywords: genetic algorithms, schema theory, OneMax problem, Wrihgt-Fisher model, Markov model

1 INTRODUCTION

We consider the probability that a population includes the optimum solution by applying Markov chain model. We call this probability as the success probability of GA. We obtained the analytical form of the transition matrix P of Markov chain, which represents the evolution of population in OneMax problem. We found that GA for the OneMax problem is equivalent to the asymmetric mutation model. There are rich studies of the asymmetry mutation model, and we can apply these results to the OneMax problem in GA. The relation between the convergence time and success probability was studied analytically and experimentally. In the analysis, we applied the Wright-Fisher model.

This paper treats the task of estimating hitting time of the optimum solution in GA for OneMax problem. This task is performed by using the Markov chain representing the behavior of population in the stationary state of evolution. The transition matrix of the stationary state is calculated by using the original transition matrix P. By dividing the states to populations which include the optimum solution and populations which include no optimum ones, we calculate the transition matrix. By doing these transformation, we find that the hitting time in the stationary state can be represented approximately by using the success probability, mutation rate and bit length. We use this theoretical result for the analysis of real GA calculations, and study the effect of success probability on the hitting time of GA.

2 MATHEMATICAL MODEL

2.1 OneMax Model

We treat the evolution process of a population with N individuals. The individuals are represented by binary strings of length ℓ, and there are $n = 2^\ell$ genotypes,

$$i = < i(\ell), \cdots, i(1) >, \quad i(k) \in \{0, 1\}.$$

We use the relative frequency $x_i(t)$ for describing the evolution

$$x_i(t) = N_i(t)/N,$$

where $N_i(t)$ is the number of individuals of genotype i at generation t.

The relative frequencies satisfy the normalization condition

$$\sum_{i=0}^{n-1} x_i(t) = 1.$$

The average fitness $\bar{f}(t)$ of the population at generation t is

$$\bar{f}(t) = \sum_{i=0}^{n-1} f_i\, x_i(t). \tag{1}$$

The OneMax fitness function f_i is defined as

$$f_i = \sum_{k=1}^{\ell} i(k). \tag{2}$$

Thus the string of all ones $< 1, 1, \ldots, 1 >$ is the optimum solution of this function.

2.2 Linkage Equilibrium

We derive here the evolution equation for the first order schema frequency. To do this, we introduce the notion of linkage equilibrium. Linkage means the correlation between the different loci in a chromosome, and if there is some correlation we call this state as linkage disequilibrium [1]. Crossover and mutation have the effect of making the population in linkage equilibrium.

The distribution of a population in linkage equilibrium is given by

$$x_i(t) \;=\; \prod_{k=1}^{\ell} h_{i(k)}(t), \tag{3}$$

where $h_{i(k)}(t)$ is a frequency of the first order schema at position k, and $i =< i(\ell), \ldots, i(1) >$. We also use the notation of $h_0^{(k)}$ and $h_1^{(k)}$ for the first order schema frequencies of bit 0 and bit 1, respectively.

2.3 Evolution Equation of the First Order Schema

In the deterministic schema theory, the evolution of the first order schema in linkage equilibrium is given by [2, 3]. The relative frequency of the first order schema at position k is determined by

$$h_1^{(k)}(t+1) \;=\; a h_1^{(k)}(t) + b, \tag{4}$$

where constants a and b are

$$a \;=\; (1-\frac{1}{\ell})\,(1-2p_m), \quad b \;=\; \frac{1}{\ell}(1-2p_m)+p_m.$$

3 ANALYSIS OF EVOLUTION RATE

3.1 Fisher's Theorem

According to Fisher's "Fundamental Theorem of Natural Selection"[3], evolution rate becomes faster if variance of fitness becomes bigger. The change $\Delta\overline{f}(t)$ of average fitness $\overline{f}(t)$ at generation t is

$$\Delta\overline{f}(t) = \overline{f}(t+1) - \overline{f}(t). \tag{5}$$

Variance of average fitness VAR(f) is

$$VAR(f) = \sum_i f_i^2 x_i(t) - \overline{f}(t)^2. \tag{6}$$

From eqs.(5) and (6),we have

$$\Delta\overline{f}(t) = \frac{1}{\overline{f}(t)} VAR(f).$$

We call $v(t) \equiv \Delta\overline{f}(t)$ as evolution rate

$$v(t) \equiv \Delta\overline{f}(t) = \overline{f}(t+1) - \overline{f}(t).$$

If we ignore mutation, evolution rate $v(t)$ is proportional to the variance of average fitness. We note that variance of fitness must be increased in order to promote the evolution.

3.2 Experiment

We examined the effect of crossover by numerical calculation for OneMax problem. The length of string is $L = 20$, population size $N = 100$ and mutation rate $p_m = 0.02$. We appled roulette selection. We averaged the results obtained by repeating 1000 same calculations. The initial state is randomly generated with $p^{(1)} = 1/\ell$. We compared three types of calculations; without crossover, one-point crossover, and uniform crossover. We investigated the change of evolution rate by crossover.

Figures 1 and 2 show values of average and variance of fitness.

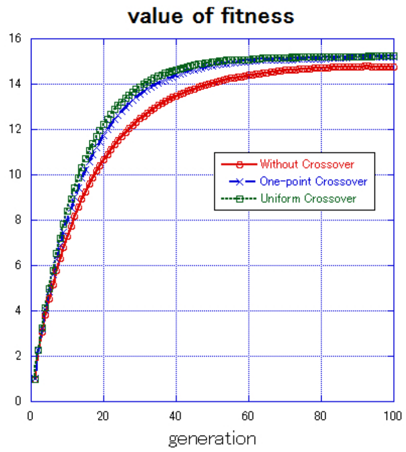


Fig. 1. Average fitness

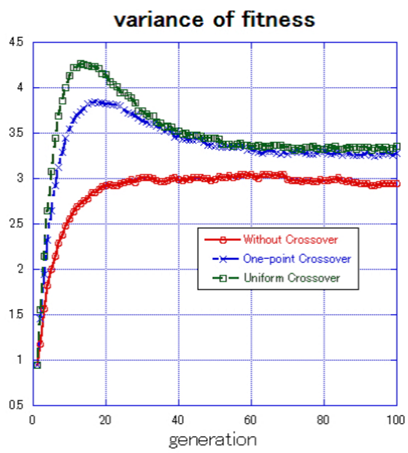


Fig. 2. Variance of fitness

We find that Fisher's theorem explains the results of Figs 1 and 2.

3.3 Linkage Analysis

The linkage disequilibrium coefficient is

$$D[m, m'] = p^{(11)}_{m,m'} - p^{(1)}_m p^{(1)}_{m'}. \tag{7}$$

Here, $p_m^{(1)}$ is the probability that the m-th bit is 1, and $p_{m,m'}^{(11)}$ is the probability that the m-th bit and the m'-th bit are 1. Using Walsh function and Walsh conversion, the variance is given by

$$VAR(f) = V_A + V_I, \tag{8}$$

where

$$V_A = \sum_{m=0}^{L} p_m^{(1)}(1 - p_m^{(1)}), \tag{9}$$

$$V_I = 2 \sum_{m<m'} D[m, m']. \tag{10}$$

Since V_A depends only on $p_m^{(1)}$, which is the function of one bit, it is not affected directly by crossover. On the other hand, V_I dependents on the correlation of two loci. Therefore, the value varies by crossover. If correlation of two loci is small, the absolute value of V_I is small. In addition,because

$$V_I \leq 0, \tag{11}$$

variance increases with weaker correlation.

3.4 Analysis of Variance

We used the same conditions as the experiment of section 3.2, and analyzed the variance of Fig.2 by linkage analysis. Figure 3 shows the change of V_A and V_I with generation.

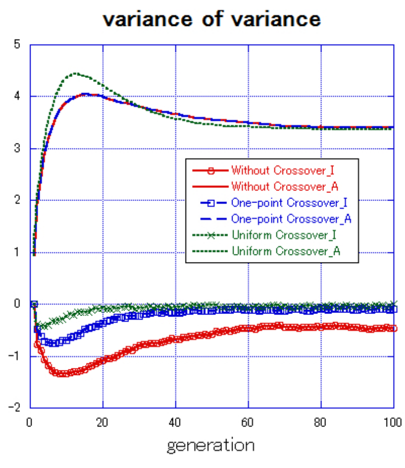


Fig. 3. Change of V_A and V_I with generation

3.5 Analysis Result

From this calculation,we found that the variance of fitness is influenced by crossover. Crossover breaks the correlation between two loci, and a population will be in linkage equilibrium. In the case of uniform crossover, we found that the variance of fitness is the largest, and evolution rate is the fastest.

4 ANALYSIS OF CONVERGENCE

4.1 Markov Model

In this study,we use the concept of Markov chain. Markov chain is a stochastic process that the transition only depend on the previous state.

One of the Markov model in population genetics is a Wright-Fisher model. Wright and Fisher put forward their evolution equation by considering the finite individuals [1]. The Wright-Fisher model treats chromosomes having one locus and two alleles, corresponding to the GA of $\ell = 1$ with genotypes $k \in \{0, 1\}$ [4]. The number of the first genotype 1 takes the values of

$$N_1 = 0, 1, \ldots, N,$$

and that of the genotype 0 is given by $N_0 = N - N_1$.

We analyze evolution processes by taking into account the effect of random sampling, and consider the fitness proportionate selection. If there are $N_1 = i$ copies of the genotype 1 at the current generation t, the probability $P(j|i)$ of N_1 taking the value of j at the next generation $t + 1$ is given by the binomial distribution

$$P_{i,j} = P(j|i) = \binom{N}{j} p_i^j (1 - p_i)^{N-j} \tag{12}$$

$$p_i = ay + b = a(\frac{i}{N}) + b,$$

where i is the number of genotype 1 at the current generation t, and j is the number of genotype 1 at the next generation $t + 1$.

The probability $P(j|i)$ specifies the process of random sampling, and the future behavior of the process only depends on its current frequencies [4]. Thus this process is a Markov chain. Let $\mu_i(t)$ be the probability that the population is in the state of $N_1 = i$ at generation t.

In the following, we use the vector notation

$$\mu^T = (\mu_0, \mu_1, \ldots, \mu_N),$$

with the normalization condition

$$\sum_{i=0}^{N} \mu_i(t) = 1. \tag{13}$$

Then the evolution process is described by

$$\mu_j(t + 1) = \sum_{i=0}^{N} \mu_i(t) P_{i,j}. \tag{14}$$

The evolution equation is given in the vector form

$$\mu(t + 1)^T = \mu(t)^T P, \tag{15}$$

where P is a matrix of the size $(N + 1) \times (N + 1)$. We know the eigenvalues of this matrix [5],

$$\lambda_0 = 1, \quad \lambda_1 = a, \quad \lambda_2 = a^2 (1 - 1/N), \ldots. \tag{16}$$

It should be noted that the second largest eigenvalue λ_1 does not depend on the population size N. We denote the left and right eigenvectors

$$\mathbf{u}_i^T P = \lambda_i \mathbf{u}_i^T, \quad P \mathbf{v}_i = \lambda_i \mathbf{v}_i, \quad (0 \le i \le N).$$

These eigenvectors satisfy the orthogonality condition

$$\mathbf{u}_i^T \cdot \mathbf{v}_j = 0 \quad (i \ne j). \tag{17}$$

The explicit form of the right eigenvector $\mathbf{v}_0$ is given by

$$\mathbf{v}_0 = (1, 1, \ldots, 1)^T, \tag{18}$$

and all elements of the left eigenvector $\mathbf{u}_0$ are positive. We adopt the normalization conditions

$$\sum_{i=0}^{N} |u_i| = 1, \quad \mathbf{u}_i^T \cdot \mathbf{v}_i = 1.$$

We consider the GA under positive mutation rate $p_m > 0$. In this case, all elements of the transition matrix P are positive, and the Markov chain of schema evolution is irreducible and aperiodic. The Markov chain theory states that an irreducible and aperiodic Markov chain converges to the stationary distribution π

$$\lim_{t\to\infty} \mu(t) = \pi, \tag{19}$$

and all elements π_i are positive [5].

The initial distribution of the first order schema can be expanded in terms of the left eigenvectors

$$\mu(t=0)^T = \sum_{i=0}^{N} C_i \mathbf{u}_i^T. \tag{20}$$

Since $\mathbf{u}_i$ is eigenvector, and from equation (15), we have

$$\mu(t)^T = \sum_{i=0}^{N} \lambda_i^t C_i \mathbf{u}_i^T. \tag{21}$$

Multiplying the right eigenvector $\mathbf{v}_0$ from right side

$$\mu(t)^T \cdot \mathbf{v}_0 = \sum_{i=0}^{N} \lambda_i^t C_i \mathbf{u}_i^T \cdot \mathbf{v}_0 = \lambda_0^t C_0 = C_0,$$

and from equations(13) and (18), we have $C_0 = 1$. At large t, $\mu(t)$ is approximately given by

$$\mu(t) \approx \mathbf{u}_0 + a^t C_1 \mathbf{u}_1, \tag{22}$$

and $\mathbf{u}_0$ is the stationary distribution π.

We treat the convergence time T_c of Markov chain. The theoretical estimation of the convergence time is given by

$$T_c(\text{eigenvalue}) = \min_t \{a^t < 0.05\}, \tag{23}$$

where a is the second largest eigenvalue of the transition matrix P.

4.2 Experiment

We compared results of the theoretical prediction with GA experiments. Crossover is uniform crossover with crossover rate $= 1$. Mutation rate is $p_m = 0.02$ for each bit. Selection is roulette selection. The length of string is $L = 20$, and population size $N = 100$. We averaged the results obtained by repeating 10000 calculations. The initial state is randomly generated with $p^{(1)} = 1/\ell$.

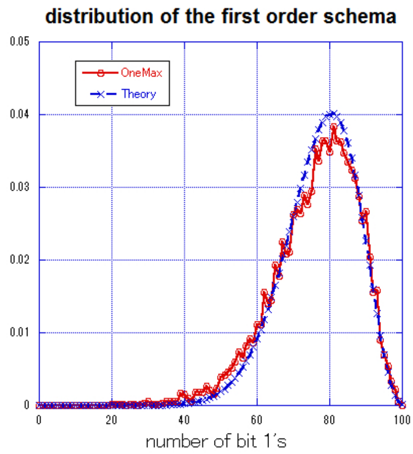


Fig. 4. Frequencies of the first order schema

Figure 4 is frequencies of the first order schema calculated by the theoretical prediction and GA experiments.

4.3 Total Variation Distance

The total variation distance between the stationary distribution and the first order schema at generation t is defined by

$$TV(t) = \frac{1}{2} \sum_{i=0}^{N} |\mu_i(t) - \pi_i|. \tag{24}$$

Figure 5 is the total variation distance TV between the stationary distribution and the first order schema. We found that the first order schema converges to the stationary distribution π.

5 APPEARANCE OF OPTIMUM SOLUTION

5.1 Theory

The distribution of appearance time of optimum solution at generation t after the convergence time T_c is

$$\phi(t) = \exp(-\lambda_m t). \tag{25}$$

Here, t is time after T_c.

Markov chain theory for the steady state predicts that the death rate λ_m is

$$\lambda_m = 1 - (1 - p_m)^L. \tag{26}$$

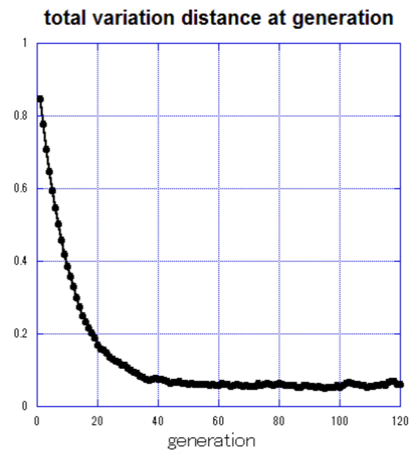


Fig. 5. Convergence of the first order schema

5.2 Experiment

We carried out experiments by setting parameters of experiment 4.2. We examined the distribution of the emergence of the optimum solution and the convergence time T_c of the stationary distribution. Figure 6 is the distribution of the hitting time for optimum solution after convergence time T_c and the theoretical prediction.

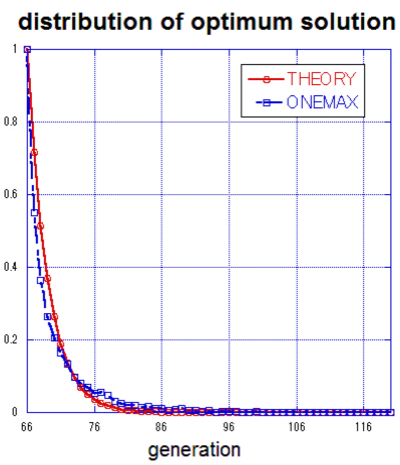


Fig. 6. Hitting time

If the mutation rate is less than 0.02, we found that the experimental results is less than the theoretical prediction.

6 SUMMARY

We studied the distribution of the hitting time of the optimum solution. To this end, we separated the task to two parts. One is the process to converge the chain into stationary state, another is the appearance of the optimum solution in a population. We examined the distribution of the optimum solution after convergence of Markov chain. We found that the appearance of the optimal solution can be predicted by using mutation rate p_m.

REFERENCES

[1] J. W. J. Ewens (2004), *Mathematical Population Genetics. I. Theoretical Introduction*, Second Edition. Springer-Verlag, New York.

[2] H. Furutani, S. Katayama, M. Sakamoto, M. Ito,(2007), Stochastic Analysis of Schema Distribution in a Multiplicative Landscape, Artificial Life and Robotics 11:101–104

[3] H. Furutani (2003), Schema Analysis of OneMax Problem –Evolution Equation for First Order Schemata. in Foundations of Genetic Algorithms 7, Morgan Kaufmann, San Francisco, 9–26

[4] J. F. Crow and M. Kimura (1970), *An Introduction to Population Genetics Theory*, Harper and Row, New York.

[5] W. Y. Tan (2002), *Stochastic Models with Applications to Genetics, Cancers, AIDS and Other Biomedical Systems*, World Scientific, Singapore.

A Model for Low-Frequency Bursts in Subthalamic Nucleus Neurons

Shigeru Kubota*[1], Jonathan E. Rubin*[2]

[1]Graduate School of Science and Engineering, Yamagata University,
4-3-16 Jonan, Yonezawa, Yamagata, 992-8510, Japan
[2]Department of Mathematics, University of Pittsburgh,
301 Thackeray Hall, Pittsburgh, PA 15260, USA
(Tel: +81-238-26-3766, Fax: +81-238-26-3299)

kubota@yz.yamagata-u.ac.jp

Abstract: Burst firing in subthalamic nucleus (STN) has been suggested to be highly associated with the motor symptoms in Parkinson's disease, which result from a loss of dopamine. Although it is clinically very important to clarify the mechanism underlying the bursting dynamics, complex interactions between STN and other brain areas make it difficult to understand. In anesthetized rats, STN neurons exhibit low-frequency (~1 Hz) bursts, which are synchronous with cortical slow oscillations and are significantly strengthened by dopamine depletion. To reproduce these low-frequency bursts, we examine a conductance-based model of an STN neuron that includes NMDA-type glutamatergic inputs reflecting cortical oscillations. In addition, the neuron model contains GABAergic inhibitory inputs, which are assumed to result from the activities of globus pallidus (GP). We show that the STN neuron model can reproduce low-frequency bursts synchronized with cortical activity, in the presence of GABAergic inhibition. In addition, we demonstrate that increased GABA activity leads to enhanced burstiness whereas increased NMDA conductances mainly augment STN firing rate. The induction of burst firing additionally decreases the coherence between STN and cortical activities. These results may give insights into how the complicated interactions between the STN, cortex, and GP can modulate the dynamics of bursting oscillations in the basal ganglia.

Keywords: Subthalamic nucleus, cortex, globus pallidus, Parkinson's disease, burst oscillations

1 INTRODUCTION

The subthalamic nucleus (STN) is a key component of the basal ganglia, a group of subcortical brain nuclei involved in motor control [1]. Bursting oscillations in the STN are suggested to be highly associated with motor disabilities in Parkinson's disease [2, 3]. Although it is clinically important to clarify the mechanisms underlying bursting dynamics, complex interactions between STN and other brain regions complicate efforts to do so.

Experiments using anesthetized rats have shown that STN neurons exhibit low-frequency bursts that are synchronized with cortical slow-wave activity (~1 Hz) [4, 5]. Cortical ablation abolished these burst oscillations and produced nearly regular firing in STN. This observation implies that the cortical inputs to STN, which are transmitted through the hyperdirect corticosubthalamic pathway, play a critical role in maintaining the oscillatory component of STN activity. Furthermore, in the case of pharmacological dopamine depletion, a condition that mimics a parkinsonian state, the firing rate of STN neurons was significantly increased to more than twice its original value [5]. The loss of dopamine also evoked low-frequency oscillations in the GABAergic interneurons in globus pallidus (GP), which are reciprocally connected with STN [5]. These data may suggest that the low-frequency bursting in STN is considerably modulated by GABA inhibition originating from GP interneurons, in addition to oscillatory inputs from the cortex.

In this study, we simulate a model STN neuron to study how STN bursts depend on cortical inputs, mediated by NMDA currents, and GABAergic inhibition. In the model, synaptic NMDA currents are activated by slow cortical oscillations consisting of active phases of high-frequency activity alternating with inactive quiescent phases, as observed in experiments [6-8]. We examine the changes in burst properties that arise with changes in the magnitudes of both the NMDA and GABA conductances. We show that enhanced GABAergic inhibition leads to stronger burstiness while enhanced NMDA activation mainly acts to increase STN firing rate. In addition, the occurrence of strong bursts is correlated with a decrease in coherence between the STN and cortical activities. These results may provide insights on how the complex interaction between STN and other associated brain areas modulates dynamical properties of bursting oscillations in the basal ganglia.

2 METHODS

2.1 STN neuron model

We constructed a conductance-based model STN neuron including synaptic NMDA and GABA currents. The

NMDA conductance is stochastically activated by the arrival of cortical inputs, while the GABA conductance has a temporally constant magnitude. The membrane potential V of the model neuron is described as

$$C\frac{dV}{dt} = -I_{leak} - I_{Na} - I_K - I_T - I_{Ca} - I_{AHP}$$
$$-I_{DIC} - I_{NMDA} - I_{GABA}, \quad (1)$$

where I_{leak} is a leak current; I_{Na} and I_K are voltage-gated Na^+ and K^+ currents, respectively, which are required for action potential generation; I_T is a low threshold T-type Ca^{2+} current; I_{Ca} is a high-threshold Ca^{2+} current; and I_{AHP} is a Ca^{2+}-activated K^+ current. I_{DIC} is a depolarization-activated inward current (DIC), which is activated by the intracellular Ca^{2+} entering through NMDA receptors (NMDARs) [9]. I_{NMDA} and I_{GABA} are the NMDA and GABA currents, respectively. The active currents other than I_{NMDA} and I_{GABA} were taken from previous models of an STN neuron [9, 10]. The NMDA current is described as

$$I_{NMDA} = g_N m_N h_N (V - V_{NMDA}), \quad (2)$$

where g_N is the peak conductance, V_{NMDA} is the associated reversal potential, and m_N and h_N are the activation and inactivation variables, respectively. The variable m_N is governed by a pair of first-order equations as follows [11]:

$$\frac{dm_N}{dt} = \alpha_m^N x_N (1 - m_N) - \frac{m_N}{\tau_m^N}, \quad (3)$$

$$\frac{dx_N}{dt} = \alpha_x^N (1 - x_N) \sum_j \delta(t - t_j^{exc}) - \frac{x_N}{\tau_x^N}. \quad (4)$$

Here, t_j^{exc} denotes the arrival time of the j th cortical input (see below). τ_m^N and τ_x^N are the time constants to regulate the decay and rise time constants of NMDA activation. The GABA current follows

$$I_{GABA} = g_G (V - V_{GABA}), \quad (5)$$

where g_G is the conductance and V_G is the reversal potential. The parameters used in the model are selected to reproduce various firing characteristic of STN neurons and the voltage-dependent properties of active currents [9, 10].

2.2 Slow cortical oscillation

Experimental findings show that during sleep and deep anesthesia, cortical neurons exhibit low-frequency oscillations (~1 Hz) with each cycle composed of two phases [6-8]. In one phase, which we refer to as the active phase, the depolarization of membrane potential induces high-frequency spiking. In the subsequent phase, which we refer to as the inactive phase, hyperpolarization rapidly suppresses action potential generation. To model the abrupt change between the two phases, we divided time into the intervals corresponding to the active and inactive phases, occurring alternately. We define T_A and T_I to be the time intervals for the active and inactive phases, respectively. Then, the values of T_A and T_I are determined as follows:

$$T_A = \frac{x_A}{f_c}(1 + 0.2 y_A), \quad (6)$$

$$T_I = \frac{(1 - x_A)}{f_c}(1 + 0.2 y_I), \quad (7)$$

where x_A = 0.4 is the fraction of each cortical period corresponding to the active phase, and f_c = 0.7 Hz is the mean frequency of the slow cortical oscillations [5]. y_A and y_I are random variables obtained independently for each cycle from a standard Gaussian distribution.

We also assumed that NMDA activation times (t_j^{exc} in Eq. 4) are determined by an inhomogeneous Poisson point process with phase-dependent (firing) rates. The frequency r of the NMDA activation is described as

$$r = \begin{cases} r_A(1 + 0.2 y_{r,1}), & \text{(active phase)} \\ r_I(1 + 0.2 y_{r,2}), & \text{(inactive phase)} \end{cases} \quad (8)$$

where $y_{r,1}$ and $y_{r,2}$ are random variables taken from Gaussian distribution. r_A = 30 Hz and r_I = 3 Hz are the mean activation frequencies for the active and inactive phases, respectively.

3 RESULTS

We examined the firing dynamics of the STN neuron model driven by low-frequency cortical oscillatory activity. As shown in Fig. 1, burst firing occurred in the presence of GABAergic inhibition, which can produce slow membrane oscillations underlying bursting [1,9,10]. Higher STN firing rates were found around the times when the active phases of the cortical oscillations occurred (Fig. 1). This result agrees well with the observation that STN bursts occur synchronously with slow cortical oscillations [4, 5].

The increased activation frequency of cortical inputs evoked NMDA activation (Figs. 1A and 1B). While the overall high levels of NMDA activation presumably helped support burst firing [9], STN burst onset occurred near the trough of NMDA activation, which built up throughout the period of cortical and STN bursting (Fig. 1C). That is, since NMDA current evolves slowly, STN activity during bursts

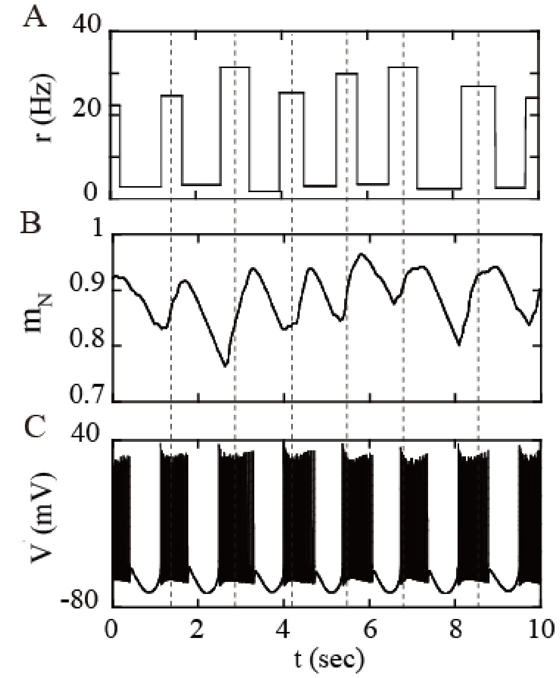


Fig. 1 An example of the response of the STN neuron model for g_N = 30 and g_G = 2 mS/cm^2. The time courses of the firing rate r of cortical inputs (A), the activation variable m_N of the NMDA current (B), and the membrane voltage V (C). The dashed lines show the center of the time period of the active phase of cortical oscillations.

in the model did not depend on the precise timing of cortical firing.

We also investigated the effects of changing the conductances of GABA and NMDA currents (Fig. 2). Enhancing GABA conductance considerably increased the coefficient of variation (CV) of interspike intervals (ISIs) (Fig. 2B), suggesting that GABAergic inhibition plays a key role in the strengthening of burstiness in our model. On the other hand, the increase in NMDA conductances was found to mainly augment the firing rate (Fig. 2A). Interestingly, the induction of burstiness by turning on GABAergic inhibition also tended to decrease the coherence between the cortical inputs and STN firing (Fig. 2C). This result indicates that the enhanced burst firing may prevent the STN neuron from abruptly tracking the change in the cortical activity phases. The depletion of dopamine, a key feature of Parkinson’s disease, is known to enhance GABA-mediated outward currents in STN [12]. Based on our results, in the case of dopamine depletion, higher levels of GABA inhibition could suppress the ability of STN to track cortical activity, which may be harmful for coordinating activities between the cortex and basal ganglia.

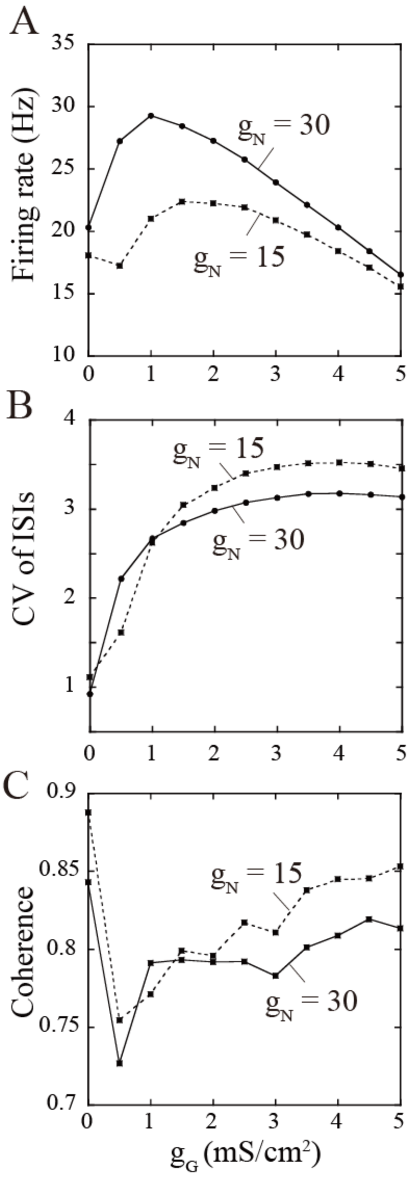


Fig. 2 Changes in the firing rate (A), the coefficient of variation (CV) of interspike intervals (ISIs) (B), and the coherence between the STN and cortical activities at 0.7 Hz (C) as function of GABA conductance g_G. The NMDA peak conductance was g_N = 15 (dotted) or 30 (solid) mS/cm^2.

4 CONCLUSIONS

In this study, we constructed a model of an STN neuron influenced by cortical slow-wave activity, which reproduced bursting oscillations. The bursting in the model was found to depend not only on NMDA associated with cortical inputs but also on GABAergic inhibition. Because NMDA is a slow current, STN firing times during bursts were not strongly linked to cortical firing times. Indeed, we observed that STN firing during bursts was less coherent than during regular firing (in the absence of

GABAergic inputs) and became more irregular with increases in GABA conductance (Fig 2B). These results may suggest a useful framework for how the complex interactions among the STN, cortex and GP shape network activity patterns, particularly in conditions of dopamine depletion when oscillations and bursting are enhanced.

REFERENCES

[1] Bevan MD, Magill PJ, Terman D, Bolam JP, Wilson CJ (2002), Move to the rhythm: oscillations in the subthalamic nucleus-external globus pallidus network. Trends in Neuroscience, 25:525–531

[2] Brown P, Oliviero A, Mazzone P, Insola A, Tonali P, Di Lazzaro V (2001), Dopamine dependency of oscillations between subthalamic nucleus and pallidum in Parkinson's disease, Journal of Neuroscience, 21: 1033-1038

[3] Levy R, Lang AE, Dostrovsky JO, Pahapill P, Romas J, Saint-Cyr J, Hutchison WD, Lozano AM (2001), Lidocaine and muscimol injections in subthalamic nucleus reverse parkinsonian symptoms, Brain 124: 2105-2118

[4] Magill PJ, Bolam JP, Bevan MD (2000), Relationship of activity in the subthalamic nucleus-globus pallidus network to cortical electroencephalogram, Journal of Neuroscience, 20: 820-833

[5] Magill PJ, Bolam JP, Bevan MD (2001), Dopamine regulates the impact of the cerebral cortex on the subthalamic nucleus-globus pallidus network, Neuroscience, 106: 313-330

[6] Amazica F, Steriade M (1995), Short- and long-range neuronal synchronization of the slow (<1 Hz) cortical oscillation, Journal of Neurophysiology, 73: 20-38

[7] Steriade M, Amzica F, Contreras D (1996), Synchronization of fast (30-40 Hz) spontaneous cortical rhythms during brain activation, Journal of Neuroscience, 16: 392-417

[8] Steriade M, Contreras D, Amzica F, Timofeev I (1996), Synchronization of fast (30-40 Hz) spontaneous oscillations in intrathalamic and thalamocortical networks, Journal of Neuroscience, 16: 2788-2808

[9] Kubota S, Rubin JE (2011), NMDA-induced burst firing in a model subthalamic nucleus neuron, Journal of Neurophysiology, 106: 527-537

[10] Terman D, Rubin JE, Yew AC, Wilson CJ (2002), activity patterns in a model for the subthalamopallidal network of the basal ganglia, Journal of Neuroscience, 22: 2963-2976

[11] Wang XJ (1999), Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory, Journal of Neuroscience, 19: 9587-9603

[12] Shen KZ, Johnson SW (2005), Dopamine depletion alters responses to glutamate and GABA in the rat subthalamic nucleus, Neuroreport, 16: 171-174

Finite-time Stabilization for Nonholonomic Chained Form Systems with Communication Delay

Hengjun Zhang and Chaoli Wang

Department of Control Science and Engineering, University of Shanghai for Science and Technology, Shanghai, 200093, PRC
(E-mail: rory110@163.com; clclwang@163.com)

Abstract: In this paper, the problem of finite-time stabilization is developed for nonholonomic chained form systems with communication delay in the input. The finite-time control laws are presented by utilizing the switching control strategy and the theory of finite-time stability, which can make the states of the nonholonomic chained form systems to converge from any non-equilibrium state to the equilibrium or a given point in a finite time. Finally, the simulation results show the effectiveness of the proposed control approach.

Keywords: Finite-time stabilization, nonholonomic chained systems, communication delay.

I. INTRODUCTION

The control and stabilization of nonholonomic systems have received considerable attention during the last few years. Nonholonomic systems with non-holonomic or non-integrable constraints are frequently used to describe the practical control systems. The representative examples of such systems are mobile robots, car-like vehicles, n-level trailer systems and so on. See the survey paper [1] for more details and references therein for more examples. So it is important from the view point of applications to study nonholonomic control systems. From a theorem due to Brockett [2], the nonholonomic system cannot be stabilized at an equilibrium point by pure smooth (or even continuous) state feedback controller.

Therefore, the stabilization problem of non-holonomic systems is still extremely challenging which has attracted a continuously increasing attention in the control community. To overcome this problem, many methods have been proposed, such as continuous time-varying feedback control laws [3], hybrid feedback control laws [4] and discontinuous feedback control laws [5] and so on. Moreover, the robust stabilization problem has been widely researched by using these valid approaches.

However, most models considered are in ideal cases without delays. From a practical point of view, delays in a control system can create significant obstacles to the stabilization problem and potentially degrade the performance of the closed-loop system. It is well known that guaranteeing the stability of a control system with time delay is a challenging problem [6]. The communication delay in the input is related to processing and connecting time for the packets arriving and which also occur when actuators and controllers are connected by networks. So it is very important to study the delay effect on nonholonomic systems. In [7], the classical set-point control problem is studied for rigid robots with input-output communication delays in the closed-loop system. The consensus problem for multi-agent systems with input and communication delays is studied based on the frequency-domain analysis in [8]. And the literature [9] presents a control strategy for stabilization of nonholonomic control systems with strongly nonlinear uncertainties and time delay.

In order to get the better convergence speed and higher steady state precision, many researchers have published literature on the finite-time stabilization. And three kinds of global finite-time feedback stabilization methods have been presented as follow: homogeneous properties, backstepping constructive approach and sliding mode technology. A finite-time controller is proposed in [10] for a class of high-order nonlinear by applying homogeneous properties of systems and backstepping constructive approach. The literature [11] considers the finite-time stability of homogeneous systems and gives Lyapunov stability theory of finite-time. Recently, based on backstepping constructive approach, finite-time stabilization is discussed in [12] for a class uncertain nonlinear systems. And the sliding mode technology is investigated in [13]. The distributed control laws in [14] are designed such that multiple non-holonomic mobile agents can be all in agreement within finite time. In [15], the problem of finite-time stability and stabilization for retarded-type functional differential equations is considered.

In this paper, we introduce a class of chained form nonholonomic systems with communication delay in the input and then study the problem of finite-time stabilization for the concerned nonholonomic systems. A global finite-time stabilization controller is presented by using switching control strategy, and making that the all states converge to a given point or the equilibrium point in a finite time.

The paper is organized as follows. In Section II, we will give a formalization of the problem considered in this article and some preliminary lemmas. Section III states our main results including controller design and stability analysis. In Section IV, the simulation results carried out to validate the theoretical framework.

Finally, the conclusions of the paper are given in Section V.

II. PROBLEM STATEMENT

A class of chained form nonholonomic systems with communication delay in the input to be considered in this paper is described by

$$\begin{cases} \dot{x}_0 = u_1(t-\tau_1), \\ \dot{x}_i = x_{i+1}u_1(t-\tau_1), \quad i=1,\cdots,n-1 \\ \dot{x}_n = u_2(t-\tau_2), \end{cases} \tag{1}$$

where $x_i(i=0,1,\cdots,n)$ and $u_j(j=1,2)$ are the state and input of systems respectively, and $u_j = u_j(t-\tau_j)$ with both $\tau_j(j=1,2)$ being positive constants.

For convenience, some symbols are given in the following at first:

$$sig^{\alpha}(x) = |x|^{\alpha} sign(x),$$

where $\alpha > 0$, $|x|$ denotes the absolute value of real number x and $sign$ denotes a sign function.

Next, we will introduce the following lemmas that are needed for our controllers later.

Lemma 2.1[11] Consider the following nonlinear system

$$\dot{y}(t) = f(y(t)) \tag{2}$$

where $V:\mathcal{D} \to R^n$ is continuous on an open neighborhood of origin and $f(0)=0$. The equilibrium $x=0$ of the system is finite-time stable if:

(i) it is asymptotically stable, in $\mathcal{N}$, an open neighborhood of origin ,with $\mathcal{N} \subseteq \mathcal{D}$;

(ii) it is finite-time convergent in $\mathcal{N}$, that is , for any initial condition $x_0 \in \mathcal{N}\setminus\{0\}$, there is a settling time $T>0$ such that every solution $x(t,x_0)$ of system (2) is defined with $x(t,x_0) \in \mathcal{N}\setminus\{0\}$ for $t\in[0,T)$ and satisfies $\lim_{x\to T} x(t,x_0)=0$, and $x(t,x_0)=0$, if $t \ge T$.

Moreover, if $\mathcal{D}=\mathcal{N}=R^n$, the origin is globally finite-time stable equilibrium.

We consider the system

$$\dot{x}(t) = Ax(t) + \sum_{i=0}^{k} B_i u(t-\tau_i), \qquad t \ge 0, \tag{3}$$

where $x(t)\in R^n$, $u(t)\in R^m$, $A \in R^{n\times n}$, $B_i \in R^{n\times m}$ and h_i are positive constants. According to the paper [15], we know that if let

$$y(t) = x(t) + \sum_{i=0}^{k} L^{\tau_i}_{(A,B_i)} u_t, \tag{4}$$

where $u_t(s) = u(t+s):[-\tau_i,0] \to R^m$ for every $s\in[-\tau_i,0]$ and

$$L^{\tau_i}_{(A,B_i)} u_t = \int_{-\tau_i}^{0} e^{A(-\tau_i - s)} B_i u(t+s)ds,$$

then from (2) and (3), we can have

$$\dot{y}(t) = Ay(t) + Bu(t), \tag{5}$$

with $B = \sum_{i=0}^{k} e^{-A\tau_i} B_i$.

Lemma 2.2 If the system (5) is finite-time stable by a feedback control law

$$u(t) = k(t) f(y(t)), \tag{6}$$

with $k(t)$ bounded and $f:R^n \to R^n$ continuous such that $f(0)=0$ and there exists a function α of class κ such that $\| f(x)\|_m \le \alpha(\|x\|_n)$, then the system (3) is finite-time stable by the feedback controller

$$u(t) = k(t) f\left(x(t) + \sum_{i=0}^{k} L^{\tau_i}_{(A,B_i)} u_t \right). \tag{7}$$

Now we consider the finite-time stabilization of the integrator systems with delay in the input described by

$$\begin{cases} \dot{x}_i(t) = x_{i+1}(t), \quad i=1,\cdots,n-1 \\ \dot{x}_n(t) = u(t-\tau), \end{cases} \tag{8}$$

which can be denoted in short by

$$\dot{x}(t) = Ax(t) + Bu(t-\tau)$$

with

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & \vdots \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

According to the Lemma 2.2 and the literature [15], it is not difficult for us to get the following lemma.

Lemma 2.3 Let $k=(k_1,\cdots,k_n)\in R^n$ such that the polynomial $s^n + k_n s^{n-1} + \cdots + k_2 s + k_1$ is Hurwitz, there exists $\varepsilon \in (0,1)$ such that for all $\alpha \in (1-\varepsilon,1)$, the system (8) is globally finite-time stable under the continuous feedback control law

$$u_{\alpha} = -k_1 sig^{\alpha_1}(y_1) - \cdots - k_n sig^{\alpha_n}(y_n), \tag{9}$$

where

$$sig^{\alpha}(y) = |y|^{\alpha} sign(y), \quad y(t) = x(t) + L^{\tau}_{(A,\bar{B})} u_t,$$

and $\bar{B} = e^{-A\tau} B$, $\alpha_1,\cdots,\alpha_n$ satisfy

$$\begin{cases}\alpha_{i-1} = \alpha_i \alpha_{i+1} / (2\alpha_{i+1} - \alpha_i), & 2 \le i \le n \\ \alpha_n = \alpha, \quad \alpha_{n+1} = 1. \end{cases} \tag{10}$$

III. CONTROLLER DESIGN

In this section, the main results will be presented. Based on the switching control strategy, the finite-time controllers u_1 and u_2 are designed to make that all states of the systems (1) can converge to the equilibrium point in a finite time. To solve the finite-time stabilization problem, we divide the systems (1) into a first-order subsystem

$$\dot{x}_0 = u(t - \tau_1), \tag{11}$$

and a n-order subsystem

$$\begin{cases}\dot{x}_i = x_{i+1} u_1(t - \tau_1), & i = 1, \cdots, n-1 \\ \dot{x}_n = u_2(t - \tau_2). \end{cases} \tag{12}$$

Theorem 3.1 The systems (1) can be finite-time stabilizable if let $k = (k_1, \cdots, k_n) \in R^n$ such that the polynomial $s^n + k_n s^{n-1} + \cdots + k_2 s + k_1$ is Hurwitz. And choosing the proper parameters $c > 0$ and $\alpha_0, \alpha_1, \cdots, \alpha_n \in (0,1)$, then take the following switching control laws:

When $t < T_1 + \tau_2$, we have

$$\begin{cases} u_1 = c, \\ u_2 = -k_1 sig^{\alpha_1}(y_1(t)) - \cdots - k_n sig^{\alpha_n}(y_n(t)), \end{cases} \tag{13}$$

where

$$y_i(t) = c^{i-1} x_i(t) + L^{\tau_2}_{(A,\bar{B})} c^{n-1} u_2, \quad \bar{B} = e^{-A\tau_2} B,$$

with the matrix A and B are same as in the Lemma 2.3.

When $t \ge T_1 + \tau_2$, we also have

$$\begin{cases} u_1 = -sig^{\alpha_0}(x_0)\left(1 + x_0(t - \tau_1)^2\right), \\ u_2 = 0. \end{cases} \tag{14}$$

Proof: When $t < T_1 + \tau_2$, because of $u_1 = c$, so the n-order subsystem (12) can be written as

$$\begin{cases}\dot{x}_i = c x_{i+1}, & i = 1, \cdots, n-1 \\ \dot{x}_n = u_2(t - \tau_2). \end{cases} \tag{15}$$

Consider the following transformation

$$\begin{cases}\tilde{x}_i = c^{i-1} x_i, & i = 1, \cdots, n \\ \tilde{u}_2 = c^{n-1} u_2. \end{cases} \tag{16}$$

With the help of transformation (16), the subsystem (15) can be expressed as

$$\begin{cases}\dot{\tilde{x}}_i = \tilde{x}_i, & i = 1, \cdots, n \\ \dot{\tilde{x}}_n = \tilde{u}_2(t - \tau_2). \end{cases} \tag{17}$$

Based on Lemma 2.3, the control law

$$\tilde{u}_2 = -k_1 sig^{\alpha_1}(\tilde{y}_1(t)) - \cdots - k_n sig^{\alpha_n}(\tilde{y}_n(t)), \tag{18}$$

where

$$\tilde{y}_i(t) = \tilde{x}_i(t) + L^{\tau_2}_{(A,\bar{B})} \tilde{u}_2, \quad \bar{B} = e^{-A\tau_2} B, \tag{19}$$

the A and B are the same as in the Lemma 2.3, and $\alpha_1, \cdots, \alpha_n$ satisfy the conditions (10). It can render the subsystem (17) globally finite-time stable. Substituting the transformation (16) into the control law (18), the states $x_1, x_2, \cdots, x_n$ and u_2 can converge to the equilibrium point in a finite time $T_1 + \tau_2$ via the control law (13).

When $t \ge T_1 + \tau_2$, the states $x_1, x_2, \cdots, x_n$ have reached on the equilibrium point, thus, we only consider the first order subsystem (11).

Construct a Lyapunov function

$$V(x_0(t)) = \frac{1}{2} x_0(t)^2. \tag{20}$$

Taking the time derivative of $V(x_0)$ along (11)

$$\begin{aligned} \dot{V} &= -|x_0|^{1+\alpha_0}\left(1 + x_0(t - \tau_1)^2\right) \\ &\le -|x_0|^{1+\alpha_0} = -2^{(1+\alpha_0)/2} V^{(1+\alpha_0)/2}. \end{aligned} \tag{21}$$

So the subsystem (11) can be stabilized within the settling time $T_2 \le |x_0(t)|^{1-\alpha_0} / (1 - \alpha_0)$ via the switching control law (14). Hence, the states $x_0, x_1, \cdots, x_n$, u_1, and u_2 can converge to the equilibrium point after time $T = T_1 + T_2 + \tau_1 + \tau_2$. This completes the proof of Theorem 3.1.

IV. SIMULATION

In this section, we will give an illustrative example of Theorem 3.1. Consider the following nonholonomic chained form systems with communication delay

$$\begin{cases}\dot{x}_0 = u_1(t - \tau_1), \\ \dot{x}_1 = x_2 u_1(t - \tau_1), \\ \dot{x}_2 = u_2(t - \tau_2). \end{cases} \tag{22}$$

Choosing $\tau_1 = \tau_2 = 0.5$ and the initial values of system (22) are $(x_0, x_1, x_2) = (-2, -1, 2)$. Selecting $c = 1$, $k_1 = 2$, $k_2 = 3$ and $\alpha_0 = 0.5$, $\alpha_2 = 0.4$, by the formula (10), we have $\alpha_1 = 0.25$.

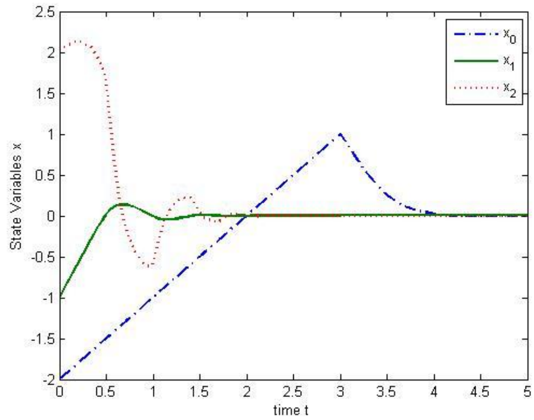


Fig.1. The state response of system (22)

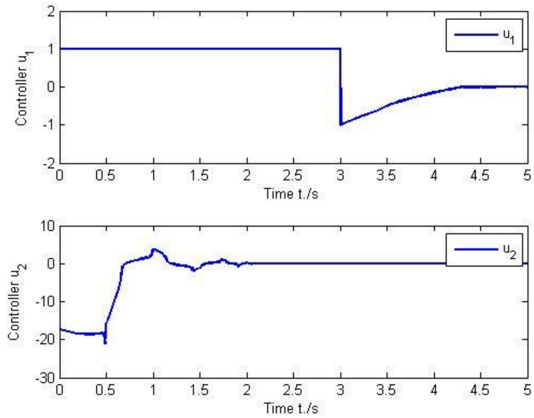


Fig.2. The response of controllers u_1 and u_2

From Fig. 1, it is easy to see that the state variables of systems can be driven to the equilibrium point in a finite time. As shown in Fig. 2, the control inputs u_1 and u_2 can also converge to the origin.

V. CONCLUSION

In this paper, a globally finite-time stabilization problem is presented for nonholonomic chained form systems with communication delays in the input. By utilizing finite-time stability theory and switching control strategy, the nonholonomic chained systems can be divided into different subsystems, then the controllers of subsystems are given with different methods such that each state of the systems can converge to the equilibrium point in a finite time. Finally, a simulation example is provided to verify that the designed controllers are effective.

ACKNOWLEDGMENT

This paper was partially supported by The National Natural Science Foundation (61374040); Key Discipline of Shanghai (S30501); Scientific Innovation program (13ZZ115); Graduate Innovation program of Shanghai (54-13-302-102).

REFERENCES

[1] I. Kolmanovsky, N. H. McClamroch (1995), Developments in nonholonomic control systems. IEEE Contr. Syst. Mag. 15(6):20-36.

[2] R. W. Brockett (1983), Asymptotic stability and feedback stabilization. in Proc. R. W. Brockett, R. S. Millman, H. J. Sussmann (Eds.), Differential Geometric Control Theory, Birkhauser, Boston. 181-208.

[3] Y. P. Tian, S. Li (2002), Exponential stabilization of nonholonomic dynamic systems by smooth time-varying control. Automatica. 38(7):1139-1146.

[4] A. P. Aguiar, A. Pascoal (2000), Stabilization of the extended nonholonomic double integrator via logic based hybridcontrol. in Proc. of 6th international IFAC symposium on robot control, ViennaCAustria. 208-213.

[5] M. C. Laiou, A. Astolfi (1999), Discontinuous control of high order generalized chained systems. Systems Control Letters. 37:309-322.

[6] J. P. Richard (2003), Time-delay systems: An overview of some recent advancesand open problems. Automatica. 39(10):1667-1694.

[7] Y. C. Liu, N. Chopra (2012), Control of robotic manipulators under input/output communication delays: theory and experiments. IEEE Transactions on Robotics. 28(3):742-751.

[8] Y. P, Tian, C. L. Liu (2008), Consensus of multi-agent systems with diverse input and communication delays. IEEE Transactions on Automatic Control. 53(9):2122-2128.

[9] Y. Y. Wu, Y. Q. Wu (2010), Robust stabilization of delayed nonholonomic systems with strong nonlinear drifts. Nonlinear Analysis: Real World Applications. 11(5):3620-3627.

[10] Y. G. Hong (2001), Finite-time stabilization and stabilizability of a class of controllable systems. Systems Control Letters. 46(4):231-236.

[11] S. P. Bhat, D. S. Bernstein (1997), Finite-time stability of homogeneous systems. in Proc. of the 1997 American Control Conference. Albuquerque. New Mexico: American Automatic Control Council. 2513-2514.

[12] X. Q. Huang, W. Lin, B. Yang (2005), Global finite-time stabilization of a class uncertain nonlinear systems. Automatica. 41(5):881-888.

[13] Y. G. Hong, G. W. Yang, D. Z. Cheng, et al. (2004), Finite- time convergent control using terminal sliding mode. Journal of Control Theory and Applications. 2(1):69-74.

[14] J. K. Wang, Z. H. Qiu, G. S. Zhang et al. (2012), Finite time consensus for multiple non-holonomic agents with communication delay. in Proc. of the 31st Chinese Control Conference. 6267-6272.

[15] E. Moulaya, M. Dambrineb, N. Yeganefarc et al. (2008), Finite-time stability and stabilization of time-delay systems. Systems Control Letters. 57(7):561-566.

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Sliding Mode Variable Structure Control for Magnetic Levitation Vehicles

Juanjuan He and Yingmin Jia

The Seventh Research Division and the Department of Systems and Control, Beihang University (BUAA), Beijing 100191, China
(E-mails: hejuanjuan@smss.buaa.edu.cn; ymjia@buaa.edu.cn).

Abstract: This paper focuses on stability control for the levitated positioning of the magnetic levitation vehicle system. For the nonlinear magnetic levitation system model, the output feedback linearization method is first employed to derive a global linearization error model. However, there exists uncertain item in the error model. To stabilize this error model, the adaptive sliding mode variable structure control method is used. Simulations show that the magnetic levitation system can be stability and track the desired signals quickly under the proposed control scheme.

Keywords: adaptive sliding mode variable structure control, magnetic levitation vehicle system, nonlinear model, output feedback linearization

1 INTRODUCTION

As the smart transportation establishing becomes a hot research issue in recent years, how to improve the operating efficiency and convenience of the entire city is becoming increasingly important. Therefore, it is necessary to establish a new high-speed and efficient means of transport. Up to 500km/h of speed, the maglev train can greatly improve the transport capability. Furthermore, magnetic levitation technology has seen a rapid development since the 20th century, and made the interesting in the field of real-life applications [1], such as transportation systems [2,3], wind tunnel levitation [4], magnetic bearing systems and antivibration table. However, the design of maglev system controller still presents plenty of formidable challenges [5] including the nonlinear solenoid model, instability, and inevitably uncertain parameters. Hence, a proper controlling mode should be designed for the suspension system to guarantee the robust stability of the entire system. For traditional methods, nonlinear magnetic levitation vehicle system model is linearized at one equilibrium point [6], which leads to the fact that the design of the controller seriously depends on the selected equilibrium point. So this type of linearization cannot be applied to the situation that the gap is a large range of variation. Thus if the traditional linearization method is applied to this situation, it may lead to the performance degradation or instability of the system. Hence, a new solution should be proposed.

In this paper, latest control method is applied in allusion to the magnetic levitation planner. Firstly, output feedback linearization technique is applied to the nonlinear magnetic levitation vehicle system model, which derives a global linearization error model. However, the simplified model is unstable and includes uncertain factors, so the adaptive sliding mode variable structure controller design method is used here, which makes the linearization error system stable. Finally, a series of simulation experiments are shown in this paper, which is used to demonstrate the effects of output feedback linearization method and adaptive sliding mode variable structure controller.

This paper is organized in five sections. In the next section, model analysis of magnetic levitation vehicle system is formulated, and output feedback linearization is proposed. Section 3 is related to the controller design. Simulation results are presented in section 4. Finally, in section 5, some concluding remarks are outlined.

2 MODEL ANALYSIS OF THE MAGNETIC LEVITATION VEHICLE SYSTEM

2.1 Nonlinear Solenoid Model Analysis

The schematic diagram of the magnetic levitation system is depicted in Fig.1. For the sake of achieving the goal of high-precision positioning, a complex magnetic levitation system vehicle model needs to be analyzed thoroughly. Before deriving the model, we make the reasonable assumption that the air-gap flux leakage is zero. Then, the idealized nonlinear solenoid model can be calculated as

$$F = \frac{\mu_0 N^2 A i^2(t)}{4x^2(t)} \tag{1}$$

where F is the electromagnetic force between the coil and suspension; the constant $\mu_0 = 4\pi \times 10^{-7} H/m$ is the absolute magnetic permeability; N is the coil turn; A is the magnetic cross section; $i(t)$ is the coil current, and $x(t)$ is the levitation height between coil and suspension.

The coil model is given by

$$u(t) = Ri(t) + \frac{d\psi(x,i)}{dt} \tag{2}$$

where u is control voltage, and R ,ψ are circuitous resistance and flux linkage respectively.

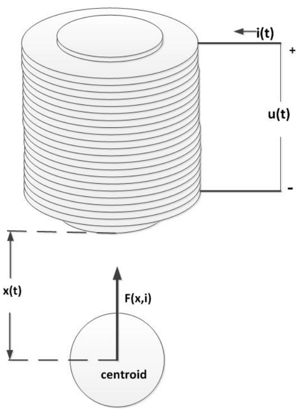


Fig. 1. Schematic diagram of magnetic levitation system

2.2 System Dynamics

The complete model of the maglev system can be derived by using $Newton's$ Law as

$$\sum \vec{F} = m\ddot{x}(t) \tag{3}$$

Firstly, to obtain the system dynamics, we substitute Eq.(1) and Eq.(2) into Eq.(3). As a result, the dynamic equations can be rewritten as

$$m\ddot{x}(t) = \frac{\mu_0 N^2 A i^2}{4x^2(t)} - mg \tag{4}$$

where m is mass of the suspension; g is acceleration of gravity; and A is magnetic area.

2.3 Analysis of Nonlinear Model

Let the states be chosen as $x_1 = x$, $x_2 = \dot{x}$, $x_3 = i$. u is the input voltage of the system changing from 0 to 5 volts, and $X = [x_1, x_2, x_3]^T$ is the state vector. Thus, the state-space model of the magnetic levitation vehicle system can be written as

$$\begin{aligned} \dot{X}(t) &= F(X) + G(X)U(t) \\ Y(t) &= x_1, \end{aligned} \tag{5}$$

where $S = \frac{\mu_0 A N^2}{2}$, and $F(X), G(X), U(t)$ are shown as follows

$$F(X) = \begin{bmatrix} x_2 \\ \frac{S}{2m}\left(\frac{x_3}{x_1}\right)^2 - g \\ \frac{x_2 x_3}{x_1} - \frac{2}{S} R x_1 x_3 \end{bmatrix}$$

$$G(x) = \begin{bmatrix} 0 \\ 0 \\ \frac{2}{S} x_1 \end{bmatrix}, U(t) = u(t).$$

2.4 Linearization of System Model

For the purpose of the afterward controller design, the feedback linearization is applied to the system model shown as Eq.(5) in this subsection. As a result, the third derivative of system outputs can be finally obtained as follows

$$\dddot{Y} = B(C + KU) \tag{6}$$

where $K = \frac{x_3}{x_1}$, $B = \frac{1}{m}$, and $C = -\frac{Rx_3^2}{x_1} - \frac{Sx_2x_3^2}{2x_1^3}$.

Obviously, when $x_1, x_3 \neq 0$, the $K \neq 0$. In physics system, $x_j,\ j \in (1,3)$, which are central levitation height of the suspension and coil current respectivelythus their product could not be zero, so the decoupling matrix K is invertible.

Thus, the decoupling control inputs can be obtained as

$$U = K^{-1}(U_c - C) \tag{7}$$

where $U_c = v$.

Then the linearized system model can be arrived at by applying Eq.(7) to Eq.(6) and finally obtained as follows

$$\dddot{Y} = BU_c = \frac{1}{m} v \tag{8}$$

3 ADAPTIVE SLIDING MODE CONTROLLER DESIGN

The error output vector is defined as $E = Y - Y_d$, y_D where is the desired position and y is the current position. Then, Eq.(8) can be rewritten as

$$\dddot{E} = \frac{1}{m} v - \dddot{Y}_d + W \tag{9}$$

where W denotes the system's uncertainty and is assumed bounded $\|W\| \leq P$. Here, we assume that P is unknown.

In order to design a controller which possesses a better ability to gain high robustness and self-tuning property, two advanced control methods have been integrated. In this paper, we will introduce the controller design and provide the stability analysis.

Based on the model Eq.(9), which is compactly re-expressed as

$$\begin{aligned} \dot{z}_1 &= \dot{E} \\ \dot{z}_2 &= \ddot{E} \\ \dot{z}_3 &= \dddot{E} = BU_c - \dddot{Y}_d + W \end{aligned} \tag{10}$$

Furthermore, we define sliding surface variable S as

$$S = z_3 + \Sigma_1 z_2 + \Sigma_2 z_1 \tag{11}$$

where Σ_1,Σ_2 are positive constants. In this paper, we try to regulate the state $Z = [z_1, z_2, z_3]^T$ to zero, in other words, the current output will achieve the goal of precise tracking of positioning. In the context of SMC, asymptotical convergence of the variable S to zero will apparently imply asymptotical convergence of Z as well to zero. To validate this, we will need to investigate the dynamics of the sliding surface variable S as follows

$$\begin{aligned} \dot{S} &= \dot{z}_3 + \Sigma_1 z_3 + \Sigma_2 z_2 \\ &= BU_c - \dddot{Y}_d + W + \Sigma_1 z_3 + \Sigma_2 z_2 \end{aligned} \tag{12}$$

In addition to the *SMC*, an adaptive controller is applied for estimating the parameters of the system online while simultaneously controlling the system. After we have the estimates of the system parameters, then these estimates can be adopted to the control command in Eq.(12) to form appropriate *SMC* with boundary layer as

$$U_c = B^{-1}[-k_1 S - \hat{P} sgn(S) + \ddot{Y}_d - \Sigma_1 z_3 - \Sigma_2 z_2] \tag{13}$$

where k_1 is a positive constant , $\hat{P}$ is the estimate of P, and $sgn(\cdot)$ is the symbolic function.

Thus, substituting Eq.(13) into Eq.(12), we obtain

$$\begin{aligned} \dot{S} &= (\Sigma_1 z_3 + \Sigma_2 z_2 - \ddot{Y}_d + W) + B[B^{-1}(-k_1 S \\ &\quad - \hat{P} sgn(S) + \ddot{Y}_d - \Sigma_1 z_3 - \Sigma_2 z_2)] \\ &= -k_1 S - \hat{P} sgn(S) + W \end{aligned} \tag{14}$$

By appropriate gains k_1, $\hat{P}$, Σ_1, and Σ_2, we can ensure the convergence of S. Hence, the state Z converges to zero, which means the tracking error E to zero, and meanwhile maintain the estimation error $\widetilde{P}$ converging to zero.

Stability analysis: we define a Lyapunov function candidate V, which is a positive definite function

$$V = \frac{1}{2} S^T S + \frac{1}{2} \frac{1}{\rho} \widetilde{P}^2 \tag{15}$$

where the estimation error is defined as $\widetilde{P} = P - \hat{P}$, $\hat{P}$ is the estimates of P , and ρ is positive.

The time derivative of the Lyapunov candidate function V can be found to be

$$\begin{aligned} \dot{V} &= S^T \dot{S} + \frac{1}{\rho} \widetilde{P} \dot{\widetilde{P}} \\ &= S^T \left[-k_1 S - \hat{P} sgn(S) + W\right] - \frac{1}{\rho} \widetilde{P} \dot{\hat{P}} \\ &\leq -S^T k_1 S - \hat{P} \|S\| + \|S\| P - \frac{1}{\rho} \widetilde{P} \dot{\hat{P}} \end{aligned} \tag{16}$$

Using the adaptive control theory to establish bounds of parameter estimates in the presence of modeling error terms, the adaptive laws is devised as

$$\dot{\widetilde{P}} = -\dot{\hat{P}} = -\rho \|S\| \tag{17}$$

After substituting Eq.(17) into Eq.(16), we can get

$$\begin{aligned} \dot{V} &\leq -S^T k_1 S - \hat{P} \|S\| + \|S\| P - \widetilde{P} \|S\| \\ &= -S^T k_1 S - \hat{P} \|S\| + \hat{P} \|S\| \\ &< -S^T k_1 S < 0 \end{aligned} \tag{18}$$

where $k_1 > 0$.

According to Lyapunov stability theory, the tracking error $E(t)$ will converge to zero. In other words, it obviously achieves the goal of precise tracking of positioning.

4 SIMULATION RESULTS

In this section, a number of typical simulation results are presented, including the transient and the steady-state responses in different situations. The simulation results are provided to demonstrate the performance of the developed magnetic levitation system with controller presented in section 3. Based on these results, we will make some conclusions which are important for the future work in this research.

In order to demonstrate the controlling performance of the magnetic levitation vehicle system more effectively, the block diagram of the system as shown in Fig.2 is constructed in this section. Base on the simulation block diagram, serials experiments are performed.

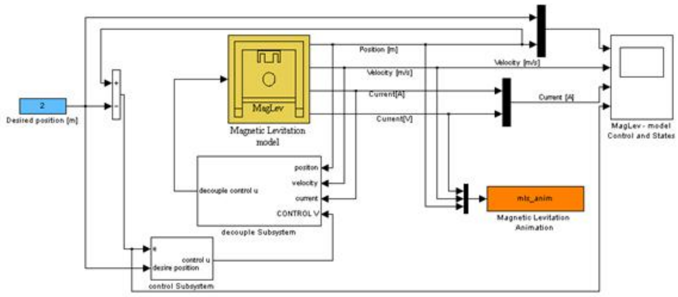


Fig. 2. Block diagram of magnetic levitation system

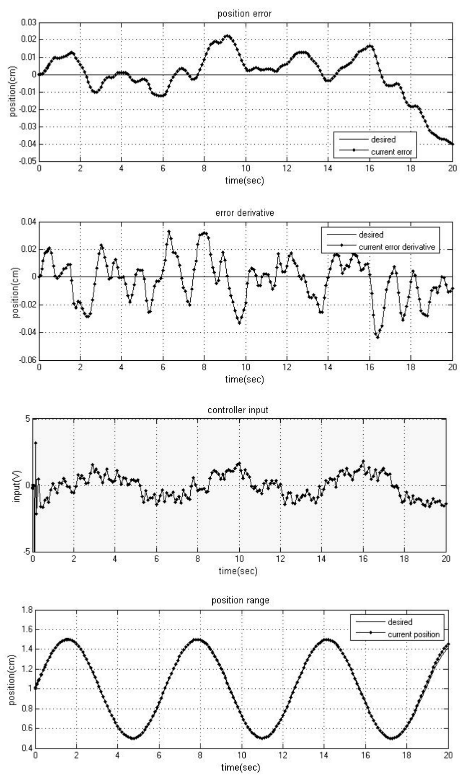


Fig. 3. Tracking response performance of central levitation height along the x-axis with uncertainty bound

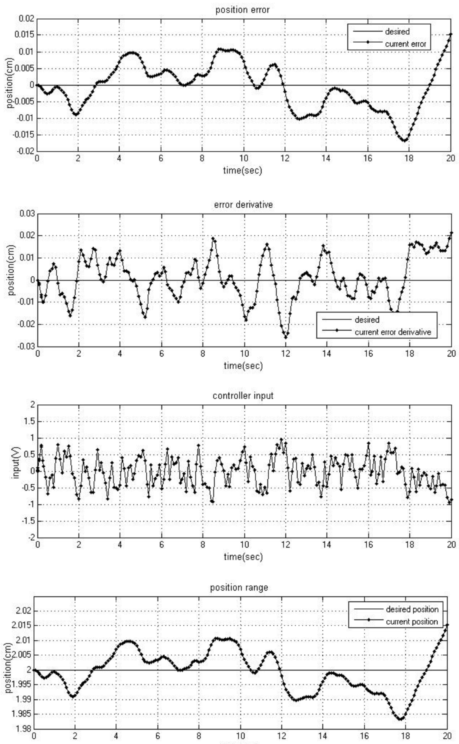


Fig. 4. Simulation results for central levitation height holding at 2cm with uncertainty bound

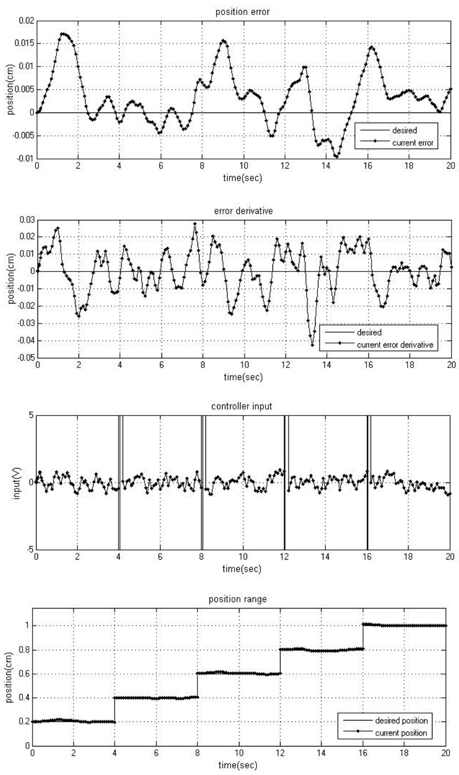


Fig. 5. Simulation results for repeating a 0.2cm step-train response in x-axis

The tracking response performance of magnetic levitation vehicle system is shown in Fig.3, where the central levitation height can track the desire signal $x(t) = 0.5sint + 1$. From the position error curve shown in Fig.3, we can see that the error can be controlled within $0.03cm$. The simulation result, as shown in Fig.4, demonstrates the ball can keep balance at the equilibrium point. The simulation results for repeating a $0.2cm$ step-train response in x-axis, as shown in Fig.5, indicates that the effective tracking performance of the proposed magnetic levitation vehicle system.

5 CONCLUSIONS

This paper does a research on the suspension control problem of the magnetic levitation vehicle system, which in nature, is served as strongly coupling, nonlinear and instability. The related analysis had been applied in to the maglev system. The linearization method and adaptive sliding mode variable structure can stable the system effectively.

6 ACKNOWLEDGMENTS

This work was supported by the National Basic Research Program of China(973 Program: 2012CB821200, 2012CB821201) and the NSFC(61134005, 60921001, 61327807).

REFERENCES

[1] D. Chao, Y. Kato, and D. Spilman(1993), ASliding Mode and Classical Control of Magnetic Levitation Systems, IEEE Control Syst. Mag.,42-48

[2] B. V. Jayawant, P.K.Sinha, and D.G.Aylwin(1976), Feedback Control Sytem for D.C. Electromagnets in Passenger-Carrying Vehicles, Int. J. Control,11:627-639

[3] M. Proise(1993), System Concept Definition of the Grumman Superconducting Electromagnetics Suspension (EMS) Maglev Design, Maglev 93 Conf. Argonne National Laboratory,11:19-21

[4] E.E. Covert, M. Vlajinac, T. Stephens, and M. Finston(1973), Magnetic Balance and Suspention System for use with Wind Tunnels, Progress in Aerospace Science,11: 27-107

[5] M.Y. Chen, M.J. Wang, and L.C. Fu(2008), Modeling and controller design of a maglev guiding system for application in precision positioning, Industrial Electronics, IEEE Transactions on, IEEE,11: 4-14

[6] H.X. Chen, Z.Q. Long, and W.S. Chang(2006), Fault tolerant control research for high-speed maglev system with sensor failure,Control and Automation, IEEE, 11:2281-2285

Variable-poled tracking control of a two-wheeled mobile robot using differential flatness

Liming Chen and Yingmin Jia

The Seventh Research Division and the Department of Systems and Control, Beihang University (BUAA),
Beijing 100191, China
(E-mails: clmtest@126.com; ymjia@buaa.edu.cn).

Abstract: This paper investigates the tracking controller design of a two-wheeled mobile robot in its kinematic model and dynamic model. Differential flatness and linear time-varying(LTV) systems' PD-spectral theory are used. Based on differential flatness, original system is transformed via a state prolongation and state transformation into a normal form to apply feedback linearization. Then using PD-spectral theory, variable poles of tracking error dynamics are assigned to realize trajectory tracking stability. Finally, simulation results are presented to demonstrate the feasibility and effectiveness of the proposed method.

Keywords: mobile robot, nonlinear control, differential flatness

1 INTRODUCTION

Wheeled mobile robots have been proven to be one of the most active areas of research since they are much useful in varieties of applications ranging from industrial settings, to military systems, to home robotics, *etc*. One of the main lines of research is the trajectory tracking problem, which is concerned with driving a mobile robot as close as possible to a desire explicit trajectory. And the tracking control approaches include backstepping [1], sliding mode control [2], linearization [3], neural network-based control [4], fuzzy control [5] and differential flatness-based control [6-8] which is also used in this paper.

Differential flatness has been introduced by Fliess *et al* [9]. It is a very useful tool for nonlinear controller design. Roughly speaking, a system is differentially flat, if there exists variables of the same dimension as inputs, called flat output, such that states and inputs can be algebraically expressed in terms of flat output and its derivatives. Moreover, this mapping is invertible, and the system is equivalent to a linear one. If the desire trajectory of flat output is given, then by performing a feedback linearization and designing a time invariant controller for the linearized system around the desire trajectory, stable tracking error dynamics are achieved. Kinematic model and dynamic model of a two-wheeled mobile robot have been proven to be differentially flat by choosing the center position of the wheel axle of the robot as the flat output. In [7,8], controllers are designed in this scheme. However, in these controllers, the parameters of the tracking error dynamics are constant, which means that the convergence speed of tracking error is fixed, and this brings limitation for this method.

In this paper, the use of PD-spectral theory of LTV systems is proposed for tracking controller design. The PD-spectral theory of LTV systems has been developed by Zhu [10], which can be seen as a natural extension of the conventional eigenvalue-eigenvector theory for LTI systems. After the state transformation based on differential flatness, the application of PD-spectral theory becomes much easier. By assigning time-varying "poles", more generalized tracking error dynamics can be obtained. And the poles can be changed at any time as we want.

The rest of this paper is organized as follows: In section 2, the kinematic model and dynamic model of a two-wheeled mobile robot are derived and analyzed with differential flatness. Section 3 presents the design of control law in both models. Simulation results are shown in section 4. Finally, concluding remarks are given in section 5.

2 MODELS OF TWO-WHEELED MOBILE ROBOT WITH DIFFERENTIAL FLATNESS

2.1 Kinematic model

Fig. 1 shows that the robot's configuration in Cartesian coordinates is given by $\mathbf{q} = [x, y, \theta]^T$, where (x, y) is the coordinates of the center of the wheel axle and θ is the heading angle of the robot. With the assumption of no-slip condition at the wheel contact points, the velocity of the wheel centers are parallel to the heading orientation and the nonholonomic constraint is given by $C(\mathbf{q})\dot{\mathbf{q}} = 0$, where $C(\mathbf{q}) = [sin\theta, -cos\theta, 0]$. Then using this constrain equation, the kinematic model of the mobile robot can be written as

$$\dot{\mathbf{q}} = S(\mathbf{q})\mathbf{v} \tag{1}$$

where $S(\mathbf{q}) = \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{v} = [v, \omega]^T$ is the vector of the heading speed v and turning speed ω of the mobile robot.

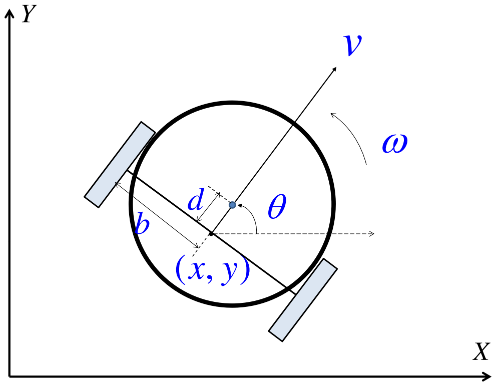


Fig. 1. The configuration of a two-wheeled mobile robot with no slip.

To outline how the kinematic model of the two-wheeled mobile robot is differentially flat, we need to select suitable flat outputs and express all state variables and inputs in terms of the flat outputs and their derivatives. The dimensions of flat outputs should be equal to that of the inputs. Here, we can choose the center position of the wheel axle (x,y) as the flat outputs (z_1,z_2).

With the chosen flat outputs, the states can be expressed as $x = z_1$, $y = z_2$, $\theta = \arctan\frac{\dot{z}_2}{\dot{z}_1}$ And the two inputs can be written as $v = \sqrt{\dot{z}_1^2+\dot{z}_2^2}$, $\omega = \frac{\dot{z}_1\ddot{z}_2-\ddot{z}_1\dot{z}_2}{\dot{z}_1^2+\dot{z}_2^2}$ It can be noticed that the expression of state θ contains the first order derivatives of both the flat outputs, $\dot{z}_1$ and $\dot{z}_2$. According to differential flatness theory, the system (1) need to be extended to 4 dimensions. Apply one prolongation of v by considering it as an additional state, then the extended system is given by

$$\begin{aligned} \dot{x} &= v\cos\theta & \dot{v} &= u_1 \\ \dot{y} &= v\sin\theta & \dot{\theta} &= u_2 \end{aligned} \qquad (2)$$

where u_1 and $u_2 = \omega$ are new inputs of the extended system. The new inputs can be calculated as

$$u_1 = \frac{\dot{z}_1\ddot{z}_1+\dot{z}_2\ddot{z}_2}{\sqrt{\dot{z}_1^2+\dot{z}_2^2}}, \quad u_2 = \frac{\dot{z}_1\ddot{z}_2-\ddot{z}_1\dot{z}_2}{\dot{z}_1^2+\dot{z}_2^2} \qquad (3)$$

The extended system (2) can be transformed via a state transformation $[x,y,v,\theta]^T \longleftrightarrow [z_1,z_2,\dot{z}_1,\dot{z}_2]^T$ into a normal form

$$\begin{aligned} \dot{z}_1 &= \dot{z}_1, & \ddot{z}_1 &= u_1\cos(\arctan\frac{\dot{z}_2}{\dot{z}_1}) - u_2\sqrt{\dot{z}_1^2+\dot{z}_2^2}\sin(\arctan\frac{\dot{z}_2}{\dot{z}_1}) \\ \dot{z}_2 &= \dot{z}_2, & \ddot{z}_2 &= u_1\sin(\arctan\frac{\dot{z}_2}{\dot{z}_1}) + u_2\sqrt{\dot{z}_1^2+\dot{z}_2^2}\cos(\arctan\frac{\dot{z}_2}{\dot{z}_1}) \end{aligned} \qquad (4)$$

Since the expressions of new inputs u_1 and u_2 in (3) are derived from the system equations (2), the state transformation ensures that (3) satisfies the equations (4). By replacing $\ddot{z}_1$ and $\ddot{z}_2$ in (3) with u_x and u_y, respectively:

$$u_1 = \frac{\dot{z}_1 u_x+\dot{z}_2 u_y}{\sqrt{\dot{z}_1^2+\dot{z}_2^2}}, \quad u_2 = \frac{\dot{z}_1 u_y - u_x\dot{z}_2}{\dot{z}_1^2+\dot{z}_2^2} \qquad (5)$$

where u_x and u_y are parameters to be designed, and substituting (5) into the system equations (4), one gets

$$\begin{aligned} \dot{z}_1 &= \dot{z}_1, & \ddot{z}_1 &= u_x \\ \dot{z}_2 &= \dot{z}_2, & \ddot{z}_2 &= u_y \end{aligned} \qquad (6)$$

Actually, if only the states of the original system (2), $[x,y,v,\theta]^T$, can be estimated, using the relationship of state transformation, the inputs (5) can be rewritten as

$$u_1 = u_x\cos\theta + u_y\sin\theta, \quad u_2 = (u_y\cos\theta - u_x\sin\theta)/v \qquad (7)$$

2.2 Dynamic model

By ignoring the mass of the wheels, the equations of motion can be derived using Euler-Lagrange method as

$$M(\mathbf{q})\ddot{\mathbf{q}} + \bar{C}(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}} = E(\mathbf{q})\tau - C^T(\mathbf{q})\lambda \qquad (8)$$

where

$$M = \begin{bmatrix} m & 0 & -dm\sin\theta \\ 0 & m & dm\cos\theta \\ -dm\sin\theta & dm\cos\theta & d^2m+I \end{bmatrix}, \tau = \begin{bmatrix} \tau_r \\ \tau_l \end{bmatrix}$$
$$\bar{C} = \begin{bmatrix} 0 & 0 & -dm\dot{\theta}\cos\theta \\ 0 & 0 & -dm\dot{\theta}\sin\theta \\ 0 & 0 & 0 \end{bmatrix}, E = \frac{1}{r}\begin{bmatrix} \cos\theta & \cos\theta \\ \sin\theta & \sin\theta \\ b & -b \end{bmatrix} \qquad (9)$$

Here, m is the robot mass, I the moment of inertia of the robot about its center of mass, d the distance between the center of mass and the center of the wheel axle, r the wheel radius, b half distance between the two wheels, τ_l and τ_r the motor torques on the wheels, and λ the constraint force.

By differentiating (1), one gets $\ddot{\mathbf{q}} = \dot{S}\mathbf{v} + S\dot{\mathbf{v}}$. Then by substituting $\ddot{\mathbf{q}}$ into (8), pre-multiplying by S^T, and using the property $S^TC^T = 0$, one can have

$$\dot{\mathbf{v}} = -(S^TMS)^{-1}S^T(M\dot{S}+\bar{C}S)\mathbf{v} + (S^TMS)^{-1}S^TE\tau \qquad (10)$$

This can finally be calculated as $\dot{\mathbf{v}} = A(\mathbf{v}) + B\tau$, where $A(\mathbf{v}) = \begin{bmatrix} d\omega^2 \\ -\frac{dm\omega v}{d^2m+I} \end{bmatrix}$, $B = \begin{bmatrix} \frac{1}{mr} & \frac{1}{mr} \\ \frac{b}{r(d^2m+I)} & -\frac{b}{r(d^2m+I)} \end{bmatrix}$. After introducing an input transformation $\bar{\mathbf{u}} = [\bar{u}_1,\bar{u}_2]^T = A(\mathbf{v}) + B\tau$, the dynamic model can be written as

$$\dot{\mathbf{q}} = S(\mathbf{q})\mathbf{v}, \quad \dot{\mathbf{v}} = \bar{\mathbf{u}} \qquad (11)$$

Here, we can also choose the center position of the wheel axle (x,y) as the flat outputs (z_1,z_2) and then all state variables and inputs can be expressed in terms of the flat outputs and their derivatives. The expression of state ω contains the second order derivatives of both the flat outputs, $\ddot{z}_1$ and $\ddot{z}_2$, which means that the system (11) need to be extended to 6 dimensions so that it can be transformed into a normal form via a state transformation. On applying one prolongation of $\bar{u}_1$, the extended system is given by

$$\begin{array}{lll} \dot{x} = v\cos\theta & \dot{v} = \bar{u}_1 & \dot{\bar{u}}_1 = \tilde{u}_1 \\ \dot{y} = v\sin\theta & \dot{\theta} = \omega & \dot{\omega} = \tilde{u}_2 \end{array} \tag{12}$$

where $\tilde{u}_1$ and $\tilde{u}_2 = \bar{u}_2$ are new inputs of the extended system. The state transformation can be written as $[x, y, v, \theta, \bar{u}_1, \omega]^T \longleftrightarrow [z_1, z_2, \dot{z}_1, \dot{z}_2, \ddot{z}_1, \ddot{z}_2]^T$. Now, we calculate the third order derivatives of the flat outputs directly using (12) as

$$\begin{bmatrix} \dddot{z}_1 \\ \dddot{z}_2 \end{bmatrix} = C + D \begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \end{bmatrix} \tag{13}$$

where

$$C = \begin{bmatrix} -2\bar{u}_1\omega\sin\theta - v\omega^2\cos\theta \\ 2\bar{u}_1\omega\cos\theta - v\omega^2\sin\theta \end{bmatrix}, D = \begin{bmatrix} \cos\theta & -v\sin\theta \\ \sin\theta & v\cos\theta \end{bmatrix} \tag{14}$$

By replacing $\dddot{z}_1$ and $\dddot{z}_2$ with $\bar{u}_x$ and $\bar{u}_y$, respectively, this yields

$$\begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \end{bmatrix} = D^{-1}\left(\begin{bmatrix} \bar{u}_x \\ \bar{u}_y \end{bmatrix} - C\right) \tag{15}$$

3 DESIGN OF CONTROL LAW

3.1 Introduction of PD-spectral theory [10]

Consider SISO LTV systems represented by the nth-order scalar LTV dynamical systems of the form:

$$y^{(n)} + \alpha_n(t)y^{(n-1)} + \cdots + \alpha_2(t)\dot{y} + \alpha_1(t)y = 0 \tag{16}$$

It can be conveniently represented as $D_\alpha\{y\} = 0$ using the scalar polynomial differential operator (SPDO)

$$D_\alpha = \delta^n + \alpha_n(t)\delta^{n-1} + \cdots + \alpha_2(t)\delta + \alpha_1(t) \tag{17}$$

where $\delta = d/dt$ is the derivative operator. The factorization of SPDO can be represented as

$$D_\alpha = (\delta - \lambda_n(t))\cdots(\delta - \lambda_2(t))(\delta - \lambda_1(t)) \tag{18}$$

where a collection $\{\lambda_k(t)\}_{k=1}^n$ is called a series D-spectrum(SD-spectrum) for D_α and an n-parameter family $\{\rho_k(t) = \lambda_{1,k}(t)\}_{k=1}^n$ is called a parallel D-spectrum(PD-spectrum) for D_α, where $\lambda_{1,k}(t)$ are n particular solutions for $\lambda_1(t)$ satisfying some nonlinear independent constrains. Actually, $\{y_k(t) = \exp(\int \rho_k(t)\mathrm{d}t)\}_{k=1}^n$ constitutes a fundamental set of solutions to $D_\alpha\{y\} = 0$.

The solution to $D_\alpha\{y\} = 0$ is uniformly asymptotically stable if
(i) all PD-eigenvalues are of polynomial order or slower, that is, an integer $m > 0$ exists such that $\lim_{t\to\infty} \frac{\rho_k(t)}{t^m} = 0$;
(ii) the extended means of real parts of PD-eigenvalues, that is $\mathrm{em}(\mathrm{Re}\rho_k(t)) = \lim_{T\to\infty} \frac{1}{T}\int_{t_0}^{t_0+T} \mathrm{Re}\rho_k(t)\mathrm{d}t$, are all negative.

3.2 Control law of kinematic model

If the desired trajectory of flat outputs (z_1, z_2) are given by $(z_{1d}(t), z_{2d}(t))$, tracking error can be introduced as $\mathbf{e} = [e_1, e_{11}, e_2, e_{21}]^T = [z_1 - z_{1d}, \dot{z}_1 - \dot{z}_{1d}, z_2 - z_{2d}, \dot{z}_2 - \dot{z}_{2d}]^T$ Using (6), parameters u_x and u_y can be designed as

$$\begin{array}{l} u_x = \ddot{z}_{1d} - \alpha_{1x}(t)e_1 - \alpha_{2x}(t)e_{11} \\ u_y = \ddot{z}_{2d} - \alpha_{1y}(t)e_2 - \alpha_{2y}(t)e_{21} \end{array} \tag{19}$$

which yields

$$\begin{array}{l} \ddot{e}_1 + \alpha_{2x}(t)\dot{e}_1 + \alpha_{1x}(t)e_1 = 0 \\ \ddot{e}_2 + \alpha_{2y}(t)\dot{e}_2 + \alpha_{1y}(t)e_2 = 0 \end{array} \tag{20}$$

PD-spectral theory is used to design these time-varying control gains to ensure the error dynamics to be uniformly asymptotically stable.

First appropriate time-varying PD-eigenvalues which satisfy (i) and (ii) are designed as ρ_1 and ρ_2. Then corresponding SD-eigenvalues λ_1 and λ_2 can be calculated as $\lambda_1 = \rho_1$, $\lambda_2 = \rho_2 + \frac{\dot{\rho}_2 - \dot{\rho}_1}{\rho_2 - \rho_1}$. Using (18), we obtain $\alpha_{1x}(t) = \lambda_2\lambda_1 - \dot{\lambda}_1 = \rho_1\rho_2 + \frac{\rho_1\dot{\rho}_2 - \rho_2\dot{\rho}_1}{\rho_2 - \rho_1}$, $\alpha_{2x}(t) = -\lambda_2 - \lambda_1 = -\rho_1 - \rho_2 - \frac{\dot{\rho}_2 - \dot{\rho}_1}{\rho_2 - \rho_1}$. $\alpha_{1y}(t)$ and $\alpha_{2y}(t)$ can be designed in the same way.

3.3 Control law of dynamic model

Similarly, if the desired trajectory of flat outputs are given by $(z_{1d}(t), z_{2d}(t))$ and tracking error $\mathbf{e} = [e_1, e_{11}, e_{12}, e_2, e_{21}, e_{22}]^T = [z_1 - z_{1d}, \dot{z}_1 - \dot{z}_{1d}, \ddot{z}_1 - \ddot{z}_{1d}, z_2 - z_{2d}, \dot{z}_2 - \dot{z}_{2d}, \ddot{z}_2 - \ddot{z}_{2d}]^T$, then parameters $\bar{u}_x$ and $\bar{u}_y$ can be designed as

$$\begin{array}{l} \bar{u}_x = \dddot{z}_{1d} - \alpha_{1x}(t)e_1 - \alpha_{2x}(t)e_{11} - \alpha_{3x}(t)e_{12} \\ \bar{u}_y = \dddot{z}_{2d} - \alpha_{1y}(t)e_2 - \alpha_{2y}(t)e_{21} - \alpha_{3y}(t)e_{22} \end{array} \tag{21}$$

By designing appropriate PD-eigenvalues ρ_1, ρ_2 and ρ_3, SD-eigenvalues can be calculated as $\lambda_1 = \rho_1$, $\lambda_2 = \rho_2 + \frac{\dot{V}_2}{V_2}$, $\lambda_3 = \rho_3 + \frac{\dot{V}_3}{V_3} - \frac{\dot{V}_2}{V_2}$ where $V_2 = \rho_2 - \rho_1$, $V_3 = \det\begin{bmatrix} 1 & 1 & 1 \\ \rho_1 & \rho_2 & \rho_3 \\ \rho_1^2 + \dot{\rho}_1 & \rho_2^2 + \dot{\rho}_2 & \rho_3^2 + \dot{\rho}_3 \end{bmatrix}$, and then control gains are obtained by $\alpha_{1x}(t) = -\lambda_3\lambda_2\lambda_1 + \lambda_3\dot{\lambda}_1 + \dot{\lambda}_2\lambda_1 + \lambda_2\dot{\lambda}_1 - \ddot{\lambda}_1$, $\alpha_{2x}(t) = \lambda_3\lambda_2 + \lambda_3\lambda_1 + \lambda_2\lambda_1 - \dot{\lambda}_2 - 2\dot{\lambda}_1$, $\alpha_{3x}(t) = -\lambda_3 - \lambda_2 - \lambda_1$. $\alpha_{1y}(t)$, $\alpha_{2y}(t)$ and $\alpha_{3y}(t)$ can be designed in the same way.

4 SIMULATION RESULT

The desired trajectory is given by $z_{1d}(t) = 3.75 \times 10^{-4}t^3 - 1.125 \times 10^{-2}t^2 + 0.15t(m), z_{2d}(t) = -3 \times 10^{-4}t^3 + 9 \times 10^{-3}t^2(m)$, over $t = [0, 20]$.

In the kinematic model, initial states are set as $[x(0), y(0), v(0), \theta(0)] = [-0.1m, 0.1m, 0.05m/s, -0.5rad]$.

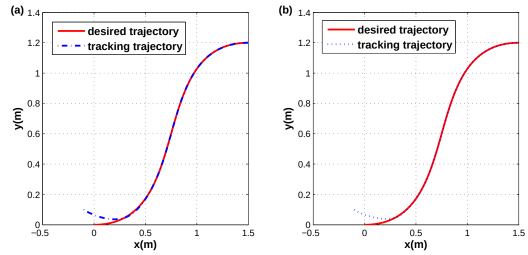


Fig. 2. Desired trajectory and tracking trajectory: (a)Kinematic model. (b)Dynamic model.

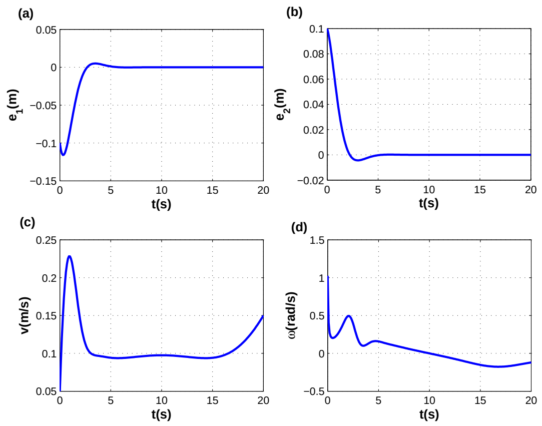


Fig. 3. Kinematic model: (a)Tracking error e_1. (b)Tracking error e_2. (c)Heading speed v. (d)Turning speed ω.

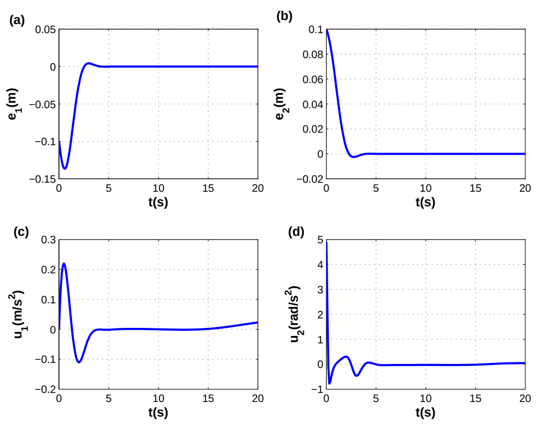


Fig. 4. Dynamic model: (a)Tracking error e_1. (b)Tracking error e_2. (c)Transformed input $\bar{u}_1$. (d)Transformed input $\bar{u}_2$.

It should be pointed out that the additional state v is in the controller. PD-eigenvalues are selected as $\rho_1 = -(1+0.25t)+(1+0.25t)i, \rho_2 = -(1+0.25t)-(1+0.25t)i$. Fig. 2(a) shows that the mobile robot gradually converges to the desired trajectory and finally moves along it. Fig. 3 demonstrates the tracking errors e_1 and e_2 which converge to zero and the control inputs v and ω.

In the dynamic model, $[x(0), y(0), v(0), \theta(0), \bar{u}_1(0), \omega(0)] = [-0.1m, 0.1m, 0.05m/s, -0.5rad, 0, 0]$. Controller is designed with PD-eigenvalues $\rho_1 = -(1+0.25t)+(1+0.25t)i, \rho_2 = -(1+0.25t)-(1+0.25t)i, \rho_3 = -2-0.25t$. Fig. 2(b) shows the tracking trajectory and Fig. 4 depicts the tracking errors and the transformed inputs $\bar{u}_1$ and $\bar{u}_2$.

5 CONCLUSIONS

In this paper, we have presented a novel method for trajectory tracking control of a two-wheeled mobile robot in its kinematic model and dynamic model. Based on differential flatness, the systems can be transformed into normal forms to use feedback linearization. The application of PD-spectral theory ensures the stability of the tracking error dynamics and establishes adjustable poles to change as we want. The simulation results confirm the effectiveness of the proposed control scheme.

6 ACKNOWLEDGMENTS

This work was supported by the National Basic Research Program of China(973 Program: 2012CB821200, 2012CB821201) and the NSFC(61134005, 60921001, 61327807).

REFERENCES

[1] R. Fierro and F. V. Lewis (1997), Control of a nonholonimic mobile robot: backstepping kinematics into dynamics, J. Robot. Syst., 14(3): 149-163

[2] J. M. Yang and J. H. Kim (1999), Sliding mode control for trajectory tracking of nonholonomic wheeled mobile robots, IEEE Trans. Robot. Autom., 15(3): 578-587

[3] D. H. Kim and J. H. Oh (1999), Tracking control of a two-wheeled mobile robot using input-output linearization, Contr. Eng. Pract., 7(3): 369-373

[4] V. Boquete, R. Garcia, R. Barea, and M. Mazo (1999), Neural control of the movements of a wheelchair, J. Intell. Robot. Syst., 25(3): 213-226

[5] T. Das and I. N. Kar (2006), Design and implementation of an adaptive fuzzy logic-based controller for wheeled mobile robots, IEEE Trans. Contr. Syst. Technol., 14(3): 501-510

[6] Chun-Hsu Ko and Sunil K. Agrawal (2010), Walk-assist robot: a novel approach to gain selection of a braking controller using differential Flatness, American Control Conference, 2799-2804

[7] Chin Pei Tang (2009), Differential flatness-based kinematic and dynamic control of a differentially driven wheeled mobile robot, in Proc. IEEE Int. Conference on Robotics and Biomimetics, 2267-2272

[8] Ji-Chul Ryu and Sunil K. Agrawal (2008), Differential flatness-based robust control of a two-wheeled mobile robot in the presence of slip, in Proc. DSCC ASME Dynamic Systems and Control Conference, 1-7

[9] M. Fliess, J. Lévine, Ph. Martin, and P. Rouchon (1995), Flatness and defect of non-linear systems: introduction theory and examples, International Journal of Control, 61(6): 1327-1361

[10] J. Zhu (1995), A unified spectral theory for linear time-varying systems–progress and challenges, in Proc. IEEE Conference on Decision and Control, 2540-2546

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Adaptive Consensus Control of Multi-Agent Systems with Large Uncertainty and Time Delays

Dongxu Zou[1a], Weicun Zhang [1b]

1. School of Automation and Electrical Engineering,

University of Science and Technology Beijing, Beijing 100083, China

[a] dongxu_zou@126.com, [b] weicunzhang@263.net

Abstract. A weighted multi-model adaptive control method is proposed to achieve consensus of multi-agent system with large parameter perturbation and communication delays, in which H^{∞} control is adopted to construct the controller set. Moreover a simple and effective weighting algorithm is also presented. The simulation results demonstrate the effectiveness of the proposed method.

Keywords: Multi-agent system; weighted MMAC; H^{∞} control

1. Introduction

During the past decade, the study of decentralized control has been focused on multi-agent systems, such as flocking or swarming behaviors [1-3], multi-robot formation control [4,5], and path planning [6] and so on.

The consensus problem, the most important and fundamental issue in the cooperation control of multi-agent system, is of theoretical value and practical significance. For a multi-agent system, consensus means that the states of all agents tend to be identical asymptotically under given protocols, based on the communication networks.

Recently, some researchers have solved the consensus problem of multi-agent systems with time-varying external disturbances and random communication delays [7]. In [8], model uncertainty was taken into consideration, and a robust H∞ controller was designed.

In this paper, the consensus problem is considered for multi-agent systems with large parameter perturbation. External disturbances and communication delays are also taken into account. Weighted MMAC (WMMAC) shows its priority of system with large uncertainty. Finally, simulation results indicate that under the proposed protocol, multi-agent systems with large parameter perturbation reach the desired consensus performance in the presence of communication delays.

2. Uncertain Multi-Agent Systems

Consider a second-order multi-agent system consisting of n identical agents with the ith one modeled by as the form of (1).

$$\begin{aligned} &\dot{x}_i(t) = Ax_i(t) + B_1\omega_i(t) + B_2u_i(t), \\ &x_i(t) = [x_{i1}(t), x_{i2}(t)]^T, \omega_i(t) \in L_2(0,\infty] \end{aligned} \quad (1)$$

$x_i(t)$ is the state of agent i, while $\omega_i(t)$ represents appropriate external disturbance that belongs to $L_2(0,\infty]$. $u_i(t)$ denotes the control protocol.

If system matrices A, B_1, B_2 are uncertain, then they are supposed to follow the forms of (2).

In (2), A_0, B_{10}, B_{20} are constant matrices and $\Delta A(t), \Delta B_1(t), \Delta B_2(t)$ are time-varying matrices. E and $F_i, i = 1,2,3$ are constant matrices with appropriate dimensions and $\Sigma(t)$ is time-varying uncertain matrix satisfying $\Sigma(t)^T\Sigma(t) \le I$. It is also assumed that (A_0, B_{20}) is stabilizable.

$$\begin{aligned} &A = A_0 + \Delta A(t), B_1 = B_{10} + \Delta B_1(t) \\ &B_2 = B_{20} + \Delta B_2(t) \\ &[\Delta A(t)\ \Delta B_1(t)\ \Delta B_2(t)] = E\ \Sigma(t)[F_1\ F_2\ F_3] \end{aligned} \quad (2)$$

The consensus of multi-agent system is that the states of agents satisfy (3) under control protocol $u_i(t)$.

$$\lim_{t\to\infty}\left(x_i(t) - x_j(t)\right) = \mathbf{0}, \forall i, j \in N \quad (3)$$

3. Weighted multi-model adaptive control

Weighted multi-model adaptive controller is made up of the following sections.

3.1 Model Set. The model set is designed as $\Omega = \{M_i \mid i = 1,2,\cdots,\mathrm{n}\}$ considering uncertainty of plant parameters and structures. Ω represents a model set with element M_i. In fact, it relies on multiple models to approach the uncertain system. Obviously, performance of control system depends on the method of model set and the amount of models directly.

3.2 Controller Set. The controller set C is shown as the form of $C = \{C_i \mid i = 1,2,\cdots,\mathrm{n}\}$, in which elements are up to corresponding models. In other words, the controller C_i is designed for model M_i.

In this paper, robust H^∞ method is selected as control strategy. Consider the multi-agent system (1), the output is supposed to follow (4). We can reformulate the system (1) as the form of (5) that satisfies (6) with the help of Kronecker product.

$$z_i(t) = x_i(t) - \frac{1}{n}\sum_{j=1}^{n} x_j(t), i = 1, \ldots, n \tag{4}$$

$$\begin{aligned} \dot{x}(t) &= (I_n \otimes A)x(t) + (I_n \otimes B_1)\omega(t) + (I_n \otimes B_2)u(t) \\ z(t) &= (L_c \otimes I_m)x(t) \end{aligned} \tag{5}$$

$$\begin{aligned} x(t) &= [x_1(t)^T, x_2(t)^T, \ldots, x_n(t)^T]^T \in \mathrm{R}^{mn} \\ \omega(t) &= [\omega_1(t)^T, \omega_2(t)^T, \ldots, \omega_n(t)^T]^T \in \mathrm{R}^{m_1 n} \\ u(t) &= [u_1(t)^T, u_2(t)^T, \ldots, u_n(t)^T]^T \in \mathrm{R}^{m_2 n} \\ z(t) &= [z_1(t)^T, z_2(t)^T, \ldots, z_n(t)^T]^T \in \mathrm{R}^{mn} \\ L_c &= \left[L_{c_{ij}} \right] \in \mathrm{R}^{n\times n}, L_{c_{ij}} = \begin{cases} (n-1)/n, i = j \\ -1/n, i \neq j \end{cases} \end{aligned} \tag{6}$$

According to the neighbors' states, the protocol of agent i can be designed as (7) to achieve the above conditions.

$$u_i(t) = K \sum_{j \in N_i(t)} a_{ij}(t)\left[x_i(t - d(\mathrm{t})) - x_j(t - d(\mathrm{t})) \right] \tag{7}$$

Substitute control protocol (7) into system (5), result in the closed-loop system (8).

$$\begin{aligned} \dot{x}(t) &= (I_n \otimes A)x(t) + (L_\sigma \otimes B_2 K)x(t - d(t)) \\ &+ (I_n \otimes B_1)\omega(t), \quad z(t) = (L_c \otimes I_m)x(t) \end{aligned} \tag{8}$$

Finally, the consensus problem of multi-agent system is converted to calculate the feedback matrix K.

Under control protocol (7), system (1) achieves consensus with a given H∞ index γ, if there are a scalar α>0, a positive define matrix $P \in R^{m\times m}$ and a matrix $Q \in R^{m_2 \times m}$ that satisfies the LMI (9).

$$\begin{bmatrix} \Psi_{\sigma i} + \Psi_{\sigma i}^T & \alpha\lambda_{\sigma i} B_2 Q & B_1 & \bar{d}\Psi_{\sigma i}^T & P \\ \alpha\lambda_{\sigma i} Q^T B_2^T & -\alpha P & 0 & \bar{d}\alpha\lambda_{\sigma i} Q^T B_2^T & 0 \\ B_1^T & 0 & -\gamma^2 I & \bar{d}B_1^T & 0 \\ \bar{d}\Psi_{\sigma i} & \bar{d}\alpha\lambda_{\sigma i} B_2 Q & \bar{d}B_1 & -\alpha P & 0 \\ P & 0 & 0 & 0 & -I \end{bmatrix} < 0$$
$$\Psi_{\sigma i} = A_0 P + \lambda_{\sigma i} B_2 Q \tag{9}$$

If LMI (9) is feasible for the maximum and minimum eigenvalues of all the connected topological graphs, the feedback matrix can be defined by $K = QP^{-1}$.

3.3 Weighting Algorithm. The method based on probability-weighted is that the weighted sum of controller set is taken as the global control input. The weight of controller can be calculated by partition theorem.

Suppose the control output of controller C_i is $u_i(k)$, the weight is $p_i(k)$, the global control input is $u(k)$, the system output is $y(k)$, the output of ith model is $y_i(k)$.

Let the output difference between ith model and plant be $e_i(k) = y(k) - y_i(k)$. Then the weighting algorithm can be designed as follows.

First, weights of all models are initialized to 1/N, if there are N models in set totally as (10).

$$p_i(0) = r_i(0) = 1/N, i = 1, 2, \ldots, N \tag{10}$$

Then, it is conductive to improve the anti-jamming capability by selecting the cumulative mean squared error as the performance index shown in (11).

$$l_i(k) = 1 + \frac{1}{k}\sum_{j=1}^{k} e_i^2(k) \tag{11}$$

Finally, the computational process of weights is denoted by (12). This method is much simpler than probability-weighted method. The performance indexes are forced to compete at every step in order to make the weights converge.

$$\begin{aligned} l_{\min}(k) &= \min_{i=1,\ldots,N} l_i(k) \\ r_i(k) &= \frac{l_{\min}(k)}{l_i(k)} r_i(k-1) \\ p_i(k) &= \frac{r_i(k)}{\sum_{i=1}^{N} r_i(k)} \end{aligned} \tag{12}$$

Simulation results

In this section, we will illustrate the consensus of multi-agent system under the method of WMMAC.

Considering a multi-agent system of four agents, the parameter matrices are designed as (13).

$$A_0=\begin{bmatrix}0 & -1\\ 2 & 1\end{bmatrix}, B_{10}=\begin{bmatrix}0\\ 1\end{bmatrix}, B_{20}=\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix},$$
$$\Delta A(t)=\begin{bmatrix}0 & 0\\ 6\left(1-e^{-t}\right) & 0\end{bmatrix}, \Delta B_1(t)=\mathbf{0}, \Delta B_2(t)=\mathbf{0} \tag{13}$$

The communication delay is assumed to be constant 0.05. The external disturbances are supposed to be band-limited white noises shown as (14).

$$\omega(t)=\begin{bmatrix}\omega_1(t) & \omega_2(t) & \omega_3(t) & \omega_4(t)\end{bmatrix}^T$$
$$=\begin{bmatrix}2\omega(t) & -1.5\omega(t) & 1.2\omega(t) & 1.8\omega(t)\end{bmatrix}^T \tag{14}$$

For simplicity, the H∞ performance index is chosen as $\gamma=1$, the constant scalar $\alpha=0.1$, $\lambda=1$ and all the nonzero weighting factors of adjacency matrix are assumed to be 1. The simplified undirected interaction graph is shown in Fig. 1.

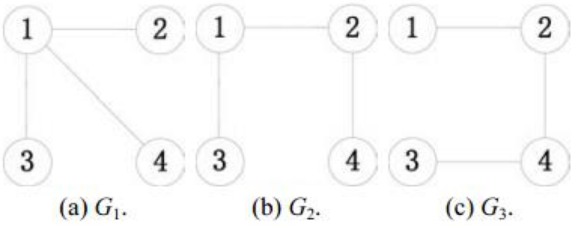


Fig. 1 simplified undirected graph

We can get the maximum and minimum nonzero eigenvalues $\lambda_{\sigma_* i_*}=0.5858, \lambda_{\sigma^* i^*}=4$ by calculating the Laplace matrices corresponding to the network graph shown in Fig. 1.

It is assumed that there are four models in the model set which is shown as (15).

$$A_1=\begin{bmatrix}0 & -1\\ 2 & 1\end{bmatrix}, A_2=\begin{bmatrix}0 & 1\\ 4 & 1\end{bmatrix}, A_3=\begin{bmatrix}0 & 1\\ 6 & 1\end{bmatrix}, A_4=\begin{bmatrix}0 & 1\\ 8 & 1\end{bmatrix} \tag{15}$$

According to the LMI (9), the feedback matrices Ki can be determined as (16) with the help of Matlab LMI toolbox.

$$K_1=\begin{bmatrix}-2.2365 & -0.3528\\ 0.1269 & -2.7761\end{bmatrix}, K_2=\begin{bmatrix}-1.4127 & 0.2621\\ 0.4986 & -2.3349\end{bmatrix}$$
$$K_3=\begin{bmatrix}-2.0759 & 0.9033\\ 1.4598 & -3.0850\end{bmatrix}, K_4=\begin{bmatrix}-2.4147 & 1.3164\\ 1.9986 & -2.9659\end{bmatrix} \tag{16}$$

The simulation results are shown in figures 2-3, in which Fig. 2 represents the trajectory of state x1, x2, and Fig. 3 represents the energy relationship between the system output $z(t)$ and the external disturbance $\omega(t)$

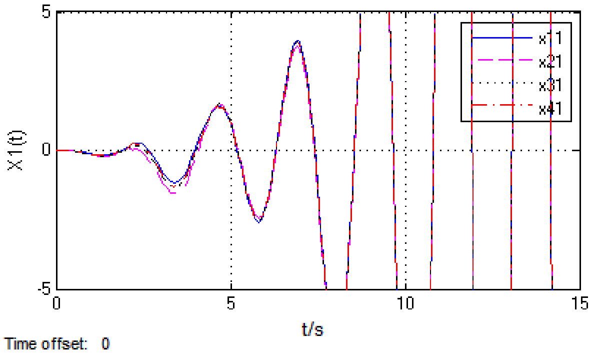


Fig. 2(a) State curve of x_1

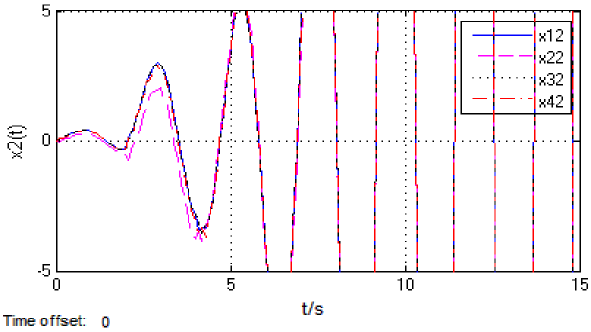


Fig. 2(b) State curve of x_2

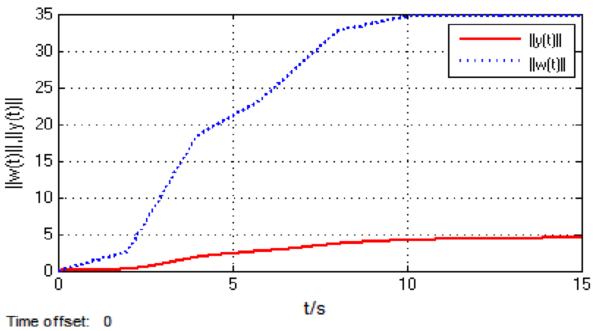


Fig. 3 Energy of $z(t)$ and $\omega(t)$

If the system states are asymptotically stable, the consensus of multi-agent system is all up to the steady-state values of each agent. Therefore, the states of agents are designed to be divergent in order to display the consensus of multi-agent system significantly.

It is easy to say that the multi-agent system achieves consensus according to Fig.2

When the state matrix A changes to

$$A=\begin{bmatrix}0 & 1\\ 7.5 & 1\end{bmatrix}$$

the weights of all models are displayed in Fig.4.

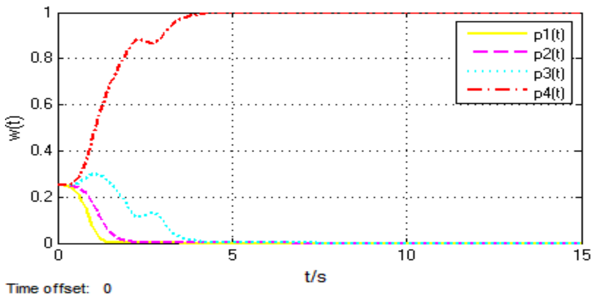


Fig.4 weights of models

In Fig.4, the weight of the 4th model is convergent to 1. Then the controlled quantity of the 4th model is chosen to be the global control input of the multi-agent system.

Conclusions

By applying the weighted MMAC methodology, we have addressed the adaptive consensus problem of multi-agent system with large parameter uncertainty and communication delays. According to the states' differences between the plant and each model, we can obtain the weight of each local controller with the weighting algorithm mentioned. Finally, the controller corresponding to the closest model to the plant will be chosen as the right controller.

References

[1] Olfati-Saber R. Flocking for multi-agent dynamic systems: algorithms and theory. IEEE Transactions on Automatic Control, 2006, 51(3):401-420.

[2] Tanner H G, Jadbabaie A and Pappas G J. Flocking in fixed and switching networks. IEEE Transactions on Automatic Control, 2007, 52(5):863-868.

[3] Su H S, Wang X F and Yang W. Flocking in multi-agent systems with multiple virtual leaders. Asian Journal of Control, 2008, 10(2):238-245.

[4] Egerstedt M and Hu X. Formation control with virtual leaders and reduced communications. IEEE Transactions on Robotics and Automation, 2001, 17(16):947-951.

[5] Olfati-Saber R and Murray R M. Distributed structural stabilization and tracking for formations of dynamic multi-agents. In Proceedings of the 41th IEEE Conference on Decision and Control, 2002, 209-215.

[6] F. Lian, Cooperative path planning of dynamical multi-agent systems using differential flatness approach, International Journal of Control, Automation, and Systems, 2008, 6(3): 401-412.

[7] W Ren, R W, Beard and D B Kingston, Multi-agent Kalman consensus with relative uncertainty, Proc. of the American Control Conference, 2005. 1865-1870.

[8] Y. Liu, Y. Jia, Robust H∞ consensus control of uncertain multi-agent systems with time delays, International Journal of Control, Automation, and Systems, 2011, 9(6): 1086-1094.

Iterative learning control for overhead crane systems

Bu Xuhui*, Wang Fuzhong, Li Sanyi, Yu Fashan
School of Electrical Engineering & Automation, Henan Polytechnic University, Jiaozuo, China.
Corresponding email: buxuhui@gmail.com

***Abstract*:** In many factories, overhead crane often transfers the same loads from one place to another following a predefined position and speed diagrams. This unique feature offers overhead crane an opportunity of improving its performance through learning iteratively. In this paper, we apply the iterative learning control approach to overhead crane systems. Based on the linearization dynamic model of overhead crane, an ILC scheme contains both feedforward learning control part and state feedback control is proposed. By providing a 2D formulation, the ILC design for crane systems can be transformed into the problem of state feedback control for 2-D systems described by Roesser models. It is shown that the proposed approach can guarantee the trolley position asymptotically converges to its desired profile with small swing angle. Simulations are illustrated to show the feasibility and effectiveness of the proposed approach.

***Keywords*:** iterative learning control; overhead crane; position tracking: convergence

I. INTRODUCTION

Overhead crane can transport a load from one place to another with high load capacity, and it is widely used in factory, wharf, etc. The purpose of control overhead crane is to transport load to the designated place following desired position and speed with small rope swing. Because high positioning accuracy, small swing angle and high safety are required, operating an overhead traveling crane is a hard work. Various controllers have been considered for overhead crane systems.

PID controller is widely used for overhead crane systems in practical systems [1-3]. However, it cannot deal with nonlinear dynamic effectively. In [4], an energy-based approach is proposed for overhead crane systems, and it can use nonlinear terms to improve the transient performance and repress the load swing angle effectively. However, it also results in sensitive to the changes of the parameters. To overcome these shortages, many other control strategies such as fuzzy logic control [5], adaptive fuzzy control [6] and adaptive sliding mode fuzzy control [7], are proposed for crane systems. Although these control strategies can obtain better tracking performance, the fuzzy rules are difficult to be built.

On the other hand, the normal operation process of a crane from the start place to final place is characterized by high repetition pattern. In many factories, loads with same mass and size are often transported from one place to another. In this progress, overhead crane operates according to their desired position or speed diagrams strictly, and their operating environments and control tasks also remain unchanged. For each operation, the trolley starts and ends at the same positions, and the system dynamic is almost repetitive due to the same mass and size of loads. The desired position or speed profile for overhead crane is predefined, and the trolley is controlled to follow that position or speed profile at each operation. Hence, we can improve tracking performance through learning iteratively or repetitively.

Iterative learning control (ILC) is an effective technique for systems with repetitive nature. Since the original work in [8], the general area of ILC has been the subject to intense research effort and widely applied in many fields [9, 10]. In this paper, an ILC design approach based 2-D system theory for overhead crane systems is presented. By providing a 2D system of the learning process, it is shown that the design problem of an ILC law for overhead crane can be transformed straightforwardly into feedback control problem of 2-D systems described by Roesser models.

This paper is organized as follows. Section 2 introduces overhead crane dynamics model and problem formulation. Section 3 presents an ILC design approach for overhead crane system by using 2D system theory. Section 4 is the simulation results, and Section 5 gives conclusions.

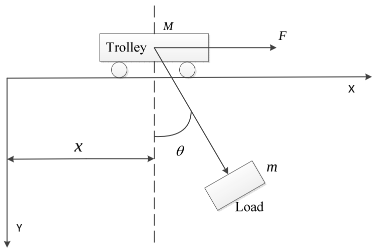


Fig. 1 Sketch of an overhead crane

II. CRANE SYSTEM DYNAMICS

As shown in Fig. 1, the overhead crane system contains a trolley, a rope and a load. The volume of the load can be ignored and the mass of the load is m. The mass of rope can be ignored compared with the load mass, and it also is assumed inflexible rod with a length of l. The trolley moves on a straight line with a mass of M. Then, using the Lagrange's equation of motion, the dynamic model of overhead crane can be given as [11]:

$$\begin{cases} (M+m)\ddot{x} + ml\ddot{\theta}\cos\theta - ml\dot{\theta}^2\sin\theta = F \\ l\ddot{\theta} + \ddot{x}\cos\theta + g\sin\theta = 0 \end{cases}, \tag{1}$$

where x, θ, F denote the trolley position, the rope swing angle, and the driving force applied to the trolley, respectively. Parameter $g = 9.8m/s^2$ is the gravity acceleration.

We first transform continuous-time differential equations to discrete-time difference equations here to make it easier to apply in overhead crane system. The sampling period is chosen as h. By using Euler Formula, (1) can be discretized as

$$\begin{cases} x_1(t+1) = x_1(t) + hx_2(t) \\ x_2(t+1) = x_2(t) + \dfrac{mgh}{M}x_3(t) + \dfrac{h}{M}u(t) \\ x_3(t+1) = x_3(t) + hx_4(t) \\ x_4(t+1) = x_4(t) - \dfrac{(M+m)gh}{Ml}x_3(t) - \dfrac{h}{Ml}u(t) \end{cases}, \quad (2)$$

where $x_1 = x, x_2 = \dot{x}, x_3 = \theta, x_4 = \dot{\theta}$ and $u = F$.

Define $\mathbf{x}(t) = [x_1, x_2, x_3, x_4]^T$, then (2) can be described as

$$\mathbf{x}(t+1) = A\mathbf{x}(t) + Bu(t), \quad (3)$$

where

$$A = \begin{bmatrix} 1 & h & 0 & 0 \\ 0 & 1 & \dfrac{mgh}{M} & 0 \\ 0 & 0 & 1 & h \\ 0 & 1 & -\dfrac{(M+m)gh}{Ml} & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ \dfrac{h}{M} \\ 0 \\ -\dfrac{h}{Ml} \end{bmatrix}.$$

When the overhead crane is operated during a finite time interval $t \in [0,T]$ repetitively, the system (3) can be described as

$$\mathbf{x}(t+1,k) = A\mathbf{x}(t,k) + Bu(t,k), \quad (4)$$

where k denotes iteration number.

In practical system, the desired trajectory $\mathbf{x}_d(t)$ of overhead crane is predefined. Hence, the control objective is to seek the input signal $u(t,k)$ that drives $\mathbf{x}(t,k)$ to track $\mathbf{x}_d(t)$ when $k \to \infty$.

III. ILC LAW DESIGN

For overhead crane system (4), we can give the following assumptions.

Assumption 1: The reinitialization condition is satisfied throughout the repeated iterations, i.e.,

$$\mathbf{x}(0,k) = 0,$$

for all k, where $\mathbf{x}(0,k) = [x(0,k), \dot{x}(0,k), \theta(0,k), \dot{\theta}(0,k)]^T$ and $x(0,k)$ is initial position, $\dot{x}(0,k)$ is initial velocity, $\theta(0,k)$ is initial rope swing angle and $\dot{\theta}(0,k)$ is initial rope swing angular velocity.

Assumption 2: For a desired trajectory $\mathbf{x}_d(t)$, there exists such an appropriate control $u_d(t)$, which can drive the overhead crane to track their desired trajectory for system (4) over the whole finite interval.

Assumption 3: The position and velocity of the trolley and the rope swing angle and velocity are all measurable.

Remark 1: For each operation, overhead crane often start at the same position with the initial velocity is 0, and the initial rope swing angle and angular velocity are also 0, which means assumption 1 is satisfied. Assumption 2 means that task assigned for control should be feasible. In practical system, the trolley position and the rope swing angle can be measured by selecting appropriate sensors. Hence, the Assumption 3 is satisfied.

In this paper, the iterative learning control law for overhead crane is constructed as follows:

$$\begin{aligned} u(t,k+1) = u(t,k) + K_1[\mathbf{x}(t,k+1) - \mathbf{x}(t,k)] \\ + K_2 e(t+1,k), \end{aligned} \quad (5)$$

where $e(t,k) = \mathbf{x}_d(t) - \mathbf{x}(t,k)$ is the tracking error.

The system (4) and (5) is essentially a 2D system. We can use the 2D analysis approach to ILC to derive an expression.

Define $\Delta u(t,k) = u(t,k+1) - u(t,k)$, from (5), we have

$$\Delta u(t-1,k) = K \begin{bmatrix} \eta(t,k) \\ e(t,k) \end{bmatrix}, \quad (6)$$

where $\eta(t,k) = \mathbf{x}(t-1,k+1) - \mathbf{x}(t-1,k), K = [K_1 \quad K_2]$.

Using (3), we can obtain

$$\begin{aligned} e(t,k+1) - e(t,k) &= \mathbf{x}(t,k) - \mathbf{x}(t,k+1) \\ &= -A\eta(t,k) - B\Delta u(t-1,k), \end{aligned} \quad (7)$$

and

$$\begin{aligned} \eta(t+1,k) &= \mathbf{x}(t,k+1) - \mathbf{x}(t,k) \\ &= A\eta(t,k) + B\Delta u(t-1,k). \end{aligned} \quad (8)$$

Equ. (7) and (8) can be rewritten as a class of 2D Roesser system as follows.

$$\begin{bmatrix} \eta(t+1,k) \\ e(t,k+1) \end{bmatrix} = \bar{A} \begin{bmatrix} \eta(t,k) \\ e(t,k) \end{bmatrix} + \bar{B}\Delta u(t-1,k), \quad (9)$$

where

$$\bar{A} = \begin{bmatrix} A & 0 \\ -A & I \end{bmatrix}, \bar{B} = \begin{bmatrix} B \\ -B \end{bmatrix}.$$

Substituting (6) into (9), we can present the ILC system as a class of 2D systems as follows

$$\begin{bmatrix} \eta(t+1,k) \\ e(t,k+1) \end{bmatrix} = (\bar{A} + \bar{B}K) \begin{bmatrix} \eta(t,k) \\ e(t,k) \end{bmatrix} = A_c \begin{bmatrix} \eta(t,k) \\ e(t,k) \end{bmatrix}, \quad (10)$$

where $A_c = \bar{A} + \bar{B}K,$

In this case, the problem of ILC design for overhead crane system can be transformed as the problem of design a feedback controller for 2-D system (9) such that the resulting closed-loop system (10) is asymptotically stable.

Remark 2: To this end, ILC design for overhead crane systems can be transformed the problem of feedback control for 2-D system. Hence, the existing ILC designs [12,13] can be used, even considering complex crane models with uncertainties, time-delays or H_∞ performance. In this paper, we give a novel design different from the existing methods [12,13].

We first give the following lemmas.

Lemma 1. If there exists a positive definite block diagonal matrix

$$Q = Q^T = Q_h \oplus Q_v = \begin{bmatrix} Q_h & \\ & Q_v \end{bmatrix},$$

satisfying

$$A_c^T Q A_c - Q < 0,$$

then 2-D discrete closed-loop system (10) is asymptotically stable.

Lemma 2. Given a symmetric matrix $\Theta \in R^{n\times n}$ and two matrices M, N of column dimension n, there exists a W such that the following condition holds

$$\Theta + MWN^T + NW^T M^T < 0,$$

if and only if the following projection inequalities with respect to W are satisfied

$$M_\perp^T \Theta M_\perp < 0, N_\perp^T \Theta N_\perp < 0,$$

where $M_\perp$ and $N_\perp$ denote arbitrary bases of the nullspaces of M and N, respectively.

Theorem 1: Consider the 2-D discrete-time system (10). If there exist a block-diagonal matrix $P = diag(P_h, P_v)$ with $P_h \in R^{n\times n}$ and $P_v \in R^{l\times l}$ and matrices S such that

$$\begin{bmatrix} -\frac{1}{2}(S+S^T) & S^T A_c^T & \frac{1}{2}S+S^T-P \\ A_c S & -P & -A_c S \\ \frac{1}{2}S^T+S-P & -S^T A_c^T & -S-S^T \end{bmatrix} < 0, \tag{11}$$

then the closed-loop system (10) is asymptotically stable.

Proof. Matrix inequality (11) can be rewritten as follows:

$$\begin{bmatrix} 0 & 0 & -P \\ 0 & -P & 0 \\ -P & 0 & 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{2}I \\ -A_c \\ -I \end{bmatrix} S \begin{bmatrix} -I & 0 & I \end{bmatrix} + \begin{bmatrix} -I \\ 0 \\ I \end{bmatrix} S^T \begin{bmatrix} \frac{1}{2}I & -A_c^T & -I \end{bmatrix} < 0. \tag{12}$$

Let $M^T = \begin{bmatrix} \frac{1}{2}I & -A_c^T & -I \end{bmatrix}, N^T = \begin{bmatrix} -I & 0 & I \end{bmatrix}$,

$$\Theta = \begin{bmatrix} 0 & 0 & -P \\ 0 & -P & 0 \\ -P & 0 & 0 \end{bmatrix},$$

and select the orthogonal complements of M and N ($M_\perp$ and $N_\perp$ respectively) as:

$$M_\perp = \begin{bmatrix} I & 0 \\ 0 & I \\ \frac{1}{2}I & -A_c^T \end{bmatrix}, N_\perp = \begin{bmatrix} I & 0 \\ 0 & I \\ I & 0 \end{bmatrix}.$$

Applying Lemma 2 we can obtain:

$$M_\perp^T \Theta M_\perp < 0, N_\perp^T \Theta N_\perp < 0, \tag{13}$$

that is

$$M_\perp^T \Theta M_\perp = \begin{bmatrix} -P & PA_c^T \\ A_c P & -P \end{bmatrix} < 0, \tag{14}$$

and

$$N_\perp^T \Theta N_\perp = \begin{bmatrix} -2P & 0 \\ 0 & -P \end{bmatrix} < 0, \tag{15}$$

Note that the condition (14) and (15), Using Schur Complement we have

$$PA_c^T P^{-1} A_c P - P < 0. \tag{16}$$

Thus, multiplying the right and the left of condition (16) by P^{-1} and taking $Q = P^{-1}$ in the resulting inequality we obtain

$$A_c^T Q A_c - Q < 0,$$

Using Lemma 1, it is obvious that the asymptotic stability of closed-loop system (10) is guaranteed. This completes the proof.

Using theorem 1 and $A_c = \bar{A} + \bar{B}K$, the synthesis of feedback controllers can be described to find (K, P, S) such that the following matrix inequality holds

$$\begin{bmatrix} -\frac{1}{2}(S+S^T) & S^T(\bar{A}+\bar{B}K)^T & \frac{1}{2}S+S^T-P \\ (\bar{A}+\bar{B}K)S & -P & -(\bar{A}+\bar{B}K)S \\ \frac{1}{2}S^T+S-P & -S^T(\bar{A}+\bar{B}K)^T & -S-S^T \end{bmatrix} < 0. \tag{17}$$

It is difficult to solve the matrix equation of inequality (17) for (K, P, S) because it is non-convex. In this case, we introduce a new variable $Y = KS$, then (17) leads to a convex sufficient condition in terms of LMI.

Theorem 2: Consider the 2-D system (10), if there exist a block-diagonal matrix a positive definite block diagonal matrix $P = diag(P_h, P_v)$ with $P_h \in R^{n\times n}$ and $P_v \in R^{l\times l}$, S, Y such that

$$\begin{bmatrix} -\frac{1}{2}(S+S^T) & (\bar{A}S+\bar{B}Y)^T & \frac{1}{2}S+S^T-P \\ \bar{A}S+\bar{B}Y & -P & -(\bar{A}S+\bar{B}Y) \\ \frac{1}{2}S^T+S-P & -(\bar{A}S+\bar{B}Y)^T & -S-S^T \end{bmatrix} < 0, \tag{18}$$

then 2-D discrete closed-loop system (10) is asymptotically stable. In this situation, a suitable control law for (6) is given by $K = YS^{-1}$.

Remark 3: Compared to the existing designs [12,13], the proposed design approach with the simple linear matrix inequality (18) obtains less computation, and the LMI condition also provides a direct approach to determine learning gains.

IV. SIMULATIONS STUDY

In this section, a simulation is performed for an overhead crane system to verify our conclusions. The sample time is chosen as $h = 0.01s$, and the system parameters are given as follows:

$$M = 5kg,\ m = 10kg,\ l = 2m,\ g = 9.8m/s^2.$$

The desired position of the crane gantry is selected as

$$x_d(t) = \begin{cases} 0.1t^2 & 0 \le t \le 5 \\ t - 2.5 & 5 < t \le 10 \\ -12.5 + 3t - 0.1t^2 & 10 < t \le 15 \end{cases},$$

For the initial state, it is assumed that $x_1(0,k) = x_2(0,k) = x_3(0,k) = x_4(0,k) = 0$ for all k. The ILC law is applied by adopting the zero initial control input $u(t,0) = 0$ for all t. Using overhead crane system parameters in (3), the 2D system matrix in (9) can be calculated.

Using Theorem 2, with the help of Matlab LMI toolbox, the LMI (18) is feasible and a gain matrix K is also derived. The results of simulation are shown in Fig. 2-Fig 4. Fig. 2 gives system output at the 10^{th}, 30^{th}, 100^{th} iterations and the desired output trajectory respectively. Fig. 3 gives the max position tracking error on the iteration domain. Fig. 4 gives the rope swing angle at 100^{th} iteration. It can be observed that by using the proposed iterative learning control law, we can guarantee the trolley position converges to desired trajectory for all $t \in [0,15]$, and the maximum swing angle is smaller than 0.035 radian (about 2 degree). Obviously, the rope swing angle by the ILC controller is much smaller than other controllers [4-7].

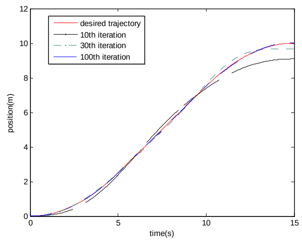


Fig. 2 Position tracking profiles

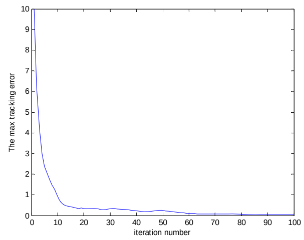


Fig. 3 Position tracking errors on iteration domain.

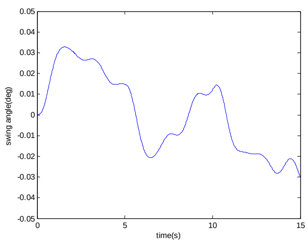


Fig. 4 Rope swing angle at 100^{th} iteration

VI. CONCLUSION

In this paper, the problem of iterative learning control for overhead crane systems has been investigated. Basing on 2D analysis of the learning process, the learning gains can be obtained by solving linear matrix inequalities. The proposed controllers can guarantee the asymptotic convergence of crane position to its desired profile with small swing angle on the whole operation time interval.

ACKNOWLEDGEMENT

This work is supported by the Program of NSFC (No. 61203065, 61240049), the program of Key Young Teacher of Henan Polytechnic University. The Doctoral Fund Program of Henan Polytechnic University (B2012-003).

REFERENCES

[1] J. Collado, R. Lozano, I. Fantoni. Control of convey-crane based on passivity. in *Proc. Amer. Control Conf.*, Chicago, IL, Jun. 2000, 1260-1264.

[2] B. Kiss, J. Levine, P. Mullhaupt. A simple output feedback PD controller for nonlinear cranes. in *Proc. IEEE Conf. Decision Control*, Dec. 2000, pp. 5097-5101.

[3] H. H. Lee. A new approach for the anti-swing control of overhead cranes with high-speed load hoisting. *Int. J. Control*, 2003, 76(15): 1493-1499

[4] Y. Fang, W. E. Dixon, D. M. Dawson, E. Zergeroglu. Nonlinear coupling control laws for an underactuated overhead crane system. *IEEE/ASME Trans. Mechatronics*, 2003, 8(5): 418-423

[5] K. A. F. Moustafa. Fuzzy control of flexible cable overhead cranes with load hoisting. *Trans. Inst. Meas. Control*, 2006, 28(4):372-386.

[6] J. Q. Yi, N. Yubazaki, K. Hirot. Anti-swing and positioning control of overhead traveling crane. *Information Sciences*. 2003, 155(1-2) :19-42

[7] D. Liu, J. Yi, D. Zhao, W. Wang. Adaptive sliding mode fuzzy control for a two-dimensional overhead crane. *Mechatronics*, 2005, 15(5): 505–522

[8] S. Arimoto, S. Kawamura, F. Miyazaki. Bettering operation of robots by learning. *J. of Robotic Systems*, 1984, 1(2):123-140.

[9] D. A. Bristow, M. Tharayil, A. G. Alleyne. A survey of iterative learning control: A learning-based method for high-performance tracking control. *IEEE Control Syst. Mag.*, 2006, 26(3): 96-114

[10] H. S. Ahn, Y. Chen, K. L. Moore. Iterative learning control: Brief survey and categorization. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, 2007, 37(6): 1099-1121

[11] B. Vikramaditya, R. Rajamani. Nonlinear control of a trolley crane system. In: Proceedings of the American control conference, Chicago, IL. 2000. 1032-1036.

[12] E. Rogers, K. Galkowski, A. Gramacki, J. Gramacki, D. H. Owens. Stability and controllability of a class of 2-D linear systems with dynamic boundary conditions. *IEEE Trans. Circuits Syst. I: Fundam. Theory Appl.*, 2002, 49(2): 181-195

[13] J. Shi, F. Gao, T. J. Wu. Robust design of integrated feedback and iterative learning control of a batch process based on a 2D Roesser system. *J. Process Control*, 2005, 15(8): 907-924.

A Hybrid Path Planning Algorithm for UGV by Combining A* and B-spline Curve Equation

Min-Ho Kim, Hee-Mu Lee and Min-Cheol Lee

Pusan National University, Busan, South Korea
Tel : 81-051-510-3081; Fax : 81-051-512-9835
Xho1995@gmail.com, slm1023@nate.com, mclee@pusan.ac.kr

Abstract: This article presents a hybrid path planning algorithm to make a smooth path for an UGV. A main theme of the proposed algorithm is to combine of the A* algorithm and the B-spline curve equation. A* is one of the well-known path planning algorithm which finds the optimal path on the given map by using the heuristic cost function. However, since A* is based on the grid map, the result path consists of straight lines with only 8 directions. It's not suitable of UGV's navigation. Therefore in this paper to overcome this issue, B-spline curve equation is proposed to make a smooth and continuous path with control points which are selected from the A* result path. And the optimal control point selection algorithm is proposed to make the hybrid path. At last, to verify the proposed algorithm, the hybrid path results are compared with A* algorithm by using a developed simulation program.

Keywords: UGV, Path Planning, A* Algorithm, B-spline Curve Equation

I. INTRODUCTION

After the development of fusion technology of mechanics, electronics and IT, there are a lot of researches about the autonomous vehicle. For example, the grand challenge has held in USA by DARPA and Google has tested an autonomous vehicle on the real road environment. [1], [2] For this kind of vehicle to move automatically, it needs the reference input point that the vehicle can follow, and this input point comes from a path planning algorithm. Therefore, researches on the path planning algorithm have increased.

Among these, A* is one of the well-known path planning algorithm developed by Peter Hart, Nils Nilsson, Bertram Raphael in 1968. [3] It is the global path planning algorithm that finds the optimal path on the given map using the heuristic cost function. Usually the map of A* algorithm consists of several grid cells, and then each cell has a cost value to be compared and evaluated to find the optimal path.

However, due to the grid map, A* algorithm gives the vehicle some reference points on the discontinuous path. Actually the result path consists of straight lines with only 8 directions. It's not suitable of real UGV's navigation. In this paper, therefore to overcome this issue, B-spline curve equation is proposed to make a smooth and continuous path with control points which are selected from the result path of A* algorithm. Using A* algorithm, the center points of selected grid cells become the way points.

In this hybrid path planning algorithm, the optimal control point selection algorithm is proposed. In the next chapter, we will explain the theory of A* algorithm and the B-spline curve equation respectively to plan the path. And we will show the A* and B-spline curve results by using the developed program. We developed the A* simulation program and the continuous path generation algorithm using B-spline interpolation equation in windows environment. [4], [5], [6] Next, the optimal control point selection algorithm will be explained. At last, to verify the proposed algorithm, we have developed the hybrid path simulation program and the path results are compared with A* algorithm to evaluate the proposed algorithm.

II. A* ALGORITHM

1. A* Algorithm and Cost Function

Before introducing the hybrid path planning algorithm, first we look around the normal A* algorithm. The A* algorithm uses the cost function like as; [3]

$$F(n) = g(n) + h(n) \tag{1}$$

In this equation, g(n) is the cost value from the start position to current position on the map. And h(n) is the heuristic cost value from the current position to the goal position. F(n) is the total cost value to be compared and evaluated for optimal path planning.

Fig.1. shows one of the result path of A* algorithm. [7] The path consists of several discontinuous lines, so if we want to use the result path into the real UGV, it should stop at the each corners to change the heading angle. In addition, a real UGV can move to any direction, the simple A* algorithm may not be effective because of the path direction limit of the A* path.

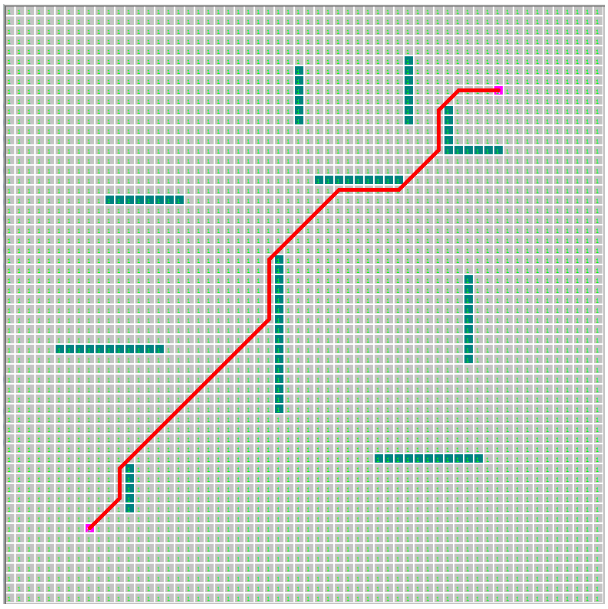


Fig.1. A Path generated with A* Algorithm

III. THE B-SPLINE EQUATION

1. The B-spline Curve Equation

To specify a path, a B-spline curve was used [8], [9], [10]. The B-spline curve can be generated by a blending function, which connects and makes the vertices of a polygon smooth. Cox and de Boor suggested a blending function shown as in equation (1);

$$N_{i,k}(u) = \frac{(u-t_i)N_{i,k-1}(u)}{t_{i+k-1}-t_i} + \frac{(t_{i+k}-u)N_{i+1,k-1}(u)}{t_{i+k}-t_{i+1}}$$
$$N_{i,1}(u) = \begin{cases} 1 & t_i \le u \le t_{i+1} \\ 0 & others \end{cases} \quad (1)$$

Where, t_i is called as knot value, which is a parameter that makes blending function nonzero. The knot value can be divided as periodic knots and non-periodic knots. In this paper, non-periodic knots shown in equation (2) were used.

$$t_i = \begin{cases} 0 & (0 \le i < k) \\ i-k+1 & (k \le i \le n) \\ n-k+2 & (n < i \le n+k) \end{cases} \quad (2)$$

Using the above blending function, B-spline curve can be defined as;

$$\mathbf{P}(u) = \sum_{i=0}^{n} \mathbf{P}_i N_{i,k}(u) \quad (t_{k-1} \le u \le t_{n+1}) \quad (3)$$

Where, each $\mathbf{P}_i$ means a pre-defined control point.

Since a blending function is difficult to express as a polynomial, Cox and de Boor suggested an algorithm to divide control points. According to their algorithm, the equation for the curve can be written as the following function $\mathbf{P}(u)$ as;

$$\mathbf{P}(u) = \sum_{i=l-k+1}^{l} \mathbf{P}_i N_{i,k}(u) \quad (4)$$

Combining equation (2) and (4), the following relation can be derived.

$$\mathbf{P}(u) = \sum_{i=l-k+2}^{l} \mathbf{P}_i^1 N_{i,k-1}(u)$$
$$\mathbf{P}_i^1 = \frac{u-t_i}{t_{i+k-1}-t_i}\mathbf{P}_i + \left(1-\frac{u-t_i}{t_{i+k-1}-t_i}\right)\mathbf{P}_{i-1} \quad (5)$$

In this case, the point $\mathbf{P}_i^1$ divides the line $\mathbf{P}_i\mathbf{P}_{i-1}$. Repeating this process, the following relation can be obtained.

$$\mathbf{P}(u) = \sum_{i=l-k+r+1}^{l} \mathbf{P}_i^r N_{i,k-r}(u)$$
$$\mathbf{P}_i^r = \frac{u-t_i}{t_{i+k-r}-t_i}\mathbf{P}_i^{r-1} + \left(1-\frac{u-t_i}{t_{i+k-r}-t_i}\right)\mathbf{P}_{i-1}^{r-1} \quad (6)$$

The point $\mathbf{P}_i^r$ is called as de Boor. Plugging (k-1) to r, the following relation can be obtained.

$$\mathbf{P}(u) = \mathbf{P}_l^{k-1} \quad (7)$$

Following the above sequence, a curve $\mathbf{P}(u)$ could be obtained from the given points $\mathbf{P}_i$. Fig. 2 shows a curve path obtained from the B-spline curve equation. In the fig. 2 the dark points are the control points which make the B-spline curve. Therefore, to build a hybrid path, a control point selection algorithm is needed.

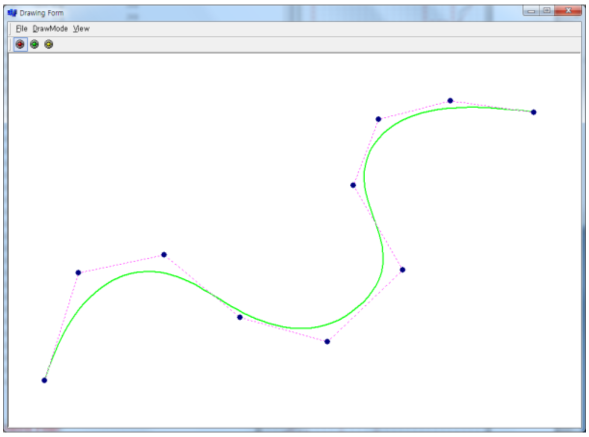


Fig.2. The Path Generated with B-spline curve equation

IV. HYBRID PATH PLANNING ALGORITHM

1. Control Point Selection Algorithm

The hybrid path is the continuous curved path which is made by combining the A* path points and the B-spline curve equation. To make this path, some points on the A* path nodes should be selected as the control points of B-spline curve equation. In this paper, we proposed a control point selection algorithm such as;

1. Select the start point on the A* path as a first control point
2. Go to the next node point on the A* path
3. If there isn't any obstacle between two points, go to the next node
4. If there is an obstacle, then select the previous node as a next control point
5. Repeat 2~4 routine, until the goal point

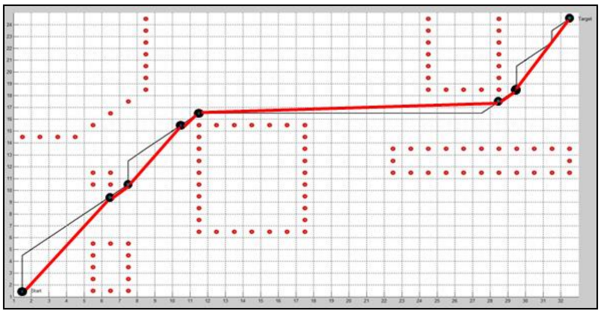


Fig.3. The Selected Points with the Proposed Algorithm

In the Fig. 3, black points are the selected points to be the control points of B-spline curve, by using those points the hybrid path is generated.

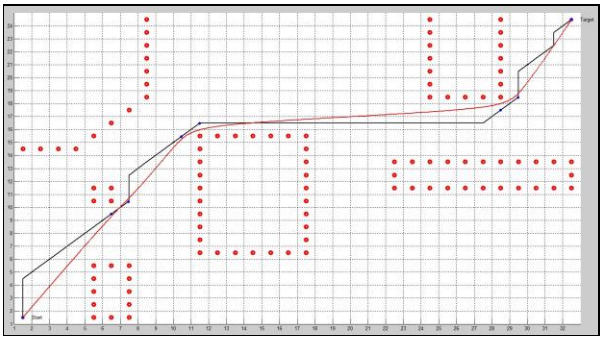


Fig.4. A B-spline Curve Path with Selected Points

Fig. 4 shows the hybrid path result. In this Fig. the black lines mean an A* path, and the red curve means a hybrid path using B-spline curve equation. B-spline curve are always in the polygon of the control points, the hybrid path will not collide the obstacles.

V. SIMULATION & RESULT

1. Path Tracking Simulation

To evaluate the proposed algorithm, we developed the path tracking algorithm. By using this program, the tracking result is compared with the A* path. In this paper, we build a 32x24 grid map with 1m x 1m grid size like as in Fig. 5., and the maximum vehicle speed is limited in 0.6m/s due to the limit of a real mobile UGV robot. [6]

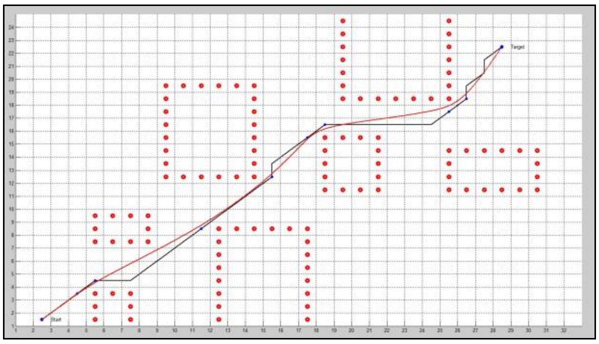


Fig.5. A* and Hybrid Path Tracking Simulation

2. The Result

The result is shown as table 1. In this table, total distance and tracking time is reduced within the hybrid path. Especially, while total distance is reduced as 5.5%, tracking time is reduced as 34.5%. It means, on the hybrid path, the UGV has no need to reduce the speed in the corner. Therefore, the hybrid path is more effective for the real UGV's navigation.

Table 1. A* and Hybrid Path Tracking Result

	A*	Hybrid	Improved
Total Distance (m)	36.5	34.5	5.5 %
Tracking Time (s)	95.4	62.5	34.5 %

VI. CONCLUSION

In this paper, we suggested the hybrid path planning algorithm by combining the A* algorithm and the B-spline curve equation. For this we proposed the control point select algorithm. With this algorithm, the continuous path can be generated with the control points which are from the A* algorithm. To evaluate the proposed algorithm, we developed the path planning and tracking simulation program. And the result shows the hybrid path is effective for the real UGV's navigation.

VII. ACKNOWLEDGMENTS

This research was supported by the MOTIE(The Ministry of Trade, Industry and Energy), Korea, under the Human Resources Development Program for Special Environment Navigation/Localization National Robotics Research Center support program supervised by the NIPA(National IT Industry Promotion Agency)." (H1502-13-1001)

REFERENCES

[1] Wikipedia (2013), DARPA Grand Challenge, http:// en.wikipedia.org/wiki/DARPA_Grand_Challenge

[2] Wikipedia (2013), Google driverless car, http:// en.wikipedia.org/wiki/Google_car

[3] P. E. Hart, N. J. Nilsson and B. Raphael (1968), A Formal Basis for the Heuristic Determination of Minimum Cost Paths. IEEE transactions of systems science and cybernetics, Vol. ssc-4, No.2:100-107

[4] M. H. Kim and M. C. Lee(2010), A Path Generation Method for Path Tracking Algorithms that use the Augmented Reality, International Conference on Control, Automation and Systems:1487-1490

[5] M. H. Kim, H. M. Lee, Y. Wei and M. C. Lee (2012), A Study of New Path Planning Algorithm Using Extended A* Algorithm with Survivability, Intelligent Robotics and Applications, Vol. 7508(59):608-617

[6] C. B. Noh, M. H. Kim and M. C. Lee (2013), 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(1):043-050

[7] M. H. Kim, C. B. Noh, J. H. Heo and M. C. Lee (2013), A Study of Path Planning Algorithm Based on the Survival Probability, Intelligent Autonomous Systems 12, Vol.193:755-763

[8] M. G. Cox (1972), The Numerical Evaluation of B-spline, J. of Approx. Theory, Vol. 6:95-108

[9] C. de Boor (1997), On Calculation with B-spline, J. of Approx. Theory, Vol. 6:76-78

[10] K. W. Lee (1999), Principles of CAD/CAM/CAE systems, Addson Wesley

The actuator device design for the implementation of haptic joystick

Dong-hyuk Lee, Sun-kyun Kang, Bo-Yeon Hwang, Ki-jung Kim, and Jang-myung Lee

Dep. of Electronic Engineering, Pusan National University, Korea
Tel : 82-51-510-7379; Fax : 82-51-514-1693
{ldh0917, sunkyun7379, boyeon1696, kijung7379, jmlee}@pusan.ac.kr

Abstract: In this paper, we designed the actuator device for haptic implementation in the general joystick. Each linear servo actuator as pillars to support the joystick mounting plate can move up and down. And this movement will generate the inclination of the plate. The haptic can be implemented by the change of the inclination. The actuator device was designed through the performance analysis of actuator device structure. The haptic will represent information from obstacle cognition device by using the designed actuator device. The haptic performance of the actuator device was verified through experiments.

Keywords: Haptic Device, Actuator Device, Tele-operation, Ultrasonic sensor, Haptic Joystick, Linear actuator control

I. INTRODUCTION

Recently, with the development of robotics industry, need in the field of remote control is being magnified. This remote control can be used in a hazardous environment that human is difficult to access directly and determine the current situation, such as the scene of the accident, cleaning the inside of the nuclear reactor and the space exploration[1]. Tele-operation, as a field of the control that users perceive and determine their environment at the outside by using a variety of sensor information attached to the robot, the technology of receiving feedback exactly from sensor information is essential. Among them, the research that users can take action and grasp the information by using the haptic joystick is in progress actively [2]. The haptic device is an essential element for the implementation of the haptic. However, the current developed haptic devices were expensive products and were produced by a haptic joystick. So, it was difficult to integrate easily into a general joystick. In this paper, a haptic joystick that a common joystick equipped with a haptic device consisting of a linear servo actuator was developed, and basic research of verifying the performance of haptic device is proceed.

In this paper, haptic device that grafts haptic system onto general joystick was designed. Haptic device was composed of linear servo actuator used as a pillar to support the joystick mounting plate, and was designed to be the optimal structure using MATLAB simulation. We implement the haptic system using the change of the mounting plate's slope caused by a linear servo actuator's up-and-down motion. To verify the performance of the designed haptic device, the haptic joystick and obstacle cognition device was constructed Tele-operation system. Obstacle cognition device obtains precise location information of obstacles by superposing the ultrasonic sensor. By using the obtained obstacle information, through the implementation of the haptic device, user can be aware of the situation.

II. Haptic device design

1. Actuator device design

Fig.1. Linear servo actuator

Figure 1 shows a linear servo actuator used to design a haptic device. In this paper, the purpose is a development of a haptic device using a linear servo actuator. Linear servo actuator is a device that an inside DC motor converts the rotary motion into the linear motion through the ball screw, which enables the linear motion. In this paper, using the linear servo actuator as a pillar of the joystick holder plate, the linear movement haptic device was designed.

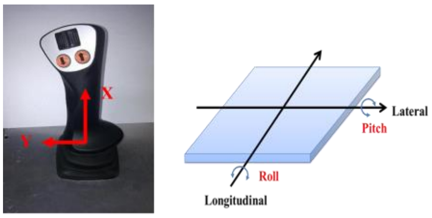


Fig.2. Joystick coordinate system and rotating coordinate system of haptic device

Figure 2 shows joystick coordinate system and rotating coordinate system of haptic device. In this paper, through the linear motion of a linear servo actuator, the haptic devices has the rotary movement of the joystick mounting plate was developed. And through the rotary movement, users could get the obstacle avoidance information. Like Figure 2, joystick fixing plate has movement of 2 degree of freedom through rotation of Roll and pitch.

2. MATLAB simulation

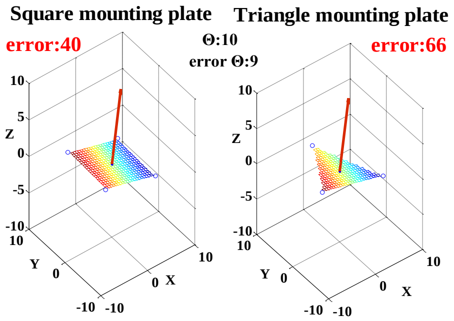


Fig.3. D.O.F of haptic device simulation

Table 1. Parameters of Figure 3

θ	Slope of haptic device
error θ	The slope of the haptic device, the first occurrence of an error
error	The haptic device is rotated, the number of angles that cannot be expressed

Figure 3 is simulation result for optimizing degree of freedom of haptic device. In this paper, haptic device is designed optimally through simulation using MATLAB. Haptic device is simulated considering two kinds of structure (Triangles and squares). Parameter of figure 3 is equal to table1.

Impossible to express the number of angles were measured during rotation according inclination of haptic device. considering encoder resolution of miniature linear servo actuators to be used for haptic device design, rotation is measured by increasing change of inclination from 0° to 10° of 1° and rotation is measured from 0° to 360° of 10°. In result of simulation impossible to express the angles are more found in triangle structure than square structure. Accordingly, triangle structure is more proper structure than square as haptic device.

III. Haptic Interface

1. Entire system Configuration

Figure 4 is entire system configuration. Haptic joystick and obstacle cognition device transmit data using Bluetooth. Obstacle cognition device estimate the distance and the angle with obstacles using ultrasonic sensor, the haptic is realized by controlling haptic device using the measured data [3].

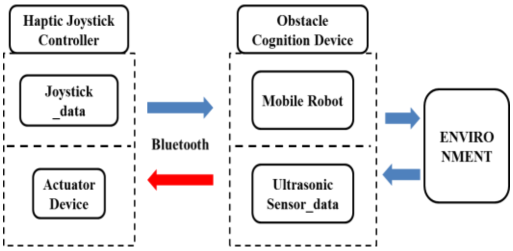


Fig.4. Entire system

2. Haptic joystick configuration

Fig.5. Haptic joystick

Figure 5 is proposed haptic joystick in this paper. Haptic joystick is configured on haptic joystick and joystick controller. Haptic joystick is configured to attach haptic device designed on the basis of Chapter 2 at joystick. Haptic device was designed by considering to implement haptic joystick combined with a variety of devices.

3. Obstacles cognition device configuration

A. Obstacles cognition device

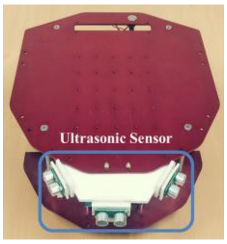


Fig.6. Obstacles cognition device

Figure 6 is obstacles cognition device. Obstacles cognition device was configured by equipped with ultrasonic wave sensor to mobile robot. Three ultrasonic sensors was equipped with the front of the mobile robot, it recognizes obstacles using ultrasonic sensor. Through arrangement of such ultrasonic sensor identify location information of obstacle in surround environment of mobile robot and it is delivered to joystick controller.

B. Superposition of ultrasonic sensor to estimate the position of obstacles

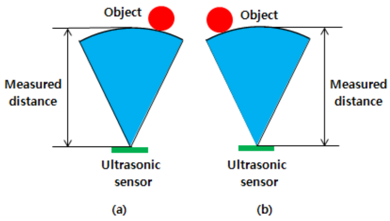


Fig.7. Position uncertainty of ultrasonic sensor

As shown in figure 7 (a) and (b), distance to obstacles in beam width is can measured, but difficulty is to know the exact position of the obstacle, this calls position uncertainty of ultrasonic sensor [4]. In order to reduce the uncertainty of locations, we nested three ultrasonic sensors to be able to recognize more precisely the location of obstacles [5].

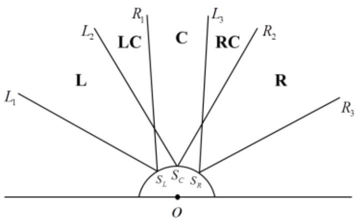


Fig.8. Position uncertainty reduction due to beam overlap

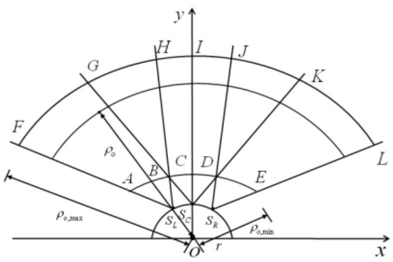


Fig.9. Three adjacent overlapped ultrasonic sensor

Figure 8 shows the ultrasonic sensors S_L, S_C and S_R which have cross beam width three nested on the circle radius is not zero. Across the full detection area of the ultrasonic sensor in the center S_C is divided three small areas (LC, C, RC) because of nested beam width of the ultrasonic sensor. The combination of ultrasonic sensors for detecting obstacles depending on the relative position of the obstacle in the beam width of the ultrasonic sensor is changed. As a result, It is possible to subdivide zones obstacle is present through combination of ultrasonic sensors for detecting obstacles. Thus it becomes possible to reduce the uncertainty of the position of the unique ultrasonic sensor. Ultrasonic sensor provides a distance of obstacle which is based on the vertex of its own each, but it has to be changed to a distance of obstacle which is based on the benchmark O of nested ultrasonic for the implementation of avoidance and obstacle detection.

The obstacle detection range of the nested ultrasonic sensor is determined by Equation (1) in consideration of the obstacle detection square area shown in Figure 9.

$$\rho_{o,\min} \le \rho_o \le \rho_{o,\max} \tag{1}$$

Here,

$$\rho_{o,\min} = \overline{OD} = d \tag{2}$$

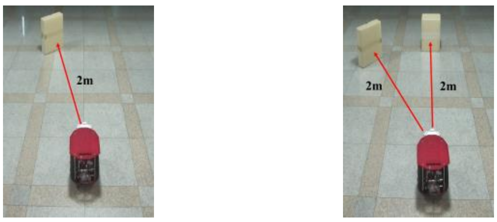
$$\rho_{o,\max} = \overline{OI} = r + \rho_{s,\max} \tag{3}$$

$\rho_{o,\min}$ in equation (2) is distance from the center of the nested ultrasonic sensor ring to the vertex of the obstacle detection square area. $\rho_{o,\max}$ in equation (3) is the distance of the obstacle on the basis of the center of ultrasonic sensor ring corresponding to the maximum detection distance of the ultrasonic sensor $\rho_{s,\max}$ [6].

IV. Experiments and Discussion

1. The experimental environment

Figure 10 is experimental environment for the performance analysis about the proposed haptic device. This experiment was carried out in the second floor lobby, Samsung University-Industrial Alliance Building in Pusan University. And two cases were assumed for performance analysis of the haptic device.



(a) Experiment 1 (b) Experiment 2

Fig.10. The experimental environment

2. The experimental results

Figure 11 is a performance analysis of the haptic device in the experiment 1. In case of that the obstacle is present on the left side of the obstacles cognition device, the performance of the haptic device is analyzed in accordance with situation that it approaches to the obstacle. In section ①, the obstacles cognition device recognizes the position of the obstacle through the nesting of the ultrasonic sensor, and haptic device was delivered degree of risk and direction to avoid to the user according to the situation. In section ②, the haptic device operates in accordance with the situation obstacles cognition device passes the obstacle closed to the left side of it.

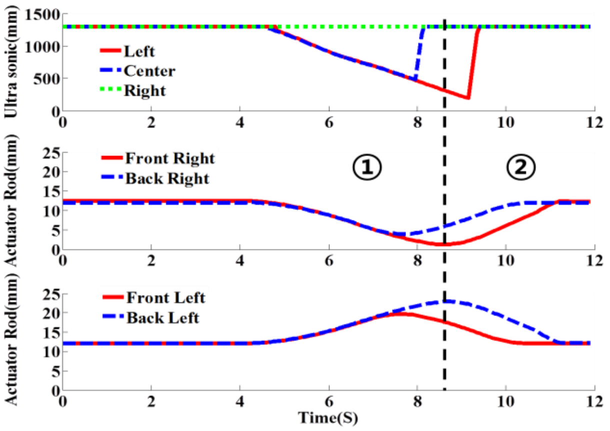


Fig.11. The performance analysis depending on recognition of obstacle (Experiment 1)

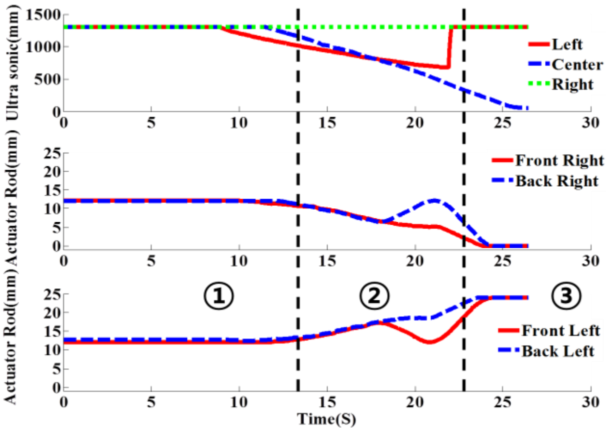


Fig.12. The performance analysis depending on recognition of obstacle (Experiment 2)

Figure 12 is a performance analysis of the haptic device in the experiment 2. Depending on the case where an obstacle is present in two locations left and center of the obstacles cognition device, in the same manner as described above and analyzed the performance of the haptic device. In section ①, the obstacles cognition device recognizes first obstacles present on the left side and haptic device was delivered degree of risk and direction to avoid to the user according to the situation. In section ②, the obstacles cognition device accurately recognizes the position of the two obstacles and haptic device was delivered degree of risk and direction to avoid to the user according to the situation. In section ③, the obstacles cognition device recognized the only obstacle in the middle through passing the obstacle on the left.

V. CONCLUSION

In this paper, it is proposed that haptic device using linear servo actuator. Haptic device was designed to be the optimal structure using MATLAB simulation. And it was verified the performance of the haptic device through experiments on a variety of situations. In the future, it is plan to research an algorithm for traveling of the mobile robot based on the position information of the obstacle the user recognized.

Acknowledgements

"This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the Human Resources Development Program for Special Environment Navigation/Localization National Robotics Research Center support program supervised by the NIPA(National IT Industry Promotion Agency)." (NIPA-2012-H1502-12-1002)

"This research was financially supported by the Ministry of Trade, Industry&Energy(MOTIE), Korea Institute for Advancement of Technology(KIAT) through the Inter-ER Cooperation Projects."

REFERENCES

[1] H. S. Kim and H. J. Kang (2005), Design of Remote Manipulator Control System using PHANToM Device. Power Electronics Conference 2005: 595-597
[2] H. Kim and J. H. Ryu (2012), The Effect of Asynchronous Haptic and Video Feedback on Teleoperation and a Comment for Improving the Performance. Institute of Control, Robotics and Systems 18(2): 156-160
[3] A. K. Ko, J. Y. Choi, H. C. Kim and J. M. Lee (2007), A Haptic Interface Using a Force-Feedback Joystick. Institute of Control, Robotics and Systems 13(12): 1207-1212
[4] Y. Choi, W. Choi, and J. Song (2011), Obstacle avoidance of a mobile robot using low-cost ultrasonic sensors with wide beam angle. Institute of Control, Robotics and Systems 17(2): 79-87
[5] S. Kim and H. Kim (2012), Comparative analysis on performance indices of obstacle detection for an overlapped ultrasonic sensor ring. Institute of Control, Robotics and Systems 18(4): 321-327
[6] S. Kim and S. Lee (2009), Positional uncertainty reduction of overlapped ultrasonic sensor ring for efficient mobile robot obstacle detection. Institute of Signal Processing and Systems 10(3): 198-206

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Design of Fuzzy Controller using Variable Fuzzy Membership Function Factors for Inverse Ball Drive Mobile Robot

Keon-woo Jeong[1], Shin-nyeong Heo, Seung-Ik Hwang, Han-Dong Yoo and Jang-myung Lee[2]

[1] Department of Electrics Engineering, Pusan National University, Busan, 616-111, Korea
Tel : +82-51-510-7379; Fax: +82-051-514-1693
E-mail: asie4u2@pusan.ac.kr
[2] Department of Electrics Engineering, Pusan National University, Busan, 616-111, Korea
Tel : +81-51-510-2378; Fax: +82-051-514-1693
E-mail: jmlee@pusan.ac.kr

Abstract: In this paper, a fuzzy controller for a inverse ball drive mobile robot is implemented to have more stable balancing capability than the conventional control system. Fuzzy control structure is chosen for a inverse ball drive mobile robot, and fuzzy membership function factors for the control system are obtained for 3 specified weights using a trial-and-error method. Next a linear Interpolation method is employed to generate fuzzy membership function factors for more stable control performance when the weight is arbitrarily selected. Through some experiments, we find that the proposed fuzzy controller using the neural network is superior to the conventional fuzzy controller.

Keywords: Inverse ball drive mobile robot, fuzzy.

I. INTRODUCTION

Recently, the interest of humen-robot interaction (HRI) is growing and many robot engineers are researching about communication ability of mobile robot with person and the path planning has been carried out. If there is one problem, mobile robot's rotation in particular place or driving to particular angle are difficult caused by the narrow space at real indoor environments. So, the new form of mobile robot that can take the place of it is required. Among the several attempts, the Inverse ball drive mobile robot meets these requirements. In this paper, the algorithm that based on general mamdani-style fuzzy, it converts to the values of the elements in the appropriate purge affiliation function depending on the load to force on the robot is proposed. When a robot comes to perform the task, its total weight is added by luggage weight. According to the changing of load, the total weight affects the balance [1]. As a method to complement this problem, a control system to obtain the value of the appropriate elements depending on the weight of the user by using a linear interpolation to obtain the value of an element belongs function of the fuzzy control system is provided. In order to confirm the performance of the control system, it is compared with typical of the fuzzy control system and the performance of the fuzzy control system proposed by using the MATLAB and Simulink.

II. SYSTEM MODELING

The section aims to describe the system structure and physical configuration of the designed inverse ball drive mobile robot, and then derive its dynamic model using Lagrangian mechanics. The modeling process is based on a simplification assumption that the inverse ball drive mobile robot is constructed by two major components: the body and the ball, as Fig. 1 shows.

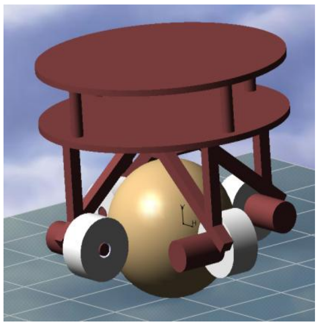


Fig.1. Inverse ball drive mobile robot

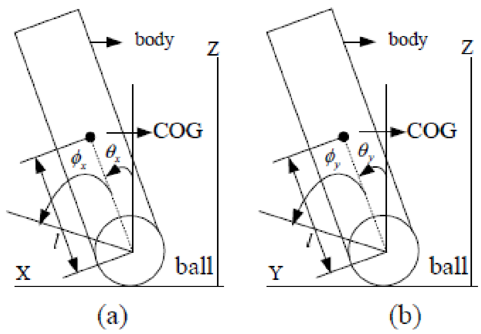


Fig.2. Simplified diagrams of the simplified robot

Consider the uncertainty term of the system parameter, the Euler-Lagrange equations are written by the following matrix form [2, 3].

$$M_{x0}(q_x)\ddot{q}_x + C_{x0}(q_x,\dot{q}_x)\dot{q}_x + G_{x0}(q_x) + D_{x0}(\dot{q}_x) + U_x = \begin{bmatrix} 0 & \tau_x \end{bmatrix}^T \quad (8)$$

where U_x is the lumped of uncertainty defined as $U_x = \Delta M_x \ddot{q}_x + \Delta C_x \dot{q}_x + \Delta G_x(q_x) + \Delta D_x = \begin{bmatrix} U_1 & U_2 \end{bmatrix}^T$ and the lumped of uncertainty is assumed to be bounded with a finite value. By defining the following state vector $X_X = \begin{bmatrix} x_{x1} & x_{x2} & x_{x3} & x_{x4} \end{bmatrix}^T$, that is $x_{x1} = \theta_x$, $x_{x2} = \phi_x$, $x_{x3} = \dot{\theta}_x$ and $x_{x4} = \dot{\phi}_x$ one obtains the subsequent nonlinear state equations

$$\dot{X}_x = \begin{bmatrix} \begin{bmatrix} x_{x3} \\ x_{x4} \end{bmatrix} \\ M_{x0}(x_X)^{-1}\left(\begin{bmatrix} 0 \\ u_x \end{bmatrix} - \alpha - \beta - \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \right) \end{bmatrix}$$

$$\alpha = C_{x0}(x_X, \dot{x}_X)\dot{x}_X + G_{x0}(x_X) \quad (9)$$

$$\beta = \begin{bmatrix} 0 \\ \mu_{\phi x0} x_{x4} + \mu_{\phi x0}\,\mathrm{sgn}(x_{x4}) \end{bmatrix}$$

where $u_x = \tau_x$. Equation (9) represents the SIMO dynamic equation of the robot in the median sagittal plane.

Similar to the previous derivation, the Euler-Lagrange equations of the robot in the median coronal plane can be derived and governed by the following matrix form

$$M_{y0}(q_y)\ddot{q}_y + C_{y0}(q_y,\dot{q}_y)\dot{q}_y + G_{y0}(q_y) + D_{y0}(\dot{q}_y) + \bar{U}_y = \begin{bmatrix} 0 & \tau_y \end{bmatrix}^T \quad (10)$$

where $\bar{U}_y = \Delta M_y \ddot{q}_y + \Delta C_y \dot{q}_y + \Delta G_y(q_y) + \Delta D_y = \begin{bmatrix} \bar{U}_1 & \bar{U}_2 \end{bmatrix}^T$

With the four state variables $X_Y = \begin{bmatrix} x_{y1} & x_{y2} & x_{y3} & x_{y4} \end{bmatrix}^T$, that is $x_{y1} = \theta_y$, $x_{y2} = \phi_y$, $x_{y3} = \dot{\theta}_y$ and $x_{y4} = \dot{\phi}_y$ one obtains the subsequent nonlinear state equations

$$\dot{X}_Y = \begin{bmatrix} \begin{bmatrix} x_{y3} \\ x_{y4} \end{bmatrix} \\ M_{y0}(x_Y)^{-1}\left(\begin{bmatrix} 0 \\ u_x \end{bmatrix} - \alpha - \beta - \begin{bmatrix} \bar{U}_1 \\ \bar{U}_2 \end{bmatrix} \right) \end{bmatrix}$$

$$\alpha = C_{y0}(x_Y, \dot{x}_Y)\dot{x}_Y + G_{y0}(x_Y) \quad (11)$$

$$\beta = \begin{bmatrix} 0 \\ \mu_{\phi y0} x_{y4} + \mu_{\phi y0}\,\mathrm{sgn}(x_{y4}) \end{bmatrix}$$

Similarly, equation (11) represents the SIMO dynamic equation of the robot in the median coronal plane

III. CONTROLLER DESIGN

1. Introduction to control system

The simulation model used in Figure 3 was designed that stated in the previous section is based on the inverse ball drive mobile robot using various parameters.

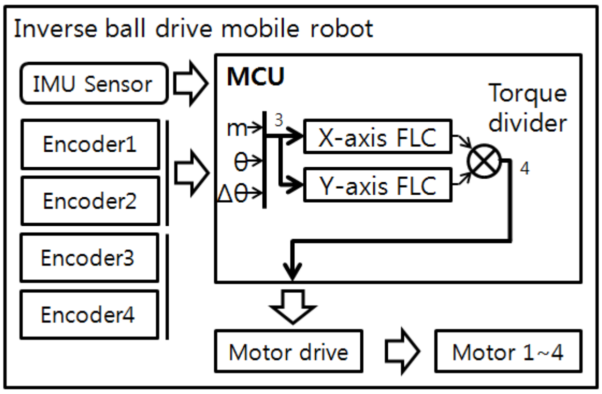


Fig.3 System Configuration

Position and velocity of the main body attached to the motor encoder can know its value from. It can find angle and angular rate for balancing from IMU sensor. Figure 4's controller was designed for balancing angle and angular velocity from the inverse ball drive mobile robot. Controller Input m is the weight of the main body, pi is angle of the main body, delta pi value is the angular velocity of the main body, and delta u value is motor drive input voltage value.

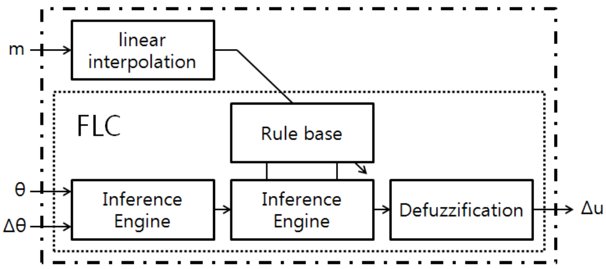


Fig.4. Block diagram of the controller

2. Fuzzy controller

Through a process of normalization, input of the Fuzzy controller comes output. When output is normalization, it is represented by linguistic rules of "IF-THEN form". Final output value comes from defuzzification. In this process, L, S and Z in Fuzzy controllers input membership function means Positive Z-axis in the direction of the robot when the size of balancing. N, Z and P means Positive X-axis in the direction of the robot when look into the front. In addition, L, M, S and Z in output membership function mean the strength of the output. N, P and Z mean Positive X-axis in the direction of the wheel. Table 1 is represented to those membership functions. The graph in Figure 5 shows fuzzy rule table [4].

Table 1. Fuzzy rule table

y \ x	LN	SN	Z	SP	LP
N	LP	MP	Z	SN	MN
Z	LP	SP	Z	SN	LN
P	MP	SP	Z	MN	LN

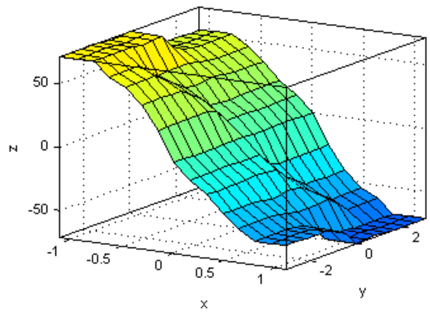


Fig.5. Surface graph of the fuzzy rule table

Mandani's MAX-MIN method was used in reasoning process and COA(center of area) method was used to derive output for defuzzification [5].

IV. SIMULATION

Simulations were run, in order to verify the validity of the proposed scheme. First of all, Tune the 3 specified weights using a trial-and-error method. These values is the specimen value of 3 fuzzy membership functions. Controller that 3 specimen values can determine fuzzy membership element value and using fuzzy membership functions for the controller element values tuned by the linear interpolation method is modeled by Simulink program.

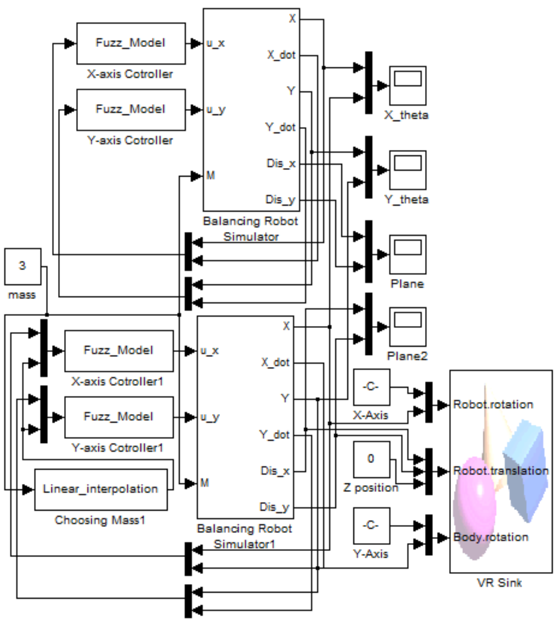


Fig.6. Schematic Design with Simulink

We applying it to the main body and analyze the inverse ball drive mobile robot's position and angle. Figure 6 is the inverse ball drive mobile robot designed by Simulink program and Figure 7 is the inverse ball drive mobile robot designed by CATIA for 3D visual simulation. It can be measured performance of controller depending on the weight. Finally, we get the value of angle and angular rate from inverse ball drive mobile robot. It was designed for transmit in real-time controller. According to this robot' s weight, we have applied with controller that 3 specimen values can determine fuzzy membership element value and using fuzzy membership functions for the controller element values tuned by the linear interpolation method. We just decide random value 4kg, 7kg and 8.1kg that can know the difference in performance instead of specimen value 3kg, 6kg, 9kg for the simulation graph.

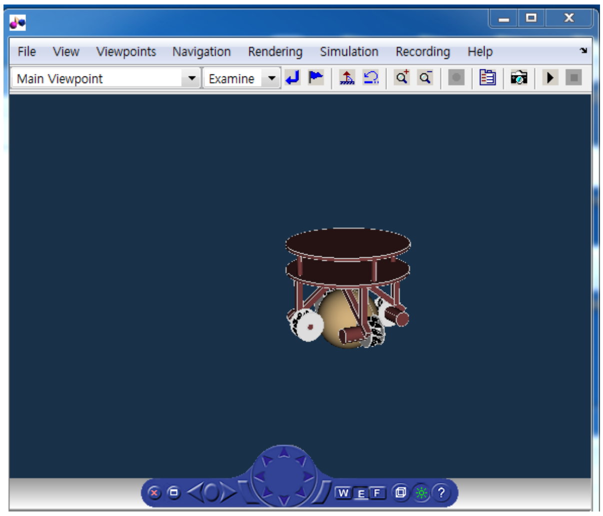


Fig.7. 3D modeling for visual simulation

The controller is the same of the proposed at the x-axis and y-axis. So the performance is compared with the existing controller and proposed controller at the x-axis.

First, Figure 8 shows angle trajectory of 4kg's weight. Simulation results of figure 8 shows well-balancing controller compared with conventional controller and proposed controller.

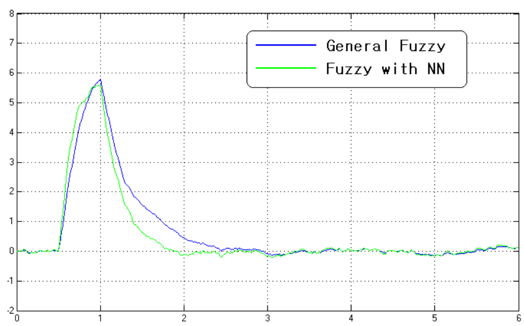


Fig.8. X-axis's angular trajectory for the weight of 4kg

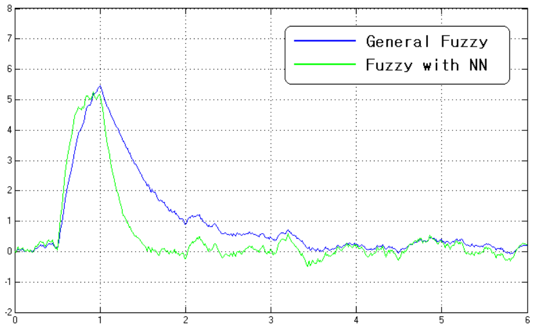


Fig.9. X-axis's angular trajectory for the weight of 7.1kg

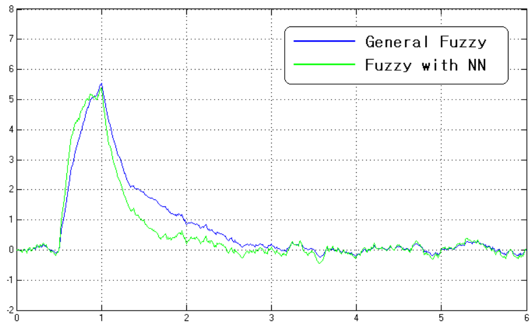


Fig.10. X-axis's angular trajectory for the weight of 8.1kg

Figure 9, 10 shows angle trajectory of 7.1kg's weight and 8.1kg's weight. Increased inertia weight makes control difficulty and it makes the more large difference in performance evident than in the case of the front. so the proposed controller than conventional controller is found to show better performance. Through those results, using fuzzy membership functions for the controller element values tuned by the linear interpolation method is better than controller that 3 specimen value can determine fuzzy membership element value.

V. CONCLUSION

In this paper, we figured out 3 values depending on weight through trial and error. This is general fuzzy controller's fuzzy membership function element for better performance. Then, by the linear interpolation method help the optimal value and apply fuzzy controller. Using simulation, we compared fuzzy controller that has 3 standard fuzzy membership fuction factor value and proposed controller with fuzzy membership fuction factor value resulted from neural network. In this simulation, the proposed controller shown good performance when 4kg, 7.1kg and 8.1kg random weight value was input. Finally, We can rise stability about general the inverse ball drive mobile robot balancing

ACKNOWLEDGEMENT

"This research was financially supported by the Ministry Of Trade, Industry & Energy(MOTIE), Korea Institute for Advancement of Technology(KIAT) and DongNam Institute For Regional Program Evaluation(IRPE) through the Leading Industry Development for Economic Region."

REFERENCES

[1] Hyunwook Kim, Seul Jung, "Experimental Studies of Controller Design for a Car-like Balancing Robot with a Variable Mass," KIIS, vol. 20, no. 4, pp.469-475 2010.

[2] U. Nagarajan and A. Mampetta, G. A. Kantor and R. L. Hollis, "State Transition, Balancing, Station Keeping, and Yaw Control for a Dynamically Stable Single Spherical Wheel Mobile Robot," *in Proc.IEEE Int. Conf. Robot. and Autom*., pp. 998-1003, 2009.

[3] U. Nagarajan, G. A. Kantor and R. L. Hollis, "Trajectory Planning and Control of an Underactuated Dynamically Stable Single Spherical Wheeled Mobile Robot," *in Proc. IEEE Int. Conf. Robot. and Autom*., pp. 3743-3748, 2009.

[4] L. A. Zadeh, "Fuzzy sets," Information and control, pp. 338-353, 1965.

[5] C. T. Lin and C. S. G. Lee, "Neural Fuzzy System," Prentice Hall, pp. 142-159, 1996.

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Error Correction of Angular Velocity for Gyroscope using Genetic Algorithm and FIR filter

Jaeyong Kim*, Hyunhak Cho**, Sungshin Kim***+

**Department of Electrical and Computer Engineering, Pusan National University, Korea*
*** Department of Interdisciplinary Cooperative Course: Robot, Pusan National University, Korea*
****Department of Electrical Engineering, Pusan National University, Korea*
Tel : 82-051-510-2367; Fax : 82-97-597-7760
arioner@pusan.ac.kr, darkruby1004@pusan.ac.kr, sskim@pusan.ac.kr

Abstract: This paper is research which sensitivity of gyro sensor is optimization. Recently, the MEMS-gyroscope is usually used at mobile robots and AGV by development of MEMS technology. However, the MEMS-gyroscope has measurement error, cumulate error, bias drift, etc. To reduce these errors, measurement of angular velocity using MEMES-gyroscope required high precision and exactly sensitivity. The sensitivity, is a constant to change angular velocity form ADC value, is changed by external factors as installed location and gradient because the MEMS-gyroscope measure angular velocity with inertia. Therefore, this paper proposes to reduce variance of ADC value and to optimize the sensitivity of MEMS-gyroscope using genetic algorithm. To verify performance of proposed method, we attached MEMS-gyroscope to automatic guided vehicle (AGV) of fork type. In experimental result, we verified that the sensitivity using the proposed method is accurate than the sensitivity in specification.

Keywords: MEMS, Gyroscope, Genetic Algorithm, Sensitivity, Optimization, Error Correction

I. INTRODUCTION

Recently, interest about the performance of sensors is increased by increasing interest about automation. And then the research to improve performance of sensors is very active. Particularly, low cost and small size MEMS-type sensors are used in various areas by the development of MEMS technology. MEMS-type sensors are used in robot area. The angular velocity of robot is measured by MEMS-type gyroscope[1,2].

The autonomous guided vehicle (AGV) has been important by the growth of logistics industry. The AGV is a kind of mobile robot. It is device which automatically transport object on given location. The core technology used in AGV is induction system. In induction system, there are wire guide and wireless guide. And there is magnetic-gyro which mixed property of wire and wireless guide. The magnetic-gyro is method which sparsely laid cylinder-type magnet each 5m~10m under the ground and then drive magnetic sensor and gyroscope sensor[3-5].

In this method, the AGV drive using variation of angle from gyroscope during straight driving. And if magnetic sensor measures magnet, the position and angle of AGV is corrected by information of magnet and perform previously inputted work. This guided method decrease install charges over 15~30 times for existed wire guided method, and is easy addition or maintenance of driving course. And if AGV should precise drive, the user has only to install magnet under the ground.

The magnetic-gyro use MEMS-type gyroscope which is one of the mechanical gyroscope. It is better small, low power and low cost than electrohydraulic, optical fiber, ring laser, etc.

However MEMS-type gyroscope is lower accuracy of angular velocity and performance of straight than typically gyroscope. Therefore, the magnetic-gyro guided method is required as high accuracy ADC (analog to digital converter) and accurate sensitivity.

The error variance of ADC is larger by noise which occurs by neighborhood environment as magnetic force, electromagnetic force. And the sensitivity of MEMS-type gyroscope is changed by external factors which installed location and slope because it measures the angular velocity using inertial.

Therefore, this paper proposes method that the error variance of ADC is reduced by FIR-filter with raw data. And the sensitivity of MEMS-type gyroscope is optimized by genetic algorithm.

II. OPTIMIZATION OF SENSITIVITY

To experiment, we installed MEMS-type gyroscope in forklift AGV with laser navigation. Follow picture is forklift AGV for experiment.

Fig.1. forklift AGV for testing

Half-EPS for passive control is in the forklift AGV. However it is converted into position controlled full-EPS for experiment. Fig. 2 is system configration of forklift AGV.

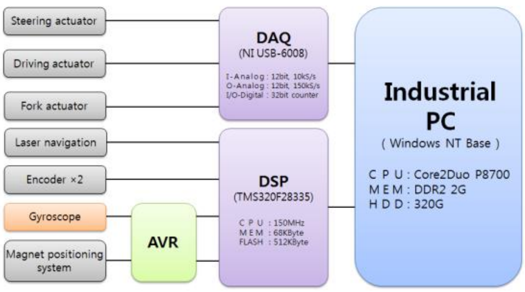


Fig.2. System configration

where industrial PC is used for rapid research and development. DAQ is used to speed control of motor. The laser navigation for global localization of AGV is installed on the top of AGV. The data of the sensors are transferred from AVR to DSP. DSP send gathered data through RS-232 communication to industrial PC. Table 1 is specification of installed sensors. Table 1 is specification of sensors.

Table 1. Specification of sensors

Sensor	Specification	
Laser Navigation (NAV200)	Supply voltage	24V
	Positioning accuracy	±4m～±25 mm
	Angular accuracy	±0.1˚
MEMS-type Gyro (ADXRS613)	Input voltage	5 V
	Range	±150 ˚ /s
	Drift	±3 %

III. OPTIMIZATION OF SENSITIVITY

1. FIR Filter

The FIR filter is finite the convergence of impulse response. FIR filter has stability because it need return loop due to non-recursive filter. And it usually used the transmission of waveform since shape of waveform of input and output is maintained. This paper performs the low-pass filter using FIR-filter from changed raw data from ADC.

FIR-filter reflects aspects which change slow on input signal because it has average output signal. However it does not reflects rapidly input signal. This is smoothing of high frequency by average operator. And it is low pass filter since low frequency is influenced by FIR-filter. The equation of FIR-filter is following.

$$y[n] = b_0 x[n] + b_1 x[n-1] + \cdots + b_N (n-N) \qquad (1)$$

$$y[n] = \frac{1}{4}x[n] + \frac{1}{4}x[n-1] + \frac{1}{4}x[n-2] + \frac{1}{4}x[n-3] \qquad (2)$$

This paper used 4 order equations. *x[n]* is input, *y[n]* is output, b_i is coefficient of filter, *N* is degree of filter.

Optimization of sensitivity using GA

The sensitivity of MEMS-type gyroscope is changed by external factors which installed location and slope. Therefore the sensitivity of MEMS-type gyroscope can be optimized at place installed gyro using measured data after installed gyro in system.

This paper proposes to optimize sensitivity of MEMS-type gyroscope using genetic algorithm with cumulated filtered data.

The genetic algorithm used crossover, reproduce operator and generation number repeated 100 times. The crossover selects 2 chromosomes which are high fitness. The reproduce pass 50% chromosome on next generation. Equation to optimize the sensitivity of MEMS-type gyro is following.

$$f(x) = \sum_{i=0}^{N} (360 - x \times O^{+}) / N \qquad (3)$$

where x is the sensitivity, O+ is sum of ADC when AGV turn 360˚ . Fig. 3. is flowchart of algorithm in this paper.

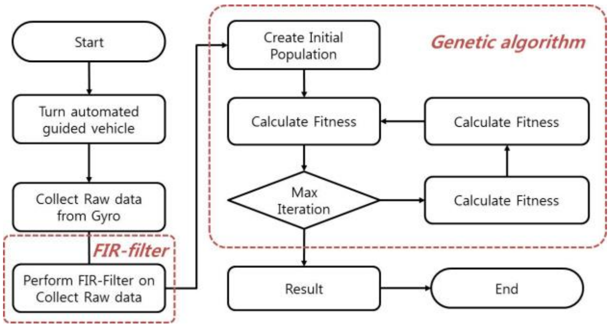


Fig.3. Flowchart of proposed algorithm

IV. EXPERIMENT AND RESULT

To verify performance of sensitivity optimization proposed in this paper, we installed MEMS-type gyroscope and laser navigation in forklift AGV. The laser navigation is used to measure 360° turn of AGV.

The experiment performed in 840×2010cm of space with 19 of reflector as Fig. 4. The AGV turn 360° while steering angle is fixed.

Fig.4. Environment of experiment

The output of MEMS-type gyroscope is angular velocity and temperature for temperature compensation with analog. To measure raw data of MEMS-type gyroscope, we changed analog data to digital data using 12bit-ADC. And output of raw data is changed as following.

$$O^{+} = O^{-} - ((2^{12}/2) - T) \times CA \qquad (4)$$

where O^{-} and T is output of gyro and temperature, respectively. In this paper, the factor of temperature is set 0.08 through experiment. Eq. (5) is angular velocity using the compensated output of MEMS-type gyroscope.

$$\Delta\theta = (C - O^{+}) \times S \qquad (5)$$

The experiment is performed at 5times according to driving speed when steering angle of AGV is fixed each 10° among 30~60° of left and right, respectively. Table 2 is average error on experiment at 5 times.

Table 2. Result of experiment

Steering angle	Driving Speed	Error(unit: degree)		
		FIR+GA: 0.0221	GA: 0.0251	Spec: 0.0312
30°	48cm/s	1.176	1.702	44.476
	43 cm/s	0.338	2.534	43.955
	38 cm/s	1.740	1.142	44.826
	33 cm/s	1.654	1.228	44.773
	28 cm/s	1.889	0.995	44.919
	23 cm/s	1.310	1.569	44.559
	average	1.351	1.528	44.585
40°	48 cm/s	3.182	6.099	47.536
	43 cm/s	4.440	7.367	48.324
	38 cm/s	1.578	1.301	44.553
	33 cm/s	1.402	1.479	44.664
	28 cm/s	0.306	2.583	45.351
	23 cm/s	1.250	1.632	44.759
	average	2.026	3.410	45.864
50°	48 cm/s	1.836	1.040	44.392
	43 cm/s	0.942	1.942	44.952
	38 cm/s	1.022	1.861	44.902
	33 cm/s	0.351	2.538	45.322
	28 cm/s	2.082	0.793	44.238
	23 cm/s	1.106	1.777	44.849
	average	1.834	1.658	44.770
60°	48 cm/s	7.628	10.581	50.321
	43 cm/s	6.908	9.856	49.870
	38 cm/s	5.932	8.871	49.258
	33 cm/s	1.224	4.125	46.309
	28 cm/s	4.376	7.303	48.284
	23 cm/s	2.668	5.582	47.214
	average	4.789	7.719	48.543
Total average		2.500	3.579	45.940

As table 2, The optimized sensitivity using proposed method was 0.0221/° /25ms, and it using GA was 0.0251V/° /25ms. The results of average error on 360° rotation are 2.500°, 3.579°, 45.940° in proposed method, GA, specification respectively.

proposed method is lower error on turn of 360° than difference method. Because of proposed method considers information on environment, and get data of low variance through FIR-filter.

V. CONCLUSION

This paper proposed optimization of sensitivity, reducing variance of output from MEMS-type gyroscope. The noise of outputs of MEMS-type gyroscope should be reduced since the outputs of MEMS-type gyroscope are influenced by external factors. And the sensitivity should be optimized in place where MEMS-type gyroscope is installed because MEMS-type gyro is inertial sensor.

Therefore this paper reduced variance of MEMS-type gyroscope using FIR-filter, and optimized sensitivity using genetic algorithm.

To the experiment, we used forklift AGV. To the result of experiment, we verified that proposed method is lower error.

ACKNOWLEDGMENTS

"This research was supported by the MOTIE(The Ministry of Trade, Industry and Energy), Korea, under the Human Resources Development Program for Special Environment Navigation/Localization National Robotics Research Center support program supervised by the NIPA(National IT Industry Promotion Agency)." (H1502-13-1001)

REFERENCES

[1] Groover, MP (2007). Automation, production systems, and computer-integrated manufacturing. Prentice Hall Press

[2] Pang, Y., Lopez De La Cruz, A., & Lodewijks, G. (2008, June). Bipolar magnetic positioning system for automated guided vehicles. In Intelligent Vehicles Symposium, 2008 IEEE (pp. 883-888). IEEE

[3] Borenstein, J. (2000). The OmniMate: a guidewire- and beacon-free AGV for highly reconfigurable applications. International Journal of Production Research, 38(9), 1993-2010

[4] Chan, C. Y. (2002). A system review of magnetic sensing system for ground vehicle control and guidance

[5] Dash, P. K., Hasan, S., & Panigrahi, B. K. (2010). Adaptive complex unscented Kalman filter for frequency estimation of time-varying signals. IET science, measurement & technology, 4(2), 93-103

Control of Robot Arm with Sterilization System for Ballast Water

Seung-Hwa Baek[1], Dong-Hyun Kim[2] and Hee-Je Kim[3]

[1]*Department of Electrical Engineering, Pusan National University, Jangjeon-dong, Geumjung-gu, Korea,* bsh6845@naver.com

[2]*Department of Interdisciplinary Program in Robotics, Pusan National University, Jangjeon-dong, Geumjung-gu, Korea,* kikkc@naver.com

[3]*Department of Electrical Engineering, Pusan National University, Jangjeon-dong, Geumjung-gu, Korea,* heeje@pusan.ac.kr

Abstract: The inadvertent transfer of harmful aquatic organisms and pathogens in the ballast water of ships has been determined to have caused a significant adverse impact to many of the world's coastal regions. Loading and discharge of ballast water is an essential part of a ship's operation, and is fundamental to maintaining safe operations under different conditions of load. However, large vessels require thousands of tonnes of water to ensure stability and manoeuvrability, and the environmental impacts of this can be considerable. These impacts result from the fact that the ballast water can contain hundreds of different species, many of which can have serious ecological, economic and public health effects if transferred to regions where they are not native. The recognition of these effects has made ballast management increasingly important for protection of the marine environment. As a means to prevent, we developed a system using high voltage- high frequency pulsed power, so called arc discharge, when this discharge happens, the electrodes are oxidized and consumed. Regarding this point, we wanted to apply robot arm to supply continuous electrode when it consumed by arc discharge.

Keywords: Ballast water treatment**,** Robot arm control, High voltage-high frequency, Pulse power.

I. INTRODUCTION

In general, a cargo ship in the sea transportation of freight interchange similar to the round-trip sailing ship , except for so many to be one-way flight one way flights to sail to the full load condition , then return voyage the ship's balance, stability and coordination to improve performance and bottom and left and right of the ship ballast tanks (ballast tank) and install a separate tank , here ballast (ballast water) (freshwater or saltwater) flows into the board and will sail to the ballast condition .

Ballast water on board ship without entering the service , when the ship's center of gravity is at the top of the back and forth violently shake from side to side and rollover risk , can not be locked by a professional Palo propulsive efficiency falls dramatically into problems .

The amount of water flowing into the ballast tanks of the ship size , ie, the displacement is set by reference to the case of large vessels over 10,000 tons of seawater flowing cases frequently appear. In this process, water inflow , which includes a fine marine vessel docked national and local seawater inflow and go to other countries and regions to release the water , the water entrained in the process of entering other fine marine life the problem has spread to areas that are .

The movement of these marine organisms are not rare phenomena in the natural world represented by the variation of species , such as red tide occurrence that happens because of the ecological disturbance causes a very serious problem .

As a result, the IMO (International Maritime Organization: IMO) in the 2004 Ballast Water Management Convention, adopted ship , and discharge of ballast water and water quality standards established in 2011 decided to apply the Convention in order .

In particular, the exotic marine life in the marine environment due to its heavy losses sustained and claims the United States, Australia, Canada, and on several occasions over the active request and the relevant international conferences International Maritime Organization (IMO) in the February 2004 ' ship ballast water and sediment management International Convention was adopted .

In response, various methods of ballast water treatment systems for the development and commercialization efforts for preemption is going on around the world , and in this way, the filtering method, ultraviolet disinfection , high-temperature heating method such as physics, chemistry, electrical processing method and apparatus is being developed.

Ballast tank to remove the fine marine organisms to filter water entering the inlet reduces the amount of water , so the processing capacity of the issues and the filter mesh (mesh) of the complete process is limited because the problem does not occur is . The chemicals put into the ballast water introduced to the marine life while the ballast tank to remove toxic chemicals remaining in the water near the water at the time of the release of the environmental damage caused by pollution is a problem .

In order to solve the above problems, before being stored in the ballast tanks of the ship, or the ship's ballast tanks, ballast water directly from the ocean ballast water before it is discharged from the discharge electrode arc generated by the ballast water sterilization equipment.

II. EXPERIMENTALS

1. Preparation

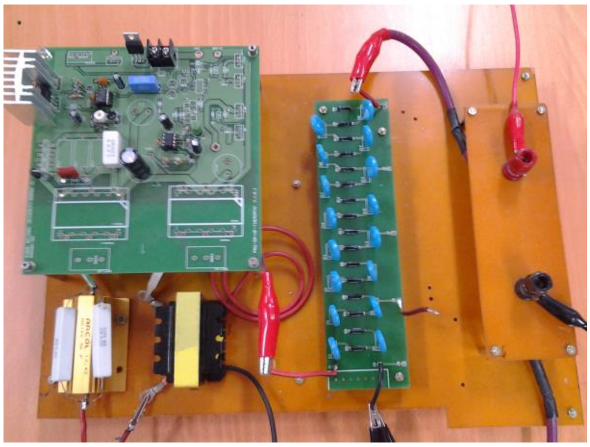
The degree of rust through an electrochemical reaction of electrolyte was checked to compare the effects of DC and high frequency, high voltage input energy. To maintain constant electrolytes, 5 liters of water and 6 grams of salt were mixed, and 200mL was used in every test. To match the amount of energy, both the DC RMS and high voltage pulsed power, two outputs were fixed to 20V RMS in the water. The experimental time for each test was 10 minutes. The surface of the copper wire and needle were observed by scanning electron microscopy (SEM). 5 liters of water and 6 grams of salt are mixed in a test bottle and 200ml of the resulting electrolyte was used for the corrosion test on copper wire and needle.

2. Circuit and system

The high-frequency, high-voltage pulse system output circuit was made to vary the voltage and frequency. A MC34063 for buck topology, IR2153 for dimming and a 2:200 transformer with Fly-back topology were applied.

The output voltage was varied up to 5kV pk-pk and the output was adjusted to make a RMS value of 20V.

For the same applied voltage, the RMS value was adjusted to 20V in water by applying a high frequency and high voltage pulse of 1kV. Copper wire and needle were weighed before and after the corrosion test, and the surface of the needle was observed by SEM at ×500 magnification to compare the effects of DC and high voltage pulsed power.



3. Electrochemical Reaction

Most of the corrosion proceeds via an electrochemical reaction involving an anode and cathode. The oxidation reaction is can be expressed as

$$M \rightarrow M^{+} + (-)$$

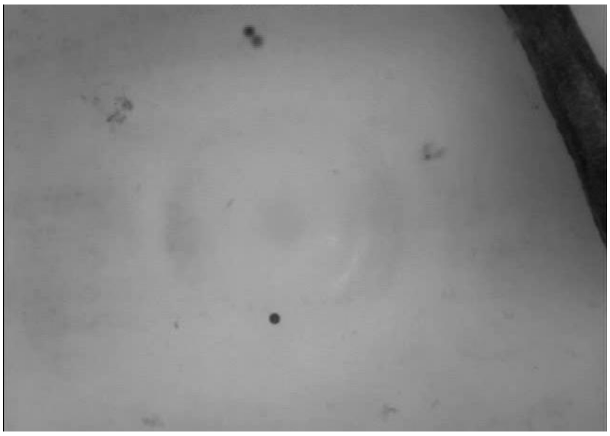
The reduction reaction can be written as follows:

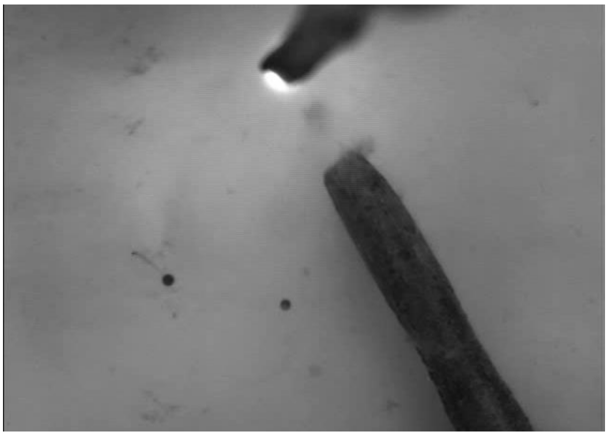
$$M^{+} + (-) \rightarrow M_2^0$$

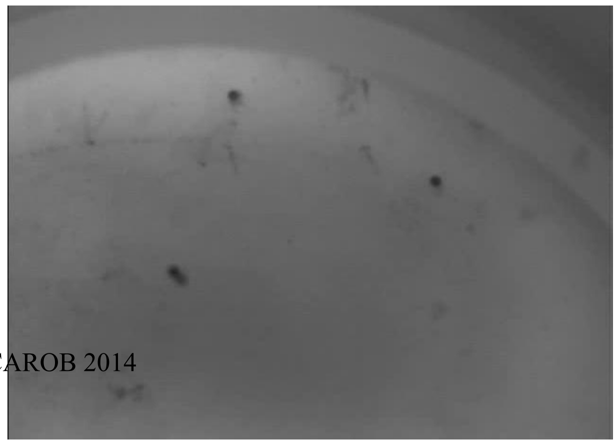
In this study, salt was used as electrolyte to boost corrosion, and the electrical corrosion reaction in salt water can be expressed as

$$Cu + 4OH \rightarrow Cu_2O + 2e(-) + 2H_2O$$

4.Results







III. CONCLUSION

The installation of ballast water management systems (or ballast water treatment systems), designed, reviewed, approved, installed and operated to satisfy an agreed upon ballast water discharge performance standard has been determined by the international marine industry to provide a more effective means to prevent, minimize and ultimately eliminate the transfer of organisms and pathogens via vessel ballast discharge, when compared to ballast water exchange.

Hart and the conjugate base of hydrogen peroxide as Hoigne HO2- That can break down the ozone that can act as initiator were found , much faster than hydroxyl and ozone decomposition OH radicals can be generated over the found. Artificially by adding hydrogen peroxide to ozone at a faster rate by the decomposition of ozone OH La How to create a radical PEROXONE AOP is called. As a result of the initial ozonide radical reactions (O3-) And HO2 - This is generated by a chain reaction, each will generate OH radicals.

$$2O3 + H2O \rightarrow 2OH + 3O2$$

Notably, hydrogen peroxide , as well as a catalyst to generate OH radicals with OH radicals updated. Raephal to as the scavenger is that action . Thus, hydrogen peroxide injection than required. Rather may cause an adverse effect on organic removal. Ground water treatment, surface water taste and odor causing material removal, etc. , but rather than using ozone to ozone. By using hydrogen peroxide and ozone effective dose to reduce the processing efficiency can be obtained. In particular, the non-biodegradable organic substances, when applied to paper wastewater containing, as well as COD, TOC removal is that you can get a fairly good test results and costs

REFERENCES

[1] Ilie Suarasan et al., "Experimental characterization of multi-point corona discharge devices for direct ozonization of liquids", J. of Electrostatics, 54 207-214 (2002)
[2] U. Kogelschatz et al., "Dielectric-Barrier Discharges. Principle and Applications", J. Phys. IV France, 7, C4-47 (1997)
[3] H.M. Jones and E.E. Kunhardt, "Development of pulsed dielectric breakdown in liquids", J. Phys. D: Appl. Phys, 28, 178 (1995)
[4] A. Abou-Ghazala, et al., "Bacterial decontamination of water by means of pulsed corona discharges", Pulsed Power Plasma Science, IEEE Conference, Las vegas (2001)
[5] S. Espie et al., "Investigation Of Dissolved Ozone Production Using Plasma Discharges In Liquid", Pulsed Power Plasma Science, IEEE Conference, Las vegas (2001)
[6] I.V. Lisitsyn et al., "Water treatment by pulsed streamer discharges", Pulsed Power Conference, IEEE Conference (1999)
[7] M. Mikula et al., "The destruction effect of a pulse discharge in water suspensions", Plasma Sources Sci. Tech., 6 179 (1997)

Performance Study for Vehicle Infotainment System

Heung-in Park, Min-gyu Kim, Jeong Hee-In, Kang-il Park, In-uk Lee, Dong-Ju Lee and Jangmyung Lee

Department of Electrics Engineering, Pusan National University, Busan, 616-111, Korea
Tel : 82-51-510-1696; Tel : 82-51-510-2378
heungin@gmail.com

Abstract: CPU performance, booting time and navigation widget are improved by the new design concepts. These points are providing some solutions that can overcome the risks of open architecture system in the Automotive. The improvement was verified with each of the test results.

Keywords: Open architecture, System performance, Booting time

I. INTRODUCTION

Vehicle infotainment system is getting a lot of function and become more complex behavior with the rapid development of electronic devices inside the vehicle. Just a few years ago, system functions that is limited to information offer services using the navigation and telematics are evolving as combining innovative technologies that have consumer IT sector devices, such as smart phones Foreign competitors, such as Ford Sync and GM On-Star, also improve the vehicle merchantability through the company's IT brands. Therefore, merchantability and innovation characteristics of IT equipment inside the vehicle becoming increasingly important for vehicle OEM's brand appreciation and vehicle marketing strategy. But, as conjunction with many external devices and home functions plant to vehicle environment, reliability and how to properly design reliability and performance condition become an important issue. For example, even if consumer's smart phone takes over 1 minute boot time, the situation is not so serious for a user. But, the case that the boot of the car multimedia system takes many time feel uncomfortable to users because of limit user functions for essential, such as navigation system, voice recognition, rearview camera, telematics, Bluetooth phone functions.

In this paper, under the vehicle reliability environmental conditions, the optimal design concept which can stably support complex Use-Case of vehicle infotainment system and the design content which can secure system performance are proposed.

And, the performance improvement method about the main optimization tuning items such as mode switching speed, boot optimization, operation stability which needs improvement under the limited system resources is proposed.

II. System Design

1. Function and feature requirements

Development system is integrated infotainment system that integrates audio, video, navigation and telematics functions of inside the vehicle. From safety/security functions to a simple convenience functions, various complex functions are realized on a single integrated system using a single CPU.

System architecture such as fig1 is designed in order that multimedia technology and a number of functions work smoothly without disconnection

2. SW Architecture

In order to ensure vehicles reliability and performance, As shown in Fig.2, add vehicle-specific Automotive Framework based on Basic Architecture Framework, classified as Automotive API list and Open API for development apps downloading. Automotive framework section is divided into service framework parts and Native section. Service framework parts performs role to control the main function and service of the same system, such as fast-boot service.

In native area, connectivity and security has been supported such as all of the infrastructure of the system.

III. Performance improvement

As mentioned above, at the system configuration and the SW architecture-based environment of the vehicle, satisfied reliability and performance should be premised. At the following section of this paper, a proposed design concept for ensuring reliability and performance in the environment of these vehicles is presented.

1. Complex Use-Case design

First, Foreground and Background motion of system to define complex motion Use-Case is restricted. It is provide that motion service of foreground and background have performance limitation is 40 percent and another 20 percent is defined as the Marginal. This basis calculated by the foregoing process is standard for the highest performance of CPU required when the voice recognition behavior exists. Thus, if applying the above-mentioned concept to application, maximum permissible share of visible or invisible on the screen of the application became 40%, so this concept can apply to other applications. For instance, a navigation to guide path showing or not showing map must satisfy such requirement.

If the full screen share is defined, it needs to restrict share about a unit application. A unit application should not exceed a maximum of 40% is defined.

After that, when each function is activated, before and after motion are provided by motion characteristics of function as Permanent, Temporary, Background mode of the three kinds of operation characteristics classified.

Then, specifications defined the service and functional priority regulations. When the trivial behavior is acted such as listening the music, watching the navigation in the car, the functional priority regulation is not necessary. This regulation is required in an emergency like as telematics automatic incident reporting. For example, when the car accident occurred and airbag was inflated, the accident story had to be reported preferentially to the emergency rescue center, and system resources should not be allocated for the low priority function processing such as entertainment function, telephone conversation.

2. Complex behavior management

We adopt process scheduling methods which use the order of priority provided by operating system, and high priority function can be conducted overbearingly in complex behavior.

The scheduling way provided OS distributes CPU resources economically depending on the number of process acted at the same time as Fig. 3.

3. Predictive control for load balancing system

In order to ensure the reliability and performance of the vehicle infotainment system, the functions of multiple provided by the system must be designed to properly scheduling priorities for each function.

Also, the management criteria functional priority for environment of the vehicle is defined as the priority of the five like Table1.

A highest priority function is that responsible for communication in the vehicle and the power of the system. Notification function is managed by the next higher one on AV, MEDIA, and the guidance voice of navigation in order that the audio output of the AV media playing can’t be uninterrupted.

The combined operation function of normal system for most situations would like to work at Group5 and Group4. However, when the resources of all of the CPU are concentrated for start-up at times system booting, the operation state of the application can be on the screen occupied state in Background state by calling the user interaction.

At this time, The CPU executes the call. If the method of static priority is used, the execution of the called application is made that does not meet the performance requirements. To deal with this problem, the priority is defined in advance so that the predictive load balancing control is implemented according to the dynamic operating conditions of application

For example, if blue-tooth connection configuration behavior is carried to execution in the blue-tooth phone application described in Fig.4

In the case that the priority of application is controlled dynamically, the switched application, which occupies the screen, has a good UI operation performance.

It compensates drawback of traditional method that the foreground operation or switching screen run slowly.

The operating system scheduler implements the highest priority first after checking on the thread of high priority function in a brief time so that its operating performance is guaranteed even at CPU 100% sharing state as Fig 5.

IV. FIGURES/TABLES

1. Figures

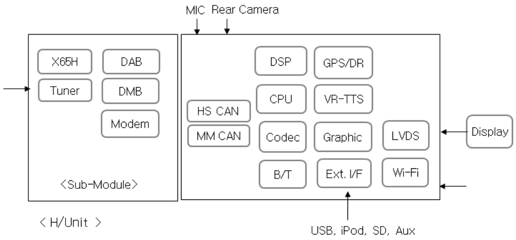


Fig.1. System Configuration

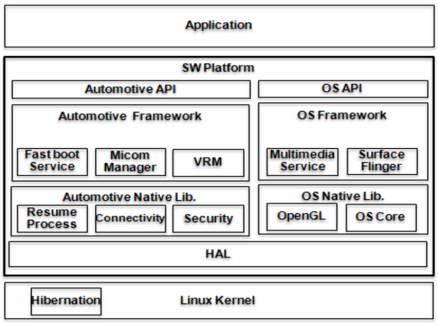


Fig.2. SW Architecture

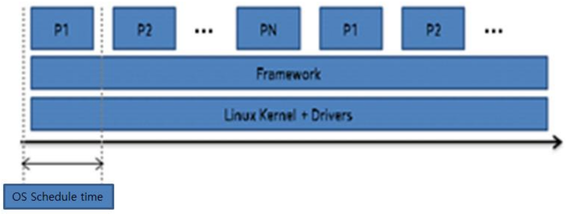


Fig.3. Operating system (GPOS) scheduling methods

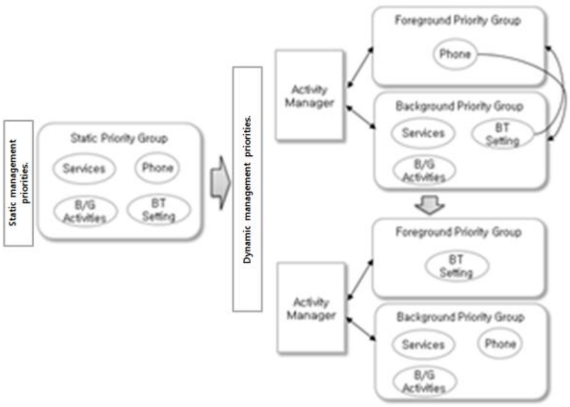


Fig.4. Dynamic management priorities.

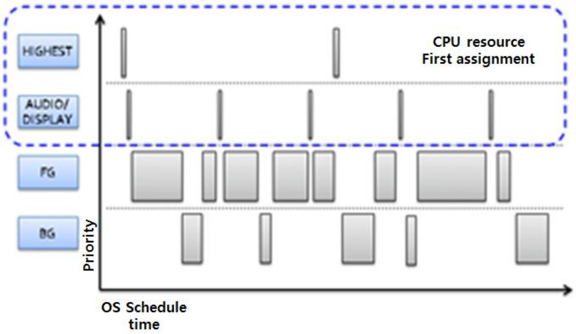


Fig.5. Secured important performance from priority management.

2. Tables

Table1. Definition of scheduling priority

Priorities	Description
Group 1	Power Management, Communications in vehicle, First priority.
Group 2	AV, Media, Navigation guidance sound, High priority for loss of sound output,
Group3	Surface Flinger : System graphic processing
Group 4	System Server, Telematics, Phonebook, Navigation, Foreground application.
Group 5	Other Background application

Table 2. The results of CPU utilization from BT phonebook download.

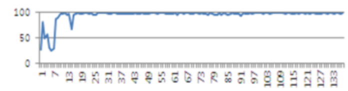
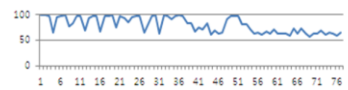
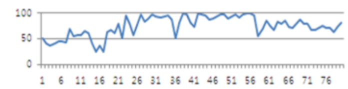
	Result	Max	Min	Ave
Initial version		99	26	95
First Improved version		99	57	80
Optimized version		99	24	75

Table 3. The result of switching performance.

	Initial Version	Optimized version
Navigation switching	3.5 sec	0.6 sec
AV/Media mode switching	3~5 sec	0.4~0.6 sec
Rear camera switching	4 sec	2 sec

VI. CONCLUSION

Both optimizing design techniques and performance improvement that presented in this paper was confirmed finally through the evaluation of the actual performance and following results were identified.

(1) Table.2 shows major improvement results for complex behavior present in this paper. This is a complex behavior in case of downloading 1000 of phonebooks by being connected to cell phone and Bluetooth hands-free while navigation is operating as the background. There were a maintaining CPU 100% and a excessive waste of time during downloading phonebooks because the initial version has a structural problem about the data processing. The primary improvement was to improve the above structural problems. And the optimization version was to upgrade through management of priority of process scheduling described above or dispersion of load giving sleep at regular intervals in the SW logic. Through these improvements solved the problem of ANR (Application No Response) that frequently occurred because other complex behaviors are very slow while downloading phonebook.

(2) The switching performance between different functions such as Navigation, AV / Media, rear camera and so on has a limit with only structural improvement like each application initialization logic and navigation.

Achieved performance improvement results of Table.3 by defining about priority of process scheduling and managing priority dynamically base on complex Use-Case design described above.

ACKNOWLEDGEMENT

“This research was financially supported by the Ministry Of Trade, Industry & Energy(MOTIE), Korea Institute for Advancement of Technology(KIAT) and DongNam Institute For Regional Program Evaluation(IRPE) through the Leading Industry Development for Economic Region.”

REFERENCES

[1] M. Rahmani, “Novel Network Architecture for In-Vehicle Audio and Video Communication Continuum Limit of Lattice Gas Fluid Dynamics”, IEEE, 2007
[2] Jose Santa, “ An Embedded Service Platform for the Vehicle Domain”, IEEE, 2007
[3] Hyeong-ju Song, Tae-yeon Kim, Ji-hoon Park, Bak Lee, Ki-young Lim, “Inside Android”, Wiki-books, 2010
[4] Fabian Hueger, “Platform Independent Applications for In-Vehicle Infotainment Systems via Integration of CE Devices”, IEEE, 2012

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Proposal of a Visualizing Method of Data Transitions to Support Debugging for Java Programs

[1]Hiroto Nakamura, [1]Tetsuro Katayama, [2]Yoshihiro Kita, [1]Hisaaki Yamaba and [1]Naonobu Okazaki

[1]University of Miyazaki, 1-1 Gakuen-kibanadai nishi, Miyazaki, 889-2192 Japan
[2]Kanagawa Institute of Technology, 1030 Shimo-ogino, Kanagawa, 243-0292 Japan
tf13006@student.miyazaki-u.ac.jp, kat@cs.miyazaki-u.ac.jp, kita@earth.cs.miyazaki-u.ac.jp, yamaba@cs.miyazaki-u.ac.jp, oka@cs.miyazaki-u.ac.jp

***Abstract*:** It takes much time to find the cause of a bug in debugging of programs. Finding the cause of a bug needs to comprehend a flow and data transitions in executing programs. It is difficult to grasp behavior in executing the programs whose behavior is unexpected by a bug. We propose a visualizing method of data transitions to support debugging for Java programs in order to improve efficiency of debugging by supporting to find the cause of a bug. We have implemented TVIS in order to confirm efficiency of the proposed method. The data transitions diagram is the most characteristic function of TVIS which shows the data transitions in executing programs as a table. It can show visually abnormal behavior: no data renewed at all, data abnormally renewed, and so on. Because abnormal behavior is detected in the data transitions diagram at first glance, it is useful for programmers in finding the cause of a bug. This paper shows that the method can support to find the cause of a bug.
***Keywords*:** programming, program slicing, visualization, debug, dynamic analysis, syntax analysis

I. INTRODUCTION

It takes much time to find the cause of a bug in debugging of programs [1]. Finding the cause of a bug needs to comprehend a flow and data transitions in executing programs. It is difficult to grasp behavior in executing the programs whose behavior is unexpected by a bug. Additionally, many programs are complicated for including a number of loops and branches. In particular, unskilled programmers need more time for this work.

Thin slicing [2] is a technique to find the cause of a bug, but sometimes its information is insufficient. Thin slicing is a kind of program slicing [3], and can analyze data transitions without superfluous information by restricting abstraction of slice to data transitions only. However, enough information to find the cause of a bug cannot be gotten because states and renewals timing of data which is not selected as slice are not shown.

This paper proposes a visualizing method of data transitions to support debugging for Java programs in order to improve efficiency of debugging by supporting to find the cause of a bug. The method supports to grasp behavior in executing programs because visualizes renewals and states of each data.

This paper has implemented TVIS which is tool to visualize data transitions for Java programs to confirm efficiency of the proposed method. Main functions of TVIS are the data transitions diagram, the renewal history table, and the slicing function. In particular, the data transitions diagram is the most characteristic function which shows the data transitions in executing programs as a table. Therefore, the method shows abnormal behavior and data renewals for understanding behavior of executing programs.

This paper visualizes the programs which include a bug in order to confirm that TVIS can support to find the cause of a bug.

II. DATA TRANSITIONS

The data transitions in this paper show the flow of variable renewals in executing a program. It expresses when and what value of each variable is renewed. Programmers can grasp behavior of a program because they can prefigure the behavior of each variable at arbitrary timing in the program execution by understanding the data transitions. Hence, finding the cause of a bug becomes easier by comparing the difference between the behavior which programmers expect and the actual behavior, if they grasp the data transitions of the program which includes a bug.

However, it is difficult to grasp data transitions of the program which includes a bug. The reason is that the program which includes a bug doesn't behave as programmer's expectation, in addition they doesn’t know the cause of a bug.

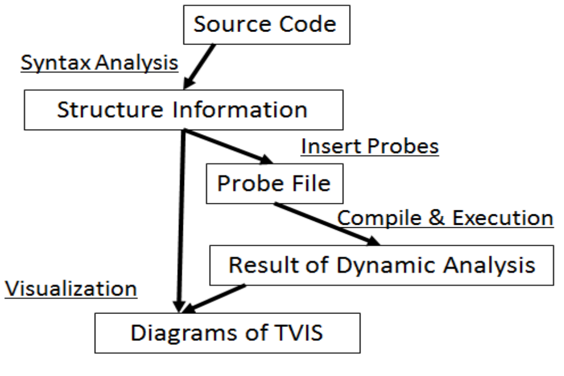


Fig.1. The process of visualizing data transitions

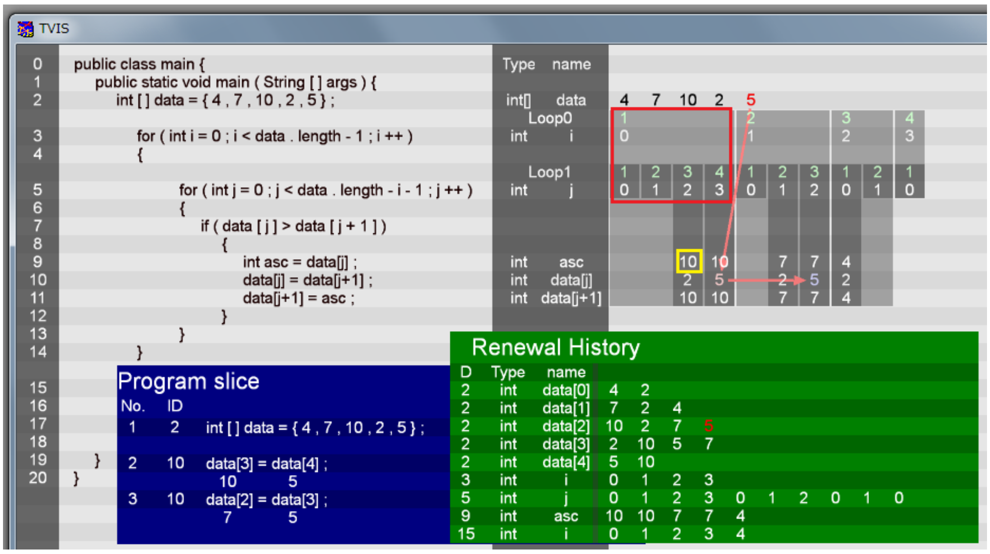


Fig.2. An example of the window of TVIS

III. VISUALIZATION

In order to confirm efficiency of the proposed method, TVIS (transitions visualization), which is tool to support debugging, has been implemented. It can visualize data transitions. Fig.1 shows the process of visualizing data transitions. The contents of the process is as follows.

1. TVIS statically analyses structure information of the program in order to choose the point which probes to do dynamic analysis are inserted in.
   TVIS does syntax analysis of the source code which is visualized and analyses the result of it for investigating structure information of the program. Structure information of a program is information which includes the following data: locations of declaring variables, locations of renewals, ranges of each variable scope, locations of loops, and so on.
2. TVIS generates the source code which includes probes, and then outputs the result of dynamic analysis of the program by executing it.
   TVIS generates the probe file which is inserted probes in the source code to get information of renewals of variables and loops from structure information. TVIS outputs the result of dynamic analysis by executing the probe file after compiling it.
3. TVIS visualizes data transitions by using the result of dynamic analysis and the structure information of a program.
   The result of dynamic analysis has information in each process in executing the program: declaration and renewals of variables, and loop. TVIS analyses data transitions which occur in each process by using the result of dynamic analysis, and shows them on the window.

Main functions of TVIS are the data transitions diagram, the renewal history table, and the slicing function. Fig.2 shows an example of the window of TVIS. The data transitions diagram is on the upper right of the window, and the renewal history table is on the green area, and slices list as the result of the slicing function is on the blue area. TVIS can support to grasp behavior of executing programs because it shows data transitions by using these functions.

The data transitions diagram is the most characteristic function which shows renewals of each variable and its timing. It is a table and indicates the number of iterations of each loop in a lateral direction and indicates the line number of data renewals in a vertical direction, and shows renewal value of variables in each loop. It is possible to understand the line and the loop where each renewal exists by using the data transitions diagram. For example, the area which is surrounded by the red frame in Fig.2 shows that *Loop1* is repeated four times in the first loop of *Loop0*. The area which is surrounded by the yellow frame in Fig.2 shows that the value of variable *asc* is assigned into 10 in the third loop of *Loop1* in the first loop of *Loop0*. Hence, the data transitions diagram can show the flow of renewals of each data. Moreover, it is possible to grasp renewals and flow of the whole by using the table form, it is easy to understand states of the other data when a suspicious state occurs.

In addition, TVIS shows the data transitions arrow which shows relations between each renewal of variables as the red arrow on the data transitions diagram. The relations mean ones between state of a variable and the state of other variables which is used for the renewal of its state. After a programmer selects a state by clicking a value on the data transitions diagram, TVIS draws a red arrow from the state which influenced its renewal to the state which was selected. When a

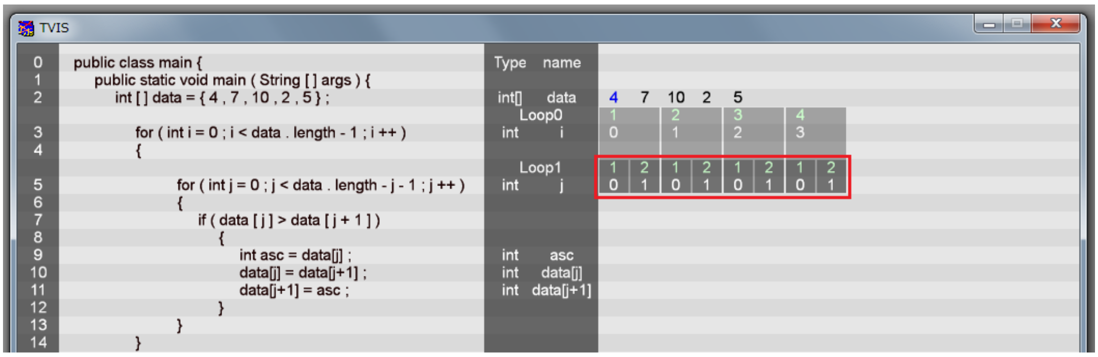


Fig.3. An example A: Visualizing the program which includes a bug

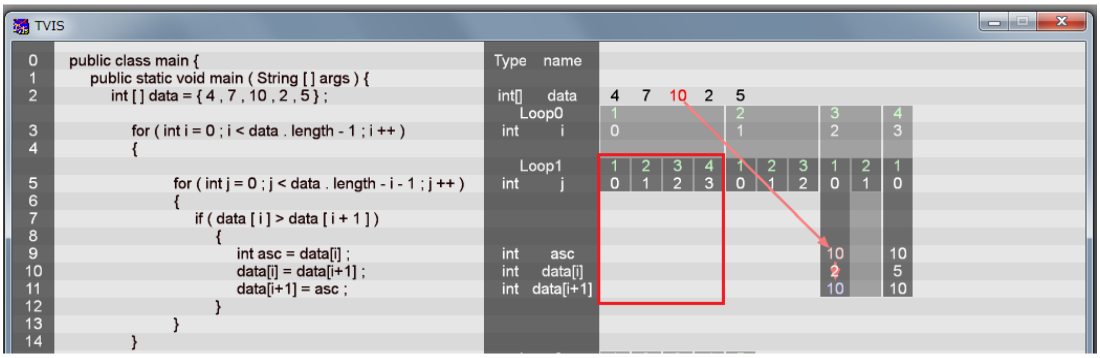


Fig.4. An example B: Visualizing the program which includes another bug

suspicious state occurs, finding the cause of it becomes easy by using the data transitions arrow.

The renewal history table shows how many renewals of each variable in executing programs happen and what its value is after renewals. It indicates the number of renewals in a lateral direction and indicates a variable name in a vertical direction. It can show information about whether abnormalities of values of variables after renewals and the number of the renewals happen.

In addition, TVIS can perform thin slicing to arbitrary states of each variable as an ancillary function. When the value on the data transitions diagram or the renewal history table is clicked, TVIS performs thin slicing, and shows the result of slicing on the blue area of the window. TVIS can perform thin slicing at one click, and it becomes easy to perform the slicing to not only the final state but also an arbitrary state of each variable by using the data transitions diagram or the renewal history table. Therefore, analysis of data transitions can be performed without superfluous slices.

IV. EXAMPLE

Two programs which include a bug are visualized by adapting them to TVIS, to confirm efficiency of the proposed method. They are bubble sort programs in Java and include different bugs.

The bubble sort program in Fig.3 sets a wrong condition of a loop. The correct loop condition in the 5th line in Fig.3 is not "*j < data.length-j-1*" but "*j < data.length-i-1*". This mistake that programmers confuse variables of similar names happens often, but they hardly notice existence of it. The program has returned the array which is identical with the original array because of this mistake.

The data transitions diagram on Fig.3 shows abnormal behavior of exchanging elements of the array from 9th line to 11th line and *Loop1* which starts at the 5th line. The area which is surrounded by the red frame shows that all iterations in *Loop1* repeated only two times and variable *j as the loop counter* didn't become more than two. The correct behavior of *Loop1* dwindles the number of iterations of the loop as shown Fig.2. Hence, we understand that the loop condition of *Loop1* is suspicious.

The bubble sort program in Fig.4 contains variables assigned to a wrong value. The variable which is used from 7th line to 11th line on Fig.4 is not variable *i* but variable *j*. This mistake occurred by merely confusing variables, but it is hardly noticed because multiple lines are wrong in succession. The program finishes process without executing until completion of sorting because of this mistake.

The data transitions diagram on Fig.4 shows abnormal behavior to exchange elements of the array. The area which is surrounded by the red frame shows that Loop1 repeated four times in the first loop of *Loop0* which starts at the 3rd line. Here, *if* statement at the 7th line, was expected to compare all elements of the array, but it never exchanged them. Hence, we understand that *if* statement at the 7th line is suspicious.

From the above two results of visualization, we have confirmed that programmers can get useful information to find the cause of a bug and grasp behavior of the program included the bug by using visualization of data transitions of TVIS.

V. EVALUATION

This paper has implemented TVIS which supports debugging for visualizing data transitions, in order to confirm efficiency of the proposed method by visualizing the program which includes bugs. It explains the difference by comparing the method with conventional methods below.

Dynamic slicing [4] which includes thin slicing is a technique to analyze behavior in executing a program by extracting processes which relate to generation of a selected state as slice. We cannot get useful information if not we appositely decide a condition of slicing, because slicing cannot show information which is not selected as slice. In case of Fig.3, it is difficult to get useful information, because slice except initialization processes is not gotten even if elements of the array are selected as a condition of slicing. However, TVIS can show useful information for programmers to grasp behavior of a program, because it shows the situation of each variable and loop by the diagram.

Breakpoint is one of the most frequently used methods for debugging [5]. Breakpoint has the problem which programmers need experience of programming to decide a location of breakpoint. Nevertheless, it is a useful method of finding the cause of a bug, and also the method which automatically generates breakpoint has reported [6]. When programmers found an abnormal process, it is difficult to understand correctly the situation of execution with only information of a point. TVIS becomes easy to grasp behavior of a program, when programmers found the abnormal process, because TVIS shows them to information of variables or loops which concern it. In a large program, efficiency of TVIS may lower in comparison with breakpoint, because information of the diagram is liable to increase.

An applicable range of programs which TVIS can visualize is small in comparison with the tool [7] which can visualize a large program by a diagram with a high abstraction level or tool [8] which can visualize multithreaded programs. TVIS, which generates a diagram of low abstraction level, may make gigantic a diagram because data to visualize even a program of moderate size may become huge [9]. Therefore, the variety of programs which TVIS can visualize may have a limit. Accordingly, TVIS need reduce spaces to show information by improving the expression format of the diagram. An applicable range of programs which TVIS can visualize is enlarged by focusing object to visualize on only an important object.

VI. CONCLUSION

This paper has proposed the visualizing method of data transitions to support debugging for Java programs in order to improve efficiency of debugging by supporting to find the cause of a bug. We have implemented TVIS which supports debugging and visualizes data transitions, in order to show efficiency of the proposed method. We visualize programs which include a bug by using TVIS, and show that TVIS can support to grasp behavior of executing programs and to find the cause of a bug.

Visualization of the proposed method can give the information which is not gotten by using conventional methods for analyzing data transitions. The data transitions diagram can show visually abnormal behavior: abnormal data renewals and so on. Also, it is easy to understanding states of other variables when abnormal behavior is found.

Therefore, the proposed method can support to find the cause of a bug by visualizing data transitions, and improve efficiency of debugging for Java programs. The method reduces time to grasp behavior in executing a program for programmers by showing data transitions to them. In particular, it is useful for unskilled programmers.

The future issues are as follows.

- Improving the format for expressing values of variables.
- Introducing the localization in visualization.

REFERENCES

[1] Roger S. Pressman (2001), Software Engineering A Practitioner’s Approach 5thEdition, McGraw-Hill Science
[2] M. Sridharan, S. J. Fink, R. Bodik (2007), Thin slicing, In Proc. the 2007 ACM SIGPLAN Conference on PLDI, pp.112-122
[3] Mark Weiser (1982), Programmers Use Slices When Debugging, Communications of the ACM Vol.25, pp. 446-452
[4] H Agrawal, JR Horgan (1990), Dynamic Program Slicing, SIGPLAN Notices, Vol.25, No.6, pp.246-256.
[5] G. C. Murphy et al (2006), How Are Java Software Developers Using the Eclipse IDE?, IEEE Software, Vol.23, No.4, pp.76-83.
[6] Cheng Zhang et al (2013), Automated Breakpoint Generation for Debugging, JOURNAL OF SOFTWARE, Vol.8, No.3, pp.603-616
[7] Steven P. Reiss, Guy Eddon (2005), From the Concrete to the Abstract: Visual Representations of Program Execution, DMS 2005, pp.315-320.
[8] Jan Lönnberg et al (2011), Java replay for dependence-based debugging, PADTAD '11, pp.15-25
[9] W. De Pauw et al (1998), Execution patterns in object-oriented visualization, In Proc. 4th COOTS, pp. 219-234

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Proposal of a Method to Build Markov Chain Usage Model from UML Diagrams for Communication Delay Testing in Distributed Systems

Zhijia Zhao[1], Tetsuro Katayama[1], Yoshihiro Kita[2], Hisaaki Yamaba[1] and Naonobu Okazaki[1]

[1]University of Miyazaki, 1-1 Gakuen-kibanadai nishi, Miyazaki, 889-2192 Japan
[2]Kanagawa Institute of Technology, 1030 Shimo-ogino, Kanagawa, 243-0292 Japan
zhao@earth.cs.miyazaki-u.ac.jp, kat@cs.miyazaki-u.ac.jp, kita@earth.cs.miyazaki-u.ac.jp, yamaba@cs.miyazaki-u.ac.jp, oka@cs.miyazaki-u.ac.jp

Abstract: As the growth of network technology with high parallelism and high reliability of distributed systems, they have been widely adopted in the enterprise and society. But, how to design for testing the real-time or a communication delay of distributed systems is not been discussed much. This paper proposes a new method to automatically build Markov Chain Usage Model. The proposed method establishes the time points and builds a new UML diagram which can test the communication delay of a distributed system to improve its reliability. We have confirmed the usefulness of the proposed method to adapt it to practical examples.

Keywords: distributed systems, communication delay, Markov Chain Usage Model, UML diagram, reliability

I. INTRODUCTION

Software testing is one of the key technologies to ensure software reliability. For decades, many software testing methods have been proposed. Building the software usage model becomes the basis of reliability evaluation and reliability testing. Software testing with Markov Chain Usage Model is an effective method to be used by programmers and testers in web sites development, to guarantee the software reliability.

Because UML (Unified Modeling Language) [1] has become a practical standard modeling language for object-oriented program, many software developers adopt UML diagrams for object-oriented development. How to build the usage model from UML diagrams has been proposed [2]. But, how to build a usage model from UML diagrams corresponding to the communication delay of distributed systems has not been discussed much.

In this paper, we aim to improve the reliability of distributed systems by testing the communication delay. To achieve this aim, we propose a new method to automatically build Markov Chain Usage Model by establishing the time points and building a new UML diagram. Specifically, we establish the time points and build a new UML diagram called De-sequence diagram based on deployment diagram and sequence diagram. Then, we propose an algorithm to generate the Markov Chain Usage Model from the new UML diagram.

II. METHOD OF BUILDING MARKOV CHAIN USAGE MODEL

1. De-sequence diagram

For testing the communication delay of a distributed system from UML diagram, we need to describe interactive time and sequence between objects of system that distributed on each node. Sequence diagram shows the dynamic partnership between objects, but it does not shows the layout of the distributed system, and it also does not describe the time of messages sent. For this reason, we combine it with deployment diagram, and set the time point to describe interactive time and sequence between objects. We call this diagram as De-sequence diagram.

As shown in Fig.1, De-sequence diagram puts the system distributed on each node as an object, and puts the life-line between time points as a state. It can describe interactive time and sequence between objects while describing dependencies of system distributed on each node.

Specifically, we show the definitions as follows:

Definition1. Let system that we will test as: Sys = ($Inst_1 \dots Inst_n$), Sys is the system, and Inst is the instance.

Definition2. Let the instance of the system as: Inst = (Na, Node, TrPre, TrPost, Trans, Pf), Na is the name of instance, Node is the node that instance belongs to, TrPre is the Pre-condition of Transition, TrPost is the Post-condition of Transition, Trans is the Transition of the system, and Pf is the probability of the Transition.

Definition3. Let message of the system as: Mes = (Mna, Sour, Targ), Mna is the name of the message, Sour is the source instance of the message, and Targ is the target instance of the message.

Before the other definitions, we propose a method as follows to divide the system into some cases for establishing the time points easily in Definition5:

Method1. Find a message from the beginning. If the source instance of the message is the same as the target instance of the next message of the first one, and the target instance of first message is the same as the source instance of the second message, we call the message and the next message pair messages. Otherwise, we call the message single message. We construct a case with two instances and pair messages or single message. Construct the other cases as above.

Definition4. Let the case we constructed in method1 as C = (C_i, Mes, Inst, CPre, CPost), C_i is the number of the case, CPre is the Pre-condition of the case, CPost is the Post-condition of the case, we use CPre and CPost to describe the sequence of cases, Mes and Inst has been defined in definition2 and definition3.

Definition5. If a message let the case have a transition, we call this message CPre, the Pre-condition of the case. If a message be sent just after the transition of the case, we call this message CPost, the Pre-condition of the case.

Definition6. Let Pre-condition of Transition as: TrPre = (Tm, Tv, La), Tm is the time mark, Tv is time violation, and La is the label of system.

Definition7. Let time mark as: Tm = (Mna, C_it_j), Mna is the name of message, and C_it_j is the count of time points.

Definition8. Let time violation as: Tv = (Tc, TvH, TvPos, Tvpro), Tc is the time constraint equation, TvH is the time violation processing, TvPos is the Post-condition of time violation, and TvPro is the probability of time violation.

Definition9. Let system label as: La = (Initial, Middle, TiVio, Final), Initial is the initial stage of system, Middle is the intermediate stage of system, TiVio is the time violation stage of system, and Final is the final stage of system.

Definition10. Let Post-condition of Transition as: TrPost = (Tm’).

Definition11. Let Trans as Transitions, it is a process to satisfy the TrPost on you are satisfied with the TrPre.

2. Definition of Markov Chain Usage Model and generation method

Re-define the Markov Chain Usage Model to generate Markov Chain Usage Model from De-sequence diagram for testing distributed systems.

Definition12. Let Markov Chain Usage Model as: MC = (S, L, Trans, S_0, F), S is the state, L is the state transition label, Trans is the transition, S_0 is the initial state, and F is the final state.

Definition13. Let state as: $s \in S$ = (Sna, La), Sna is the state name, and La is the state mark.

Definition14. Let state transition label as: $l \in L$ = (C_it_j, Tc, Mna, Pf), C_it_j is the time point, Tc is the time constraint equation, Mna is the massage name, and Pf is the probability of state transition.

Definition15. Let state transition as: $S \times L \rightarrow s$, Trans(s, l) = v, it means state s use l(state transition label) to get to state v.

After that we propose a method to generate Markov Chain Usage Model from De-sequence diagram.

Method2. Choose one of cases we constructed to perform the following operations.

1. Set initial state and final state.
2. Set Pre-condition of transition to be the event of state transition, Post-condition of transition to be arriver state, and enter the state name.
3. Construct the state transition label like l = <C_it_j, Tc, Mna, Pf>.
4. The transition of time violation accordance with the provisions of Tv, and the probability of time violation is been set at 0.1 this time.

Finish other cases like this. At last, we propose a method to combine each Markov Chain Usage Model of cases we constructed.

Method3. First, use the CPre and CPost (defined in Definition5) to describe sequence relationship of cases. If the CPost of case A is the same as the CPre of case B, that case A is the foregoing case of case B, and case B is the immediate successor case of case A. Then, combine them together to build Markov Chain Model for whole system. Take Markov Chain Usage Model initial state of the case to be the final state of its foregoing case, and take the final state of the case to be the initial state of its immediate successor case. If the case does not have the foregoing case or immediate successor case, maintain the status quo.

III. APPLICATION EXAMPLES

We use E-commerce system and Factory automation system as application examples to confirm the usefulness of the proposed method. We generate the test cases with roulette test case automatic generation method [3] based on two kinds of Markov Chain Usage Model.

1. E-commerce System

E-commerce system [4] have three parts distributed in three nodes, it is SS(Supplier Server) in node1, DS(Data Server) in node2, and DC(Data Control) in node3. SS sent commodity query to DS, DS received the message and let DC to query commodity information, DC queried the commodity information and sent it to DS, DS received the information and sent it to SS. From node2 to node3 have three second network delay.

We generate the test cases for E-commerce by building Markov Chain Usage Model from De-sequence diagram as follows:

1. Construct the De-sequence diagram for the E-commerce system shown in Fig.1 use the Definition1 to Definition11.
2. Divide the system into three cases shown in Fig.2 use the Method1.
3. Choose the first case to generate its Markov Chain Usage Model shown in Fig.3 use the Definition12 to Definition15 and Method2.
4. Generate Markov Chain Usage Model for other cases like the first case.
5. Combine them together to build Markov Chain Usage Model for whole system shown in Fig.4 use the Method3.
6. Generate test cases with roulette method (one time to run it for time violation this time).

Then, we can have two test cases:{cc, cq, tv, tvh, cq, qm, cm} and {cc, cq, qm, cm}. {tv, tvh, cq} in test cases means to test the communication delay. The Markov Chain Usage Model built by the conventional method has only one test case:{cc, cq, qm, cm}.

2. Factory Automation System

Factory automation system [4] have five parts distributed in three nodes, SS(Supplier Server) is in node1, PS(Parts Supply) is in node2, WC(Workstation Controller), WC1(Workstation Controller1) and PC(Package Controller) are in node3. Have an order request, C will sent parts query to PS, PS feedback information to C, C is waiting until parts are enough. Then C ask WC whether can work or not and waiting until WC can work. After that, C sent a request to PS for giving parts to WC. When WC finished its work, it ask WC1 whether can work or not and waiting until WC1 can work, then WC sent commodity to WC1. When WC1 finished its work, it ask PC whether can work or not and waiting until PC can work, then sent commodity to PC. From node1 to node2 have two second network delay, from node2 to node3 have three second network delay, and from node1 to node3 have four second network delay.

We use the same method with E-commerce system. Construct De-sequence diagram for the Factory

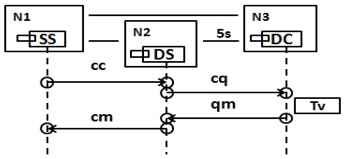
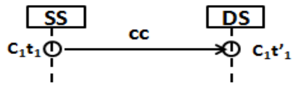
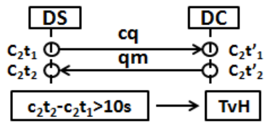


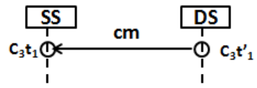
Fig.1. De-sequence diagram of E-commerce system



(case1)



(case2)



(case3)

Fig.2. Cases of E-commerce system

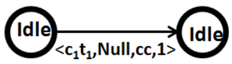


Fig.3. Markov Chain Usage Model of case1

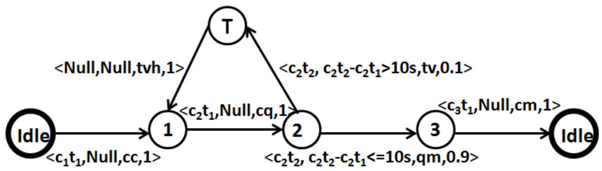


Fig.4. Markov Chain Usage Model of E-commerce system

automation system shown in Fig.5, divide the system into eight cases, generate Markov Chain Usage Model for each case, and combine them together to build Markov Chain Usage Model for whole system shown in Fig.6.

Then, we can have four test cases:{cqp, tv, tvh, cqp, qmp, cqwc, tv, tvh, cqwc, qmwc, sr, sd, cqwc1, qmwc1, sdwc, cqpc, qmpc, sdwc1}, {cqp, tv, tvh, cqp, qmp, cqwc, qmwc, sr, sd, cqwc1, qmwc1, sdwc, cqpc, qmpc, sdwc1}, {cqp, qmp, cqwc, tv, tvh, cqwc, qmwc, sr, sd, cqwc1, qmwc1, sdwc, cqpc, qmpc, sdwc1}, and {cqp, qmp, cqwc, qmwc, sr, sd, cqwc1, qmwc1, sdwc, cqpc, qmpc, sdwc1}. The Markov Chain Usage Model built by the conventional method has only one test case:{cqp, qmp, cqwc, qmwc, sr, sd, cqwc1, qmwc1, sdwc, cqpc, qmpc, sdwc1}.

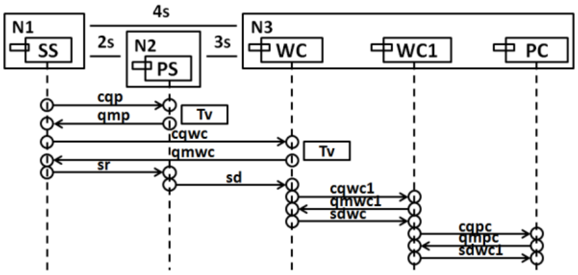


Fig.5. De-sequence diagram of Factory automation system

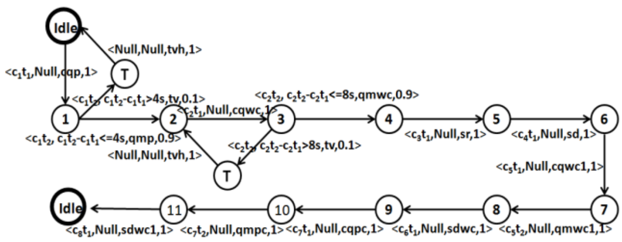


Fig.6. Markov Chain Usage Model of Factory automation system

IV. DISCUSSTION

Reference [3] is also proposed a method to build Markov Chain Usage Model from UML diagram, but the method does not discuss the condition of distributed systems. In this paper, we have established the time points to describe the interactive time and sequence between objects and built a new UML diagram to describe the dependencies of distributed systems. We have got the test case included {tv, tvh, cq} in the example of the E-commerce system, and {tv, tvh, cqp}, {tv, tvh, cqwc} in the example of the Factory automation system that the method of reference [2] cannot have. Hence, from comparing results we can know, our proposed method can find the fault of communication delay that the conventional method cannot find. Consequently, we have confirmed that the proposed method can improve the reliability of distributed systems.

V. CONCLUSION

In this paper, we aim to improve the reliability of distributed systems by testing the communication delay. To achieve this aim, we have proposed a new method to automatically build Markov Chain Usage Model by establishing the time points and building a new UML diagram for testing communication delay of a distributed system. By adapting this method to practical application examples, we have confirmed that we can test the communication delay of distributed systems that the conventional method cannot test, and that the proposed method can improve the reliability of distributed systems.

Future issues are as follows:

- Get the probability of time violation.
  We did not discuss the probability of time violation, just set it at 0.1 in this paper. We need to propose a method to get the probability of time violation.
- Get the Coverage of time violation
  We did not discuss the coverage of time violation, just run it one time in this paper. We need to think the coverage of time violation to let the test case accurate even more.

REFERENCES

[1] Grady booch, James Rumbaugh, and Ivar Jacobson, The Unified Modeling Language User Guide Second Edition, Addison-Wesley, 2005.
[2] Wu Caihua, Liu Juntao, Peng Shirui, Li Haihong, Deriving Markov Chain Usage Model from UML Model (in Chinese). Journal of Computer Research and Development, 49(8): 1811-1819, 2012.
[3] Lei Hang, Chen Limin: Test Case Generation Based on Markov Chain Usage Model (in Chinese). Journal of University of Electronic Science and Technology of China, 40(5): 732-736, 2011.
[4] Hassan Gomaa, Designing Concurrent, Distributed, and Real-Time Applications with UML, Addison-Wesley, pp: 567-622, 2000.

Proposal of a Supporting Method for Debugging to Reproduce Java Multi-threaded Programs by Petri-Net

Shoichiro Kitano[1], Tetsuro Katayama[1], Yoshihiro Kita[2], Hisaaki Yamaba[1] and Naonobu Okazaki[1]

[1]*University of Miyazaki, 1-1 Gakuen-kibanadai nishi, Miyazaki, 889-2192 Japan*

[2]*Kanagawa Institute of Technology, 1030 Shimo-ogino, Kanagawa, 243-0292 Japan*

kitano@earth.cs.miyazaki-u.ac.jp, katayama@cs.miyazaki-u.ac.jp, kita@earth.cs.miyazaki-u.ac.jp, yamaba@cs.miyazaki-u.ac.jp, oka@cs.miyazaki-u.ac.jp

Abstract: It is difficult to implement the multi-threaded programs. The reason is that the behavior of each thread is non-deterministic. Also it is difficult to reproduce the situation in which an incident occurs. This paper proposes a supporting method for debugging to reproduce Java multi-threaded programs by visualizing the behavior of the programs with Petri-net. Moreover, we have confirmed the effectiveness of our method by implementing a tool.

Keywords: multi-threaded program, debugging, Petri-net, Java, reproducibility

I. INTRODUCTION

In recent year, many computers are adopted multi-core CPUs. For using such resources effectively, the demand of multi-threaded programs increases.

In multi-threaded programs, it is hard work for even expert programmers to implement them, and easier to embed bugs than single-threaded programs[1]. Most of such bugs are discovered in latter half of the development process or in executing the programs by users. Therefore, it is difficult to fix the programs. It is necessary to remove the bugs at the unit testing to resolve this problem.

One of the testing methods executes a program with plural interleaving by putting off a timing of execution of each thread. Hereby, we can discover potential bugs in a multi-threaded program.

However, even if this testing method shows that bugs exist, in multi-threaded programs, it is hard to discover the cause of the bugs in debugging because the behavior of multi-threaded programs is non-deterministic.

This paper proposes a supporting method for debugging to reproduce multi-threaded programs by Petri-net to improve efficiency of debugging work for the multi-threaded programs written in Java language.

Specifically, the proposed method gives reproducibility to multi-threaded programs, generates data file for execution path of multi-threaded program, and simulates the behavior of the program by Petri-net based on the data file.

Here, ordinal Petri-net cannot express the behavior of multi-threaded programs written in Java completely. Therefore, we extend Petri-net so that it can treat with the behavior of Java multi-threaded programs.

In addition, we confirm validity of our method by implementing a supporting tool that tests multi-threaded programs with plural interleaving automatically and simulates their execution path with Petri-net.

II. SUPPORTING METHOD USING PETRI-NET

1. Supporting method

This section proposes a supporting method for debugging that improves the efficiency of discovering the cause of bugs by giving reproducibility to multi-thread programs. The procedure is following.

1. Generating the data file of the execution path in the situation in which an incident occurs
   The data has ID of threads executing process, executed processes, timing of processes, and ID of generated threads.
2. Generating a Petri-net model from a tested Java program
   We explain the method to model Java programs by Petri-net in the next section.
3. Simulating the behavior of Java program which has tested
   This process uses the data file and a Petri-net model.

2. Modeling Petri-net for Java Programs

This paper models Java programs by Petri-net using the conversion rule as follows. Fig.2 is an example of a

Petri-net model for the program of Fig.1 which is an example of a multi-threaded program written in Java.

- a statement is converted into a place
- a reserved word 'synchronized' is converted into a place
- a state of waiting lock is converted into a place
- a beginning of the method is converted into a place
- a transition to the next statement is converted into a transition
- a thread is converted into a token
- a locked instance is converted into a token

```
public class Example {

    static class ThreadExample extends Thread {
        public void run() {
            Object lock = new Object();
            synchronized (lock) {
                System.out.println("in synch");
            }
            System.out.println("lock release");
        }
    }

    public static void main(String[] args) {
        System.out.println("Hello");
        new ThreadExample().start();
        System.out.println("World");
    }
}
```

Fig.1. Example of a multi-threaded program written in Java

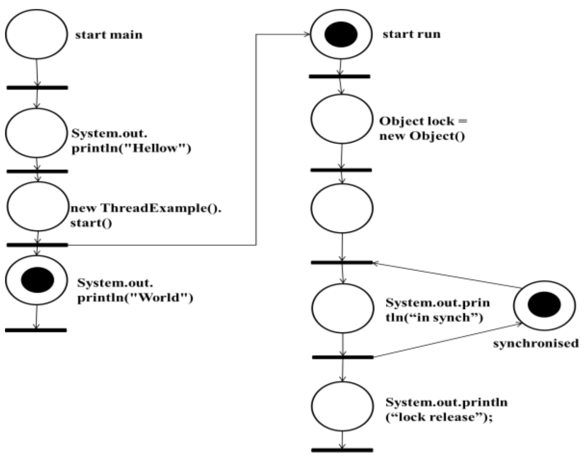


Fig.2. Petri-net modeling of Fig.1

3. Extending Petri-Net

To improve the efficiency of discovering the cause of bugs, we extend Petri-net. This extension can support to understand the behavior of Java programs more intuitively.

A. Identifying Tokens by ID

Because of inability of Petri-net to identify the each token which passes as the same path, we enable Petri-net to identify each token by adding ID to tokens.

B. Expressing the Role of Tokens by Color

The modeling rule which we have described in section II cannot make us to understand that the token expresses thread or locked instance intuitively. Thus, we define a token expressing thread is a white circle and a locked instance is a black circle as shown in Fig.3. This definition makes us easy to understand the role of tokens.

Fig.3. Model for identifying a thread and a locked instance

C. Expressing Locked Instances

At present, it is hard to understand our Petri-net when many locked instances exist in the model or many threads share one locked instance. Also, there is a problem not to know which thread locks which instance.

As an example of these problems, Fig.4 expresses that there are threads numbered ID 1,3,4 which use the locked instance numbered ID 10 and thread numbered ID 2 which use the locked instance numbered ID 20. However, such information is not described in the Petri-net model. As shown in Fig.5, it is impossible to confirm the cause of waiting threads with ID3,4 when the thread with ID1,2 have moved to the next place.

Hence, we extend a token expressing a thread to describe ID which should be locked by each thread. Furthermore, in order to express the state of a thread locking an instance, we extend a token expressing a thread to describe the locked token on the thread token. This extension can express the state of releasing the lock in synchronized block happened by wait() method which is a method of java.lang.Object class.

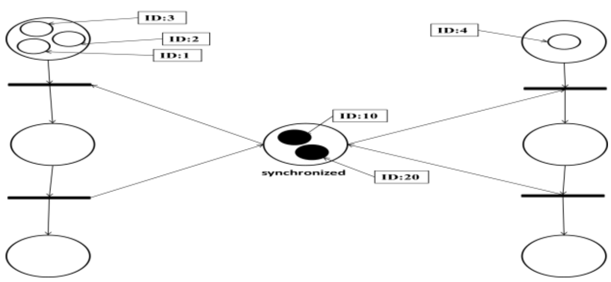


Fig.4. Example of Petri-net with ID

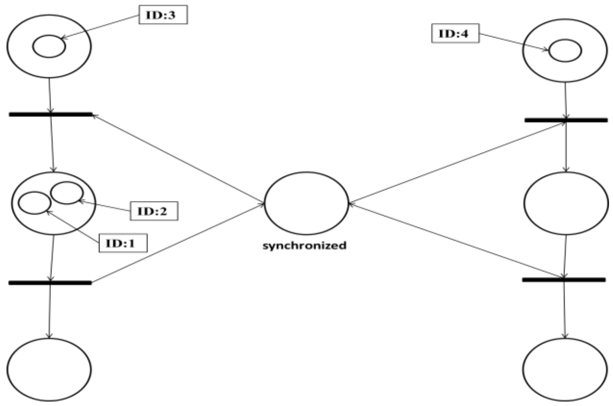


Fig.5. State of waiting for a locked instance

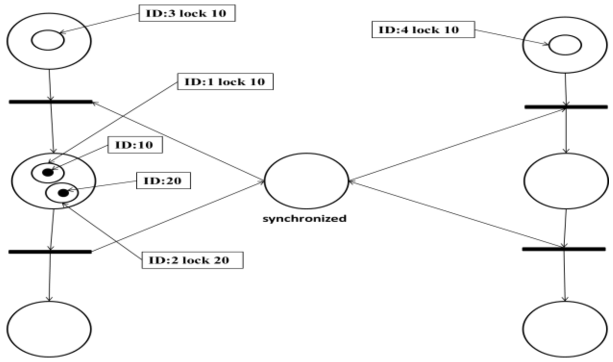


Fig.6. Extended Petri-net for Java

III. CONFIRMATION

We confirm the effectiveness of proposed supporting method for debugging by implementing a tool.

1. The method of confirmation

As an example, we use a multi-threaded program without the process of synchronization written in Java. Fig.7 shows the code of the example.

This program outputs strings by three classes succeeding java.lang.Thread class. FirstNamePrinter class outputs “Shoichiro”, SpacePrinter class outputs “ ”, and SurnamePrinter outputs “Kitano”. The expected result is a string “Shoichiro Kitano”. However, because of this program without process of synchronization, the result is often different from the expected result.

Fig.8 shows a part of the code generated by our tool from the tested code. An unexpected result is discovered by executing the code. Furthermore, our tool can generate the data file of an execution path in testing the code.

By using a data file generated by executing the code shown in Fig.8, we have confirmed the effectiveness of our proposed method by letting students in our laboratory use our tool.

```
public class NakedNamePrinter {

    private final String firstName;
    private final String surName;

    public NakedNamePrinter(String firstName, String surName) {
        this.firstName = firstName; //NakedNamePrinter 1
        this.surName = surName; //NakedNamePrinter 2
        new FirstNamePrinter().start(); //NakedNamePrinter 3
        new SpacePrinter().start(); //NakedNamePrinter 4
        new SurnamePrinter().start(); //NakedNamePrinter 5
    }

    private class FirstNamePrinter extends Thread {
        public void run(){
            System.out.print(firstName); //FirstNamePrinter 1
        }
    }

    private class SpacePrinter extends Thread {
        public void run() {
            System.out.print(' '); //SpacePrinter 1
        }
    }

    private class SurnamePrinter extends Thread {
        public void run() {
            System.out.println(surName); //SurnamePrinter 1
        }
    }

    public static void main(String[] args) {
        new NakedNamePrinter("Shoichiro", "Kitano"); //main 1
    }
}
```

Fig.7. Tested code

```
public class NakedNamePrinter {

    private final String firstName;
    private final String surName;

    public NakedNamePrinter(String firstName, String surName) {
        long t;
        t = System.nanoTime();
        ExecutionWriter.addExecutionData(t, “start  NakedNamePrinter");

        this.firstName = firstName;
        t = System.nanoTime();
        ExecutionWriter.addExecutionData(t, "this.firstName = firstName");

        this.surName = surName;
        t = System.nanoTime();
        ExecutionWriter.addExecutionData(t, "this.surName = surName");

        new FirstNamePrinter().start();
        t = System.nanoTime();
        ExecutionWriter.addExecutionData(t, "new FirstNamePrinter().start()");

        new SpacePrinter().start();
        t = System.nanoTime();
        ExecutionWriter.addExecutionData(t, "new SpacePrinter().start()");

        new SurnamePrinter().start();
        t = System.nanoTime();
        ExecutionWriter.addExecutionData(t, "new SurnamePrinter().start()");
    }
```

Fig.8. A part of the generated code

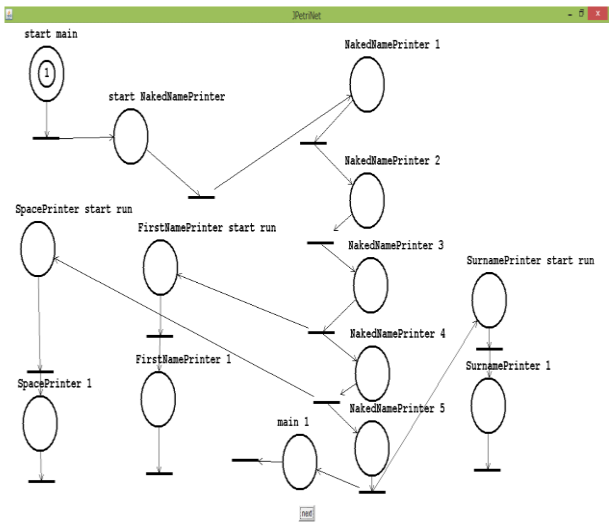


Fig.9. Petri-net generated by tool

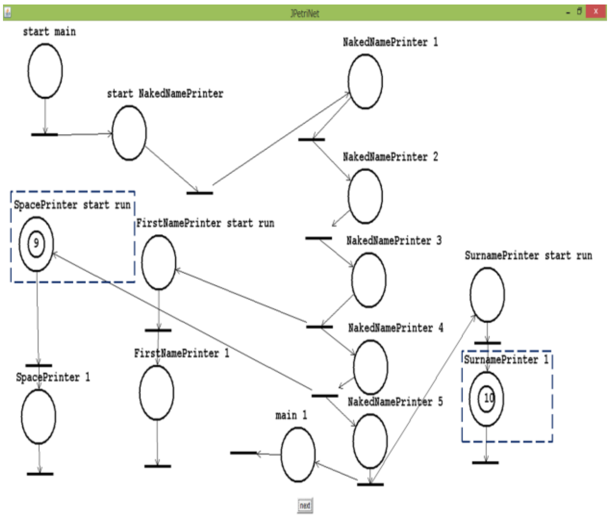


Fig.10. Situation in which an incident occurs

2. Result

Fig.9 shows Petri-net generated by our tool. In Fig.9, the label with number added to the place corresponds with the comment in Fig.7, and the place expresses statement with corresponded the comment.

Fig.10 shows the reproduction of situation when an incident occurs. Primarily, the place should be marked in order of “FirstNamePrinter 1”, “SpacePrinter 1”, “SurnamePrinter 1” in this Petri-net. In Fig.10, Because of a token is on “SpacePrinter start run”, it realizes that “SpacePrinter 1” is not executed and “SurnamePrinter 1” has been executed earlier than “SpacePrinter 1” by the marking of a place with “SurnamePrinter 1”. Here, we can understand that threads are not executed in order our expectation. Thus, we have noticed that there is a defect in synchronized process between threads.

From this, because of reproducing the situation in which an incident occurs and understanding the behavior of a program graphically, it is easy to discover a cause of a bug. Thus, we have confirmed our method is effectively.

3. Comparison with Related Work and Existing Tool

As related work, some researches discover potential bugs of multi-threaded programs by automatically generated Petri-net models[2][3][4]. In these researches, it is possible to discover the existence of bugs by analyzing Petri-net statically. However, this approach cannot discover the cause of bugs.

There is a tool for multi-threaded programs that discover potential bugs by executing the unit testing automatically[5]. This tool cannot reproduce the situation in which an incident occurs to test the multithreaded program by plural interleaving at random. Therefore, discovering cause of bugs is depended on users.

On the other hand, our debugging method can reproduce the same situation in many times, because our method uses the data file of an execution path that was really executed. It makes us easy to discover the cause of bugs. Therefore, it may be said that our proposing method is effective for removing the potential bug of the multi-thread program.

IV. CONCLUSION

In this paper, we have proposed supporting method for debugging to reproduce Java multi-threaded programs by Petri-net. And we have extended Petri-net to understand the behavior of Java programs more. We think that this extension improves the efficiency of debugging. Furthermore, we have shown the effectiveness of our method by adapting a tool to an example and comparing with related works and the existing tool.

Future issues are as follows.

- Improving of the adaptation range of our tool
- Confirming effectiveness of extended Petri-net

REFERENCES

[1]J. K. Ousterhout, Why Threads Are A Bad Idea (for most purposes), Presentation given at the 1996 Usenix Annual Technical Conference, January 1996 http://www.softpanorama.org/People/Ousterhout/Threads/

[2]Krishna M. Kavi, Alireza Moshtaghi, Deng-jyi Chen(2002), Modeling Multithreaded Applications Using Petri Nets International Journal of Parallel Programming, Vol. 30, No. 5, October 2002: 353-371

[3] Govindarajan, R, Suciu, F., Zuberek, W.M, Timed Petri net Models of Multithreaded Multiprocessor Architectures, IEEE Proceedings of the Seventh International Workshop on Petri Nets and Performance Models, June 1997: 153-162

[4] Hongwei Liao, Yin Wang, Hyoun Kyu Cho, Jason Stanley, Terence Kelly, Stéphane Lafortune, Scott Mahlke, Spyros Reveliotis, Concurrency bugs in multithreaded software: modeling and analysis using Petri nets, Discrete Event Dynamic Systems, June 2013, Vol. 23, Issue 2: 157-195

[5] ConTest - A Tool for Testing Multi-threaded Java Applications, https://www.research.ibm.com/haifa/projects/verification/contest/

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Proposal of a Supporting Method to Generate a Decision Table from the Formal Specification

Kenta Nishikawa*, Tetsuro Katayama*, Yoshihiro Kita†, Hisaaki Yamaba* and Naonobu Okazaki*

*University of Miyazaki, 1-1 Gakuen-kibanadai nishi, Miyazaki, 889-2192 Japan
†Kanagawa Institute of Technology, 1030 Shimo-ogino, Kanagawa, 243-0292 Japan
nishikawa@earth.cs.miyazaki-u.ac.jp, kat@cs.miyazaki-u.ac.jp, kita@earth.cs.miyazaki-u.ac.jp, yamaba@cs.miyazaki-u.ac.jp, oka@cs.miyazaki-u.ac.jp

Abstract: In recent years, the software quality becomes more important because the system becomes large scale and high performance. In general, many defects are embedded in the upstream process of the software development. As one reason of the above, specifications include ambiguous description. As a means for writing specifications strictly, formal methods are proposed. By the way, as one of test design techniques, the decision table is proposed. However, it takes much time and effort to extract test items and understand contents written on specifications in designing manually the decision table. This paper proposes a supporting method to generate a decision table from the formal specification in order to improve efficiency of the test design with formal methods. We have implemented a supporting tool to generate a decision table. It automatically generates a skeleton decision table from the formal specification. By using the tool, it is considered that the efficiency of the test design is improved.

Keywords: formal method, VDM++, test design, decision table, automatic generation.

I. INTRODUCTION

In recent years, the software quality cannot be maintained with the conventional software development methods because the system becomes large scale and high performance. At the same time, effect of defects in the system becomes one of the major social problems with the economy and life [1].

Hence, the software quality becomes more important. A demand for reliability and safety of the system is becoming increasingly.

In general, many defects are embedded in the upstream process of the software development. As one reason of the above, each step in the software development process moves to the next step with specifications included ambiguous description. Therefore, specifications should be written strictly. As a means for writing specifications strictly, formal methods [2] are proposed. The formal methods are a means for using strict specifications in each step in the software development process. They express the system with a specification description language based on mathematical logic. Using the formal methods can remove defects or ambiguity of the specifications. They attract attention as a means to improve software quality.

By the way, as one of the test design techniques, the decision table [3] is proposed in the testing process of the software cycle. The decision table uses a matrix divided the logical relationships in specifications into items of conditions and actions. However, it takes much time and effort to extract test items and understand contents written on specifications in designing manually the decision table. It is no exception even if you write strict specifications with formal methods.

This paper proposes a supporting method to generate a decision table from the formal specification in order to improve efficiency of the test design with formal methods. This paper uses the formal specification description language VDM++ which is the lightweight formal methods VDM (Vienna Development Method) to write the formal specification. The proposed method supports to generate a decision table by extracting the logical relationship of the conditions and actions from a specification written with VDM++. We have implemented a supporting tool to generate a decision table. It automatically generates a skeleton decision table from the formal specification. By using the tool, it is considered that the efficiency of the test design is improved.

Here, the skeleton decision table which the tool generates has condition items, action items, and combinations of truth-values for the condition items. This means that a user must add to write truth-values in action items of the generated skeleton decision table to complete the decision table.

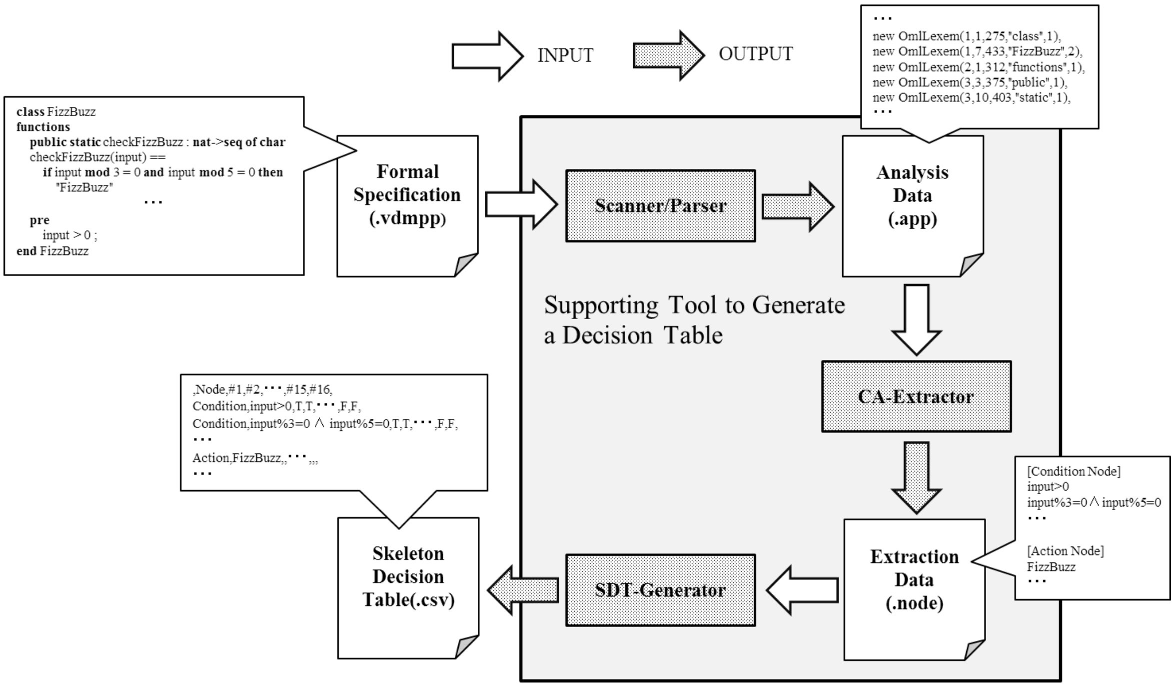


Fig. 1. The flow of supporting method to generate a decision table from the formal specification

II. SUPPORTING METHOD TO GENERATE A DECISION TABLE FROM THE FORMAL SPECIFICATION

Fig.1 shows the flow of the proposed method. The proposed method consists of three parts as follows.

i. Scanner/Parser
ii. CA-Extractor
iii. SDT-Generator

We explain the flow of the proposed method as follows.

1. A user prepares a formal specification after it is finished the syntax checking and type checking.
2. A user runs the supporting tool to generate a decision table after the user specified a formal specification.
3. The supporting tool to generate a decision table runs Scanner/Parser, which uses a formal specification specified by the user as an input.
4. Scanner/Parser generates the analysis data by analyzing the formal specification.
5. CA-Extractor generates the extraction data extracted conditions and actions of the formal specification from the analysis data by generated Scanner/Parser.
6. SDT-Generator generates a skeleton decision table from the extraction data by generated CA-Extractor, and output it as a csv file.

We explain each part as follows.

1. Scanner/Parser

Scanner/Parser is a VDM++ analyzer for Overture implemented by Marcel Verhoef [4]. Scanner/Parser reads the formal specification specified by a user. Scanner/Parser generates the analysis data (.app) from the formal specification.

The analysis data has analysis information (e.g., token, token ID, and abstract syntax tree) needed by CA-Extractor.

2. CA-Extractor

CA-Extractor extracts conditions and actions from the analysis data generated by Scanner/Parser, and writes them to extraction data (.node). CA-Extractor extracts conditions and actions only from pre-conditions expressions, post-conditions expressions and if-then-else expressions.

Table 1 shows extraction processes of conditions and actions.

3. SDT-Generator

SDT-Generator generates a skeleton decision table from the extraction data by generated CA-Extractor, and outputs it as a csv file. Output format of a skeleton decision table is adopted CSV format because it is versatile.

Table 2 shows an example of a skeleton decision table which SDT-Generator outputs. SDT-Generator describes condition items and action items on the second column of a skeleton decision table. SDT-Generator describes truth-values in condition items on and after the third column of a skeleton decision table. The skeleton decision table of truth-values in action items is empty, because the supporting tool to generate a decision table cannot generate truth-values in action items.

Table 1. Extraction processes of conditions and actions

■ Extraction of conditions
● Conditions extraction processes are pre-conditions expressions, post-conditions expressions and if-then-else expressions. ● To extract conditions, CA-Extractor extracts between from a start token (if, else, pre, post) to an end token (then, ;) as a condition. e.g., if (elseif) *conditionA* then → *conditionA* pre *conditionB* ; → *conditionB*
■ Extraction of actions
● Actions extraction processes are if-then-else expressions. ● To extract actions, CA-Extractor extracts just after a start token (then or else) as an action. e.g., else *actionA* → *actionA*

Table 2. An example output of the skeleton decision table

	Node	#1	#2	…	#15	#16
Condition	A>0	T	T	…	F	F
Condition	B=true	T	T	…	F	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Action	NAT	null	null	null	null	null
⋮	⋮	⋮	⋮	⋮	⋮	⋮

III. APPLICATION EXAMPLE

We have implemented a supporting tool to generate a decision table, in order to realize the proposed method. It supports the test design. It automatically generates a skeleton decision table from the formal specification.

We confirm that it works properly by adapting it to an example. Specifically, we confirm two items as follows.

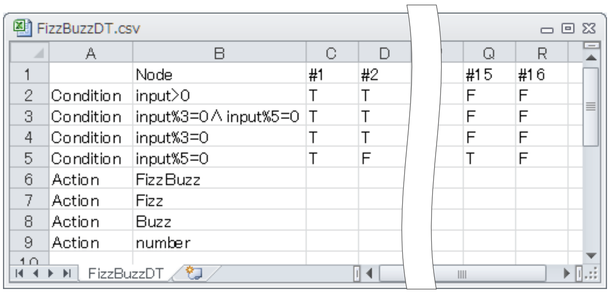
i. To extract conditions and actions from a formal specification

ii. To generate truth-values in condition items

Fig.2 shows the formal specification which is used as an application example. It is a specification of FizzBuzz, which is described in VDM++. FizzBuzz is a group word game. Players take turns to count

```
1 class FizzBuzz
2 functions
3   public static checkFizzBuzz : nat->seq of char
4   checkFizzBuzz(input) ==
5     if input mod 3 = 0 and input mod 5 = 0 then
6       "FizzBuzz"
7     elseif input mod 3 = 0 then
8       "Fizz"
9     elseif input mod 5 = 0 then
10      "Buzz"
11    else
12      "number"
13  pre
14    input > 0 ;
15 end FizzBuzz
```

Fig. 2. The formal specification of FizzBuzz



	A	B	C	D	Q	R
1		Node	#1	#2	#15	#16
2	Condition	input>0	T	T	F	F
3	Condition	input%3=0∧input%5=0	T	T	F	F
4	Condition	input%3=0	T	T	F	F
5	Condition	input%5=0	T	F	T	F
6	Action	FizzBuzz				
7	Action	Fizz				
8	Action	Buzz				
9	Action	number				

Fig. 3. A part of the skeleton decision table

incrementally, replacing any number divisible by three with the word "fizz", and any number divisible by five with the word "buzz".

The formal specification of FizzBuzz in Fig.2 is described a pre-condition (i.e., in the 13th and 14th line) which means a value of the input is more than zero.

Fig.3 shows a part of the skeleton decision table generated by the supporting tool, and it is the result of applying the formal specification to the supporting tool.

We have confirmed that the skeleton decision table in Fig.3 is satisfied i and ii. Therefore, we have confirmed that the supporting tool to generate a decision table works properly.

IV. DISCUSSION

This paper has proposed a supporting method to generate a decision table from the formal specification in order to improve efficiency of the test design with formal methods. We have implemented a supporting tool to generate a decision table, in order to realize the proposed method. It automatically generates a skeleton decision table from the formal specification. By using the tool, it is considered that the efficiency of the test design is improved.

We discuss our proposed method in this chapter.

1. Problems of the supporting tool to generate a decision table

We show problems of the supporting tool to generate a decision table as follows.

- Generating truth-values in action items

The supporting tool to generate a decision table cannot automatically generate truth-values in action items. This means that a user must add to write truth-values in action items of the skeleton decision table, in order to complete the decision table. If condition items increases, it takes much time and effort to describe manually truth-values in action items.

- Supporting to each syntax: cases expressions, for loop, and while loop, and so on

There is a limit to the formal specification applicable to the supporting tool to generate a decision table. Specifically, the supporting tool does not support to all syntax in VDM++. Hence, targets of conditions and actions extracted from a formal specification are pre-conditions expressions, post-conditions expressions and if-then-else expressions.

2. Evaluation of the proposed method

We confirm usefulness of the proposed method by using the generated skeleton decision table in chapter III. First, we have completed the decision table by describing truth-values in action items of the skeleton decision table by hand. Second, we have implemented FizzBuzz in Java from the formal specification used as an application example in chapter III.

We have tested FizzBuzz program by using the completed decision table, and at the result, we have been able to test all combinations of conditions and actions in FizzBuzz. Therefore, we have confirmed the usefulness of the proposed method.

3. Discussion about the supporting tool to generate a decision table

The supporting tool to generate a decision table extracts conditions and actions from a formal specification, and automatically generates a skeleton decision table. It is able to reduce the time and effort when testers manually design the testing.

Few researches of test design from the formal specification are reported, and the method is not well established. In addition, some tools to automatically generate a decision table are proposed [5, 6], but no tool to support the test design from a formal specification such as our proposed method.

By using the supporting tool to generate a decision table, it is considered that the efficiency of the test design from a formal specification is improved.

V. CONCLUSION

This paper proposed a supporting method to generate a decision table from the formal specification in order to improve efficiency of the test design with formal methods. We have implemented a supporting tool to generate a decision table, in order to realize the proposed method. The supporting tool generates a skeleton decision table from the formal specification by extracting conditions and actions of the formal specification.

We have confirmed that it generates the skeleton decision table and truth-value in conditions and actions of the skeleton decision table. By using the supporting tool to generate a decision table, it is considered that the efficiency of the test design from a formal specification is improved.

Future issues are as follows.

- Generating truth-values in action items
- Supporting to each syntax: cases expressions, for loop, and while loop, and so on
- Comparing with a decision table described by hand

REFERENCES

[1] John Fitzgerald, Peter Gorm Larsen, Paul Mukherjee, Nico Plat, Marcel Verhoef (2005), Validated Designs for Object-Oriented Systems, Springer.
[2] Shin Nakajima (2007), Formal Methods as Software Engineering Tools (in Japanese), NII Technical Report, NII-2007-007J.
[3] ISO 5806, Specification of single-hit decision tables.
[4] A Scanner/Parser for the Overture Toolset, http://overturetool.hosting.west.nl/twiki/bin/view/Main/OvertureParser/ (accessed November 29, 2013).
[5] CEGTest, http://softest.jp/tools/CEGTest/ (accessed November 29, 2013).
[6] PictMaster, http://sourceforge.jp/projects/pictmaster/ (accessed November 29, 2013).

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Development of a Campus Guide System based on Augmented Reality

M. Tabuse and A. Tada

Graduate School of Life and Environmental Sciences, Kyoto Prefectural University
1-5 Hangi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan
Tel : 81-75-703-5453; Fax : 81-75-703-5453
e-mail:tabuse@kpu.ac.jp

***Abstract*:** This paper presents a campus guide system based on augmented reality (AR). AR is a technology of displaying CG overlaid on a real scene. AR is a fusion of a real world and a virtual world of human interfaces. We propose a campus guide system based on markerless AR. This system displays information of a university overlaid on a real scene using a note PC and a web camera. In addition, this system presents a variety of events, for example a university festival and a welcome party, as 3D virtual objects overlaid on a real scene. Using this system, visitors can understand a university life more deeply. We have developed a campus guide system based on Parallel Tracking and Multiple Map system (PTAMM) and construct 3D virtual objects using Visual Structure from Motion System (VisualSFM). And we have confirmed the effectiveness of the system in our campus.

Keywords: Campus Guide System, Augmented Reality, AR, PTAMM, VisualSFM**.**

I. INTRODUCTION

Recently, augmented reality (AR) has been actively researched. AR is a technology of displaying CG overlaid on a real scene. Therefore, AR is a fusion of a real world and a virtual world of human interfaces. There are three types of AR, AR with GPS and an attitude sensor, marker based AR and markerless AR. However, GPS does not indicate location information with sufficient accuracy for AR. In marker based AR, we must set up a particular marker and it is eyesore. Therefore, markerless AR is the most promising.

In this paper, we propose a campus guide system based on markerless AR. This system displays information of a university overlaid on a real scene using a note PC and a web camera. In addition, this system presents a variety of events, for example a university festival and a welcome party, as 3D virtual objects overlaid on a real scene. Using this system, visitors can understand a university life more deeply. We have developed a campus guide system based on Klein’s Parallel Tracking and Multiple Map system (PTAMM) [1] and constructed 3D virtual objects using Wu’s Visual Structure from Motion System (VisualSFM) [2]. And we have confirmed the effectiveness of the system in our campus. First, we have captured images of locations to guide and create 3D maps in our system. Then, we have moved along the guide course while capturing a campus and confirmed that the system display the guide contents properly.

II. AUGMENTED REALITY

AR is a technology of displaying CG overlaid on a real scene. In order to realize AR, it is an important matter how to overlay digital information in the real environment. In particular, the estimate of a position and a pose of a camera is the most important processing. The methods to estimate a position and a pose of a camera are as follows.

(1) The method using GPS and an attitude sensor [3][4].
(2) The method using a particular marker [5].
(3) The makerless method [1].

When using GPS and an attitude sensor, the position of the camera is estimated based on the position information of GPS. However, it is difficult to display CG overlaid on a real scene precisely because GPS does not indicate location information with sufficient accuracy for AR. When using a particular marker, AR determines the coordinate based on the marker and displays CG in this coordinate. However, the markers are eyesores in a real scene, so that it is difficult to place particular markers on building and streets. In markerless AR, AR detects natural feature points everywhere in a real scene, estimates camera position and pose, determines the coordinate and displays CG in this coordinate. One of the most promising and stable markerless AR is PTAMM [1], which an extension of PTAM (Parallel Tracking and Mapping) [6]. Thus, we develop a campus guide system based on PTAMM.

III. PTAM AND PTAMM

1. PTAM

PTAM splits tracking and mapping into two separate tasks, processed in parallel threads.

A. Initial 3D map

At first a camera position and pose are selected by user and its image and FAST corner feature list become the first keyframe. Then the camera is moved to a new position and its image and feature list become the second keyframe. From these keyframe, an initial 3D map is built using a stereo technique so that a dominant plane defines $z=0$ surface in the world.

B. Tracking

Tracking thread receives images from the camera and maintains a real-time estimate of the camera position and pose relative to the 3D map using feature lists. The method of estimate of the camera position and pose is as follows.

1. A new image is obtained from the camera, and a prior pose estimate is generated from a motion model.
2. 3D map feature points are projected into the image according to the prior pose estimate.
3. The camera pose is updated from feature points matching in the image.

C. Mapping

As the camera is moved, new images and features are added as new keyframes whenever (a) the camera has translated by a minimum distance from any previous keyframe, (b) around 20 frames have passed since the previous keyframe. Then mapping thread expands 3D map.

2. PTAMM

Since PTAM is a system for small-scale space, it can not be applied to applications for multiple rooms and large-scale space. PTAMM is an extension of PTAM for large-scale space. It builds multiple maps of the environment. Then it selects the correct local map close to a current camera image. The maps built by PTAMM are stored in PC. Thus we can display particular 3D virtual objects overlaid on the scene according to the selected maps.

IV. CAMPUS GUIDE SYSTEM

We have developed an AR campus guide system based on PTAMM. This system displays information of a university overlaid on a real scene using a note PC and a web camera. In addition, this system presents a variety of events, for example a university festival and a welcome party, as 3D virtual objects overlaid on a real scene. Using this system, visitors can understand a university life more deeply.

1. 3D CG character

We create a 3D CD mascot character of the university using 3D modeling software Metasequoia [7] shown in Figure 1. This character plays a role of guide of the university in the AR system.

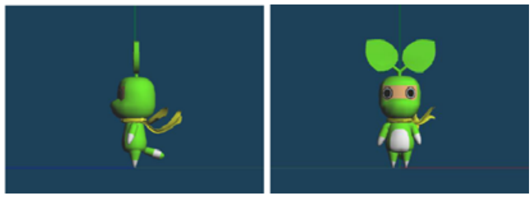


Fig. 1. 3D mascot character

2. 3D virtual object

In order to construct 3D virtual objects we have used VisualSFM [2]. VisualSFM is a GUI application for 3D reconstruction using structure from motion (SFM). SFM is reconstructing algorithm of 3D structure from 2D image sequences which consist of camera images of an object while changing the viewpoint of the camera.

In our system we have prepared 40 image of the stage of our university festival 2012. Figure 2 shows examples of images and figure 3 shows 3D object reconstructed by VisualSFM.

V. EXPERIMENTS

We have installed the AR campus guide system into a note PC and moved along the guide route with a PC and a web camera while capturing images of locations to guide and created 3D maps in our system. The guide route is from the university main gate to a library through the side of the No.1 building. Then we have moved along the same route with the PC and a web camera. We have verified that the system selects the 3D maps and displays guide contents overlaid on a real scene appropriately.

In our experiments we have used a note PC (Panasonic CF-SX1; CPU : Core i5 2540, Memory : 4GB), a web camera (Logicool HD Webcam C270). For programming, we use Microsoft Visual C++ 2008, markerless AR software PTAMM, image processing software library OpenCV2.4.3.

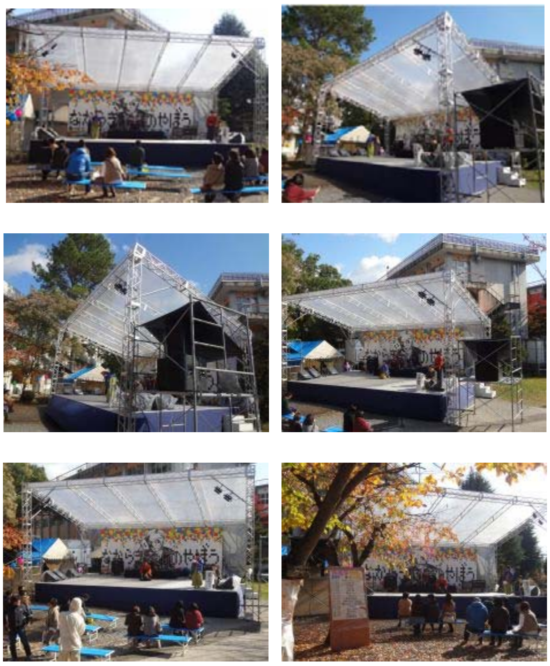


Fig. 2. Examples of images

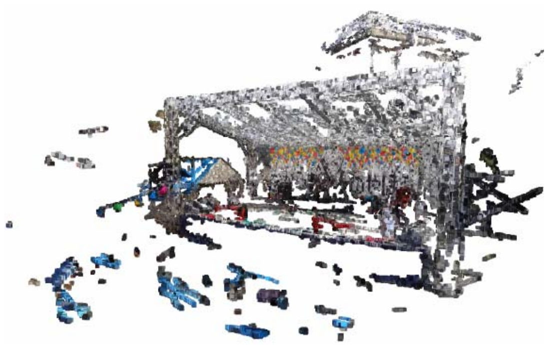


Fig. 3. 3D virtual object

VI. Results

We have made 3D maps at (1) the main gate, (2) the bulletin board near the lecture building, (3) the street in front of No.1 building and (4) the tree in front of the lecture building. Then we have created guide contents for each 3D map. Figs. 4 - 9 show guide contents overlaid on a real scene. In Fig. 4 and Fig. 5 the mascot character bows and the outline of the university is shown. In Fig. 6 we turn the camera to the No.1 building. Then the system finds the next map, shows a tag of the No.1 building and displays the outline of the No.1 building in Fig. 7. Fig. 8 shows a state of searching the next map. In Fig. 9 the system finds the next map and displays a related guide contents.

Fig. 10 shows the virtual object overlaid on a real scene. Creating a guide contents, we can change the size and orientation of a virtual object. Since AR estimates a camera poison and pose, a virtual object seems to be in actual space, shown in Fig. 11.

VII. CONCLUSION

In this paper we have developed an AR campus guide system based on PTAMM. We have confirmed that the system selects correct maps in the guide route and displays the guide contents for each scene properly.

In our experiments we have found the following subjects.

(1) It is difficult to create an initial map especially on an outdoor scene.

(2) We find the feature which is easy to detect on an outdoor scene.

(3) A 3D virtual object is made more realistic.

(4) We make simple and clear guide contents.

We will solve these subjects and develop a useful AR guide system.

REFERENCES

[1] Castle R O, Klein G, Murray D W, Video-rate Localization in Multiple Maps for Wearable Augmented Reality, Proc of 12th IEEE International Symposium on Wearable Computers (ISWC), pp.15-22 (2008).

[2] Wu C, VisualSFM : A Visual Structure from Motion System, http://ccwu.me/vsfm/

[3] Sekai Camera, Tonchidot Corp., http://sekaicamera.com

[4] Layar, Layar B.V., http://www.layar.com

[5] Kato H, Billinghurst M, Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System, Proc. of 2nd International Workshop on Augmented Reality (IWAR), pp.85-94 (1999).

[6] Klein G, Murray D, Parallel Tracking and Mapping for Small AR Workspaces, Proc. Sixth IEEE and ACM International Symposium on Mixed and Augumented Reality (ISMER), pp.225-234 (2007).

[7] Metasequoia, tetraface Inc., http://www.metaseq.net.

Fig. 4. AR at the main gate.

Fig. 5. AR at the bulletin board near the lecture building.

Fig. 6. AR at No.1 building (1).

Fig. 7. AR at No.1 building (2).

Fig. 8. The state of searching the next map.

Fig. 9. AR at the next map.

Fig. 10. The virtual stage overlaid on a real scene.

Fig. 11. Approach to the virtual stage.

Facial Expression Analysis while Using Video Phone

T. Asada[1], Y. Yoshitomi[1], A. Tsuji[2], R. Kato[1], M. Tabuse[1],

N. Kuwahara[2], and J. Narumoto[3]

1: Graduate School of Life and Environmental Sciences Kyoto Prefectural University,
1-5 Nakaragi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan
t_asada@mei.kpu.ac.jp, yoshitomi@kpu.ac.jp, r_kato@mei.kpu.ac.jp, tabuse@kpu.ac.jp,
2: Graduate School of Science and Technology, Kyoto Institute of Technology,
Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan, diff.dim0505@gmail.com, nkuwahar@kit.ac.jp
3: Graduate School of Medical Science, Kyoto Prefectural University of Medicine,
Kajii-cho, Kawaramachi-Hirokoji, Kamigyo-ku, Kyoto 602-8566, Japan, jnaru@koto.kpu-m.ac.jp

Abstract: To improve the quality of life of elderly people living in a home or healthcare facility, especially in a rural area, we have been developing a method for analyzing the facial expressions of a person using a video phone system (Skype) to talk to another person. The video is recorded and analyzed by image processing software (OpenCV) and the newly proposed feature vector of facial expression, which is extracted in the mouth area by applying 2D-DCT. The facial expression intensity, defined as the norm of the difference vector between the feature vector of the neutral facial expression and that of the observed expression, can be used to analyze a change of facial expression. The judgment of speaking is performed by using the intensity of the sound wave. The experimental results show the usefulness of the proposed method. We intend to use the proposed method of facial expression analysis to develop a method for estimating the emotions or mental state of people, particularly elderly patients.

Keywords: Facial expression analysis, Video phone, Mouth area extraction**,** OpenCV, and Skype.

I. INTRODUCTION

In Japan, the average age of the population has been increasing, and this trend is expected to continue. Because of this trend, the number of older people with dementia and/or depression, especially those living in rural areas, is increasing very rapidly. Due to the mismatch between the number of patients and the number of healthcare professionals, it is difficult to provide adequate psychological assessments and support for all patients.

Information and communication technology (ICT) is a promising method for overcoming the difficulty caused by the lack of adequate healthcare. In Japan, the first inexpensive connection to the Internet became available only recently in rural areas and high-quality free software, such as Skype [1], is being distributed.

To improve the quality of life (QOL) of elderly people living in a home or a healthcare facility, we propose a method for analyzing the facial expressions of a person using a video phone system (Skype) to speak with another person. In the present study, the phone video is recorded and analyzed by image processing software (OpenCV) and the newly proposed feature vector of facial expression, which is extracted in the mouth area. Moreover, the judgment of speaking is performed by using the intensity of the sound wave.

II. PROPOSED METHOD

1. System overview and outline of the method

As already mentioned, the video phone is Skype [1]. VodBurner (Netralia Pty Ltd.) [2] is introduced for recording the audio and video dialogue. Tapur [3] is also introduced for recording the audio data. Conversations are recorded for the analysis of facial expression. The recorded data are analyzed by image processing software, Open Source Computer Vision Library (Open CV, Intel), for real-time computer vision [4] and the newly proposed feature vector of facial expression described in this paper. The Y component obtained from each frame in the dynamic image is used for measuring the facial expression intensity. The proposed method consists of (1) extraction of the mouth area, (2) measurement of facial expression intensity, and (3) judgment of utterance. In the following subsections, these three are explained in detail.

2. Extraction of mouth area from a dynamic image

First, the face area is extracted from each frame in the dynamic image by the classifiers for a front-view face included in OpenCV. The Haar-like feature parameters and Adaboost algorithm for learning are used as the classifiers [5]. It is assumed that the distance between a subject and the camera generally remains constant during a Skype conversation. However, we observed that the size of the face area tends to increase when the face deviates from a front-view image. In this case, it is difficult for OpenCV to extract the face area. Therefore, the minimum size of dynamic image within a specified period is assumed to be the most likely front view. In the present study, we set one second as the period. Next, by using OpenCV, the mouth area is extracted for the frame selected by the face-area size criterion described above. The mouth area is selected because the difference between the facial expressions of neutral and happy distinctly appear in this area.

3. Measurement of facial expression intensity

For the Y component of the frame selected by the processing described above, the newly proposed feature vector of facial expression is extracted in the mouth area by applying 2-dimensional Discrete Cosine Transform (2D-DCT) for each domain of 8×8 pixels.

The mouth area used for measuring the facial expression intensity has n blocks of 8×8 pixels, where n is obtained as $n = [a/8] \times [b/8]$, a and b denote the number of pixels of the mouth area in the face area obtained by OpenCV in the vertical and horizontal directions, respectively, and $[x]$ equals the maximum integer that does not exceed x.

The high-frequency components of the 2D-DCT coefficients tend to express a minute change in the data, and thus result in the presence of noise. Therefore, we select 15 low-frequency components of the 2D-DCT coefficients, except for a direct current component, as the feature parameters for expressing facial expression (Fig. 1). This selection of 2D-DCT coefficients is popular among researchers in facial expression recognition [6].

Because we do not know the combination of the specific face location and the frequency component of the 2D-DCT coefficients to successfully recognize a facial expression, we adopt the strategy described below.

To gather useful information from the mouth area, we obtain the absolute value of 2D-DCT coefficients, then we obtain the mean of the absolute value for each 2D-DCT coefficient component in the mouth area (Fig. 2). The number of 2D-DCT coefficient components is 15. Therefore, we obtain 15 values as the elements of the feature vector. The facial expression intensity, defined as the norm of the difference vector between the feature vector of the neutral facial expression and that of the observed expression, can be used for analyzing a change of facial expression.

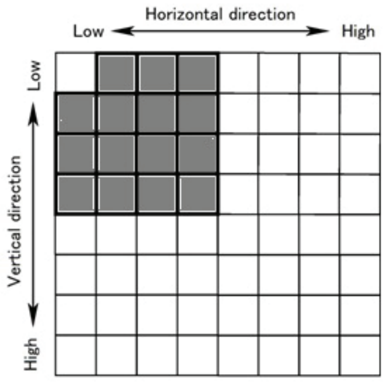


Fig. 1. Special frequency bands used for the analysis

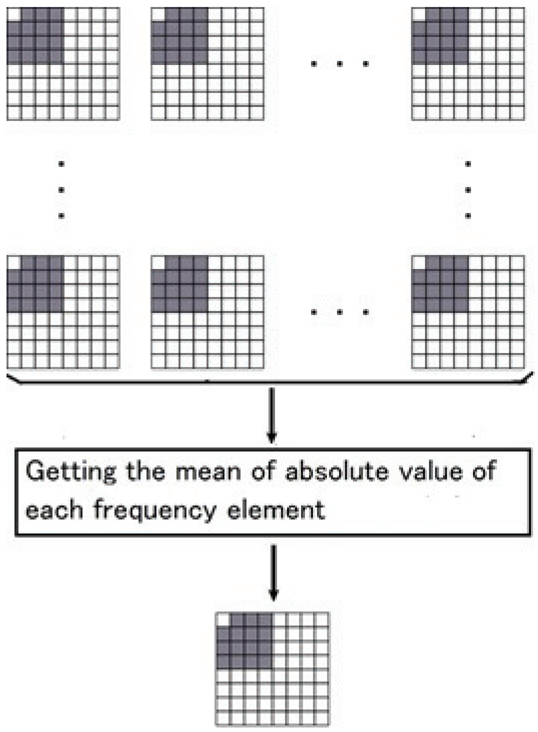


Fig. 2. Schematic diagram of the DCT feature parameter calculation in the mouth area [7]

4. Judgment of utterance

Combining the video signal obtained from Skype with the sound signal, we can distinguish the facial expression with speaking from that without speaking. Based on the method reported in [8]–[10], the sound data are smoothed and sampled to erase noise. The judgment of speaking is performed by using a threshold of the sound intensity. The threshold is determined by the average and the standard deviation of the sound intensity when the subject does not speak in the sound environment where Skype is used. The thresholds for the sound data values are set as $\bar{x}_s - 14\sigma_s$ and $\bar{x}_s + 14\sigma_s$, where $\bar{x}_s$ and σ_s express the average and the standard deviation, respectively, of the sound data value for one second under the condition of no utterance.

Then, every sampled data that falls within $[\bar{x}_s - 14\sigma_s, \bar{x}_s + 14\sigma_s]$ are considered to be the range of no utterance. When at least one sampled datum has a value outside $[\bar{x}_s - 14\sigma_s, \bar{x}_s + 14\sigma_s]$, our system judges that the sound data contain an utterance.

III. EXPERIMENT

1. Condition

Two males (subjects A and B) in their 20s participated in the experiment. Using Skype, the two subjects held a conversation for approximately 80 seconds. The videos saved by VodBurner were transformed into AVI files, and WAV files were saved by Tapur. The AVI files were used for measuring the facial expression intensity. The WAV files were used for judgment of an utterance.

2. Results and discussion

Facial expression intensity changes of subjects A and B during their conversation were recorded (Fig. 3). The timing of utterances (Fig. 4), and the timing of no utterances (Fig. 5) are shown. In both Figs. 4 and 5, face images and mouth images show the characteristic timing positions for the facial expression intensity.

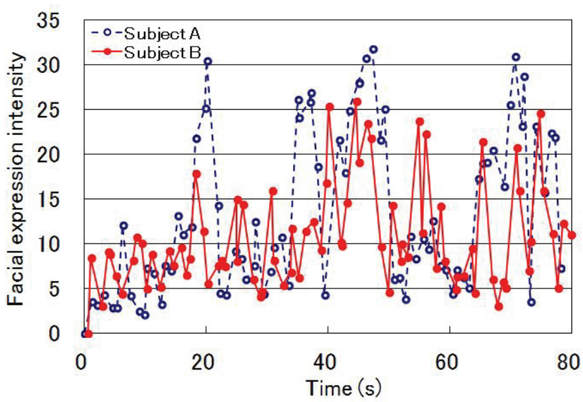


Fig. 3. Facial expression intensity change of subjects A and B during their conversation

Fig. 4. Facial expression intensity changes (upper graph), face images (lower left side), and mouth image (lower right side) of subjects A and B at the timing of utterances during their conversation

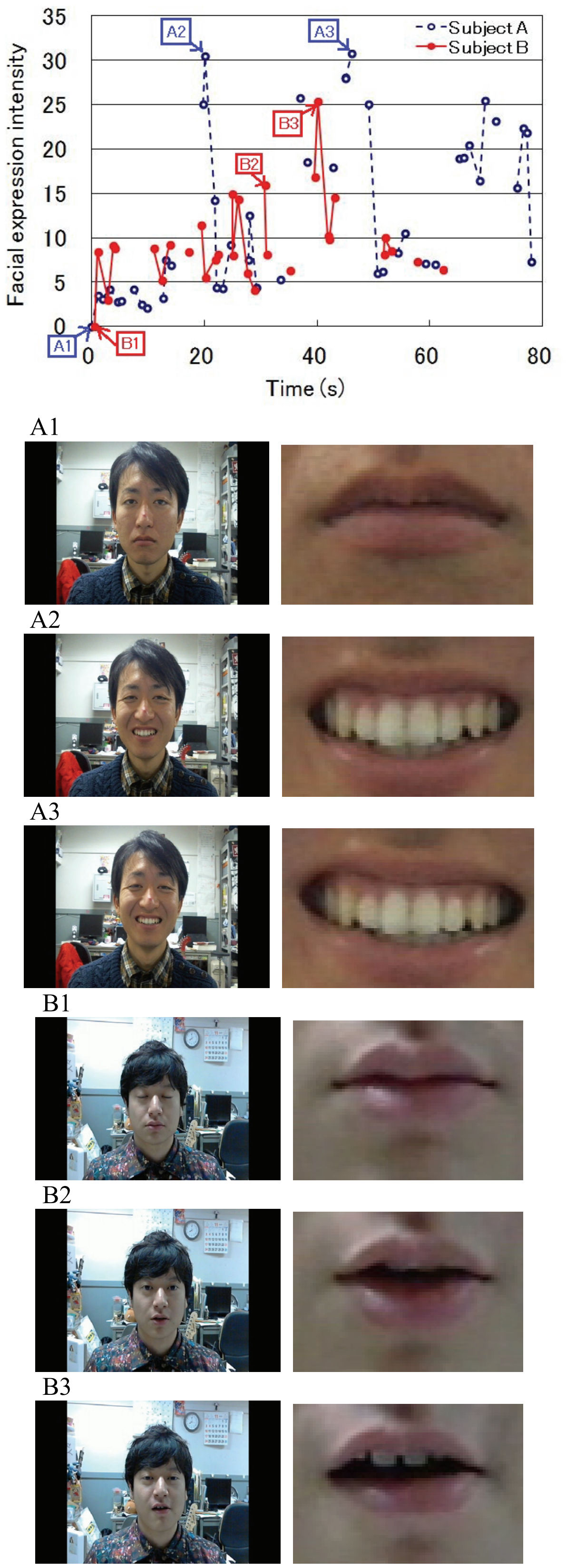


Fig. 5. Facial expression intensity changes (upper graph), face images (lower left side), and mouth images (lower right side) of subjects A and B at the timing of no utterances during their conversation

Subject A expressed four local peaks of facial expression intensity at approximately 20, 37, 47, and 71 seconds from the start point, while subject B expressed six local peaks of facial expression intensity at approximately 40, 45, 55, 65, 71, and 75 seconds from the start point (Fig. 4). For the timing of utterances during their conversation, subject A expressed four local peaks of facial expression intensity at approximately 18, 37, 47, and 71 seconds from the start point, while subject B expressed five local peaks of facial expression intensity at approximately 45, 55, 65, 71, and 75 seconds from the start point (Fig. 5). For the timing of no utterances during their conversation, subject A expressed three local peaks of facial expression intensity at approximately 20, 46, and 70 seconds from the start point, while subject B expressed one local peak of facial expression intensity at approximately 40 seconds from the start point (Fig. 6). Subject B mainly made an utterance during the second half of the conversation (Fig. 4), while the facial expression intensity increased gradually for the timing of no utterances during the first half of the conversation (Fig. 5). Subject A expressed two local peaks of facial expression intensity at approximately 18 and 37 seconds from the start point, and subject B expressed no local peaks of facial expression intensity when the data were limited to the timing of utterances during the first half of the conversation (Fig. 4). Just after the only local peak of facial expression intensity having no utterance at approximately 40 seconds from the start point, subject B expressed continual peaks of facial expression intensity with utterances.

The images of the face and mouth areas at the characteristic timing points show that the proposed method can quantitatively express the facial expression (Figs. 4 and 5).

IV. CONCLUSION

We proposed a method for analyzing the facial expressions of a person while speaking with a video phone system (Skype). The recorded video is analyzed by image processing software (OpenCV) and the newly proposed feature vector of facial expression, which is extracted in the mouth area by applying 2D-DCT. The facial expression intensity, defined as the norm of the difference vector between the feature vector of the neutral facial expression and that of the observed expression, can be used to analyze the change of facial expression. The judgment of speaking is performed by using the intensity of the sound wave. The experimental results show the usefulness of the proposed method.

Acknowledgment

We would like to thank the subjects who participated in the experiments. This research is supported by SCOPE (122307003) of the Ministry of Internal Affairs of Communications of Japan and COI STREAM of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

[1] Skype. http://www.skype.com/ Accessed 5 November 2013

[2] VodBurner. http://www.vodburner.com/ Accessed 1 December 2013

[3] Tapur. http://www.tapur.com/jp/ Accessed 6 December 2013

[4] Open CV. http://opencv.org/ Accessed 1 December 2013

[5] Viola P and Jones MJ (2001) Rapid Object Detection Using a Boosted Cascade of Simple Features. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.1, pp.511-518

[6] Sakaguchi T and Morishige S(1997) Real-Time Facial Expression Recognition Based on the 2-Dimensional DCT. Trans IEICE J80-D-Ⅱ(6):1547-1554

[7] Yoshitomi Y, Tabuse M, Asada T (2012) Facial expression recognition using thermal image processing. In: Carvalho VH (Ed) Image processing: methods, applications and challenges. Nova Science Publisher, New York, pp. 57-85

[8] Ikezoe F, Nakano M, Yoshitomi Y, et al (2005) Facial expression recognition using thermal face image automatically acquired in speaking (in Japanese). In: Proceedings of Human Interface Symposium 2005, Fujisawa, Kanagawa, Japan, Sep.15-17, 2005, pp 7-12

[9] Nakano M, Yoshitomi Y, and Tabuse M (2006) Efficient facial expression recognition using thermal face image in speaking and its application to analysis of individual variations (in Japanese). In: Proceedings of Human Interface Symposium 2006, Kurashiki, Okayama, Japan, Sep.25-28, 2006, pp 1151-1156

[10]Nakano M (2008) Robust facial expression recognition for various speakers (in Japanese). Master's thesis, Dept. of Environmental Information Graduate School of Human Environmental Sciences Kyoto Prefectural University, pp 1-70

Method of Facial Expression Analysis Using Video Phone and Thermal Image

Y. Yoshitomi, T. Asada, R. Kato, and M. Tabuse

Graduate School of Life and Environmental Sciences Kyoto Prefectural University,
1-5 Nakaragi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan
yoshitomi@kpu.ac.jp, {t_asada, r_kato}@mei.kpu.ac.jp,tabuse@kpu.ac.jp

Abstract: To improve the quality of life of elderly people living in a home or healthcare facility, especially in a rural area, we have been developing a method for analyzing the facial expressions of a person using a video phone system (Skype) to speak with another person. In the present study, we proposed a method for analyzing facial expressions of a person using the video phone system to talk to another person. The recorded video is analyzed by thermal image processing and the newly proposed feature vector of facial expression, which is extracted in the mouth area by applying 2D-DCT. The facial expression intensity, defined as the norm of the difference vector between the feature vector of the neutral facial expression and that of the observed expression, can be used to analyze a change of facial expression. The judgment of utterance is performed by using the intensity of the sound wave. The experimental results show the usefulness of the proposed method. We intend to use the proposed method of facial expression analysis to develop a method for estimating the emotions or mental state of people, especially elderly patients.
Keywords: Facial expression analysis, Video phone, Area of mouth and jaw, Thermal image, and Skype.

I. INTRODUCTION

In Japan, the average age of the population has been increasing, and this trend is expected to continue. Because of this trend, the number of older people with dementia and/or depression, especially those living in rural areas, is increasing very rapidly. Due to the mismatch between the number of patients and the number of healthcare professionals, it is difficult to provide adequate psychological assessments and support for all patients.

Information and communication technology (ICT) is a promising method for overcoming the difficulty caused by the lack of adequate healthcare. In Japan, the first inexpensive connection to the Internet became available only recently in rural areas and high-quality free software, such as Skype [1], is being distributed.

Although the mechanism for recognizing a facial expression has received considerable attention in the field of computer vision research, it still falls far short of human capability, especially from the viewpoint of robustness under widely varying lighting conditions. One reason for this lack of robustness is that nuances of shade, reflection, and local darkness influence the accuracy of facial expression recognition through the inevitable change of gray levels. To avoid this problem and develop a method for facial expression recognition that is applicable under widely varying lighting conditions, we use images produced by infrared rays (IR), which reveal the thermal distribution of the face [2]-[13].

To improve the quality of life (QOL) of elderly people living in a home or healthcare facility, we have been developing a method for analyzing the facial expressions of a person using a video phone system to speak with another person. In the present study, we proposed a method for analyzing the facial expressions of a person using the Skype video phone system. Instead of a visible ray image, we use an IR image that is applicable under widely varying lighting conditions. Moreover, the judgment of utterance is performed by using the intensity of the sound wave.

II. PROPOSED METHOD

1. System overview and outline of the method

As already mentioned, the video phone is Skype [1]. VodBurner (Netralia Pty Ltd.) [14] is introduced for recording the audio and video dialogue. Tapur [15] is also introduced for recording the audio data. Conversations are recorded for the analysis of facial expression. The recorded data are analyzed by thermal image processing and the newly proposed feature vector of facial expression described in this paper. The proposed method consists of (1) extraction of the area of the mouth and jaw, (2) measurement of facial expression intensity, and (3) judgment of utterance. In the following subsections, these three are explained in detail.

2. Extraction of area of mouth and jaw from a dynamic image

The frame extracted every 0.1 second in the dynamic image is used for thermal image processing. Six face areas (Fig. 1) are extracted by the thermal image processing reported in our study [11]. The area of the mouth and jaw is selected because the difference between the facial expressions of neutral and happy distinctly appears in this area.

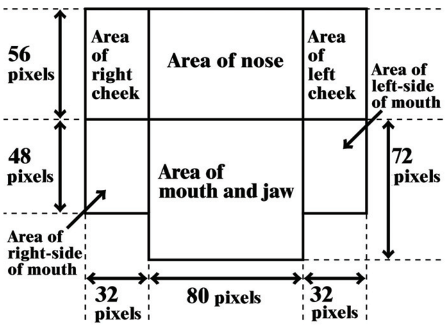


Fig. 1. Blocks for extracting face areas in the thermal image [11]

3. Measurement of facial expression intensity

For the extracted frame, the newly proposed feature vector of facial expression is extracted in the area of the mouth and jaw by applying a two-dimensional discrete cosine transform (2D-DCT) for each domain of 8×8 pixels.

The high-frequency components of the 2D-DCT coefficients tend to express a minute change in the data, and thus result in the presence of noise. Therefore, we select 15 low-frequency components of the 2D-DCT coefficients, except for a direct current component, as the feature parameters for expressing facial expression (Fig. 2). Because we do not know the combination of the specific face location and the frequency component of the 2D-DCT coefficients to successfully recognize a facial expression, we adopt the strategy described below.

To gather useful information from the mouth area, we obtain the absolute value of 2D-DCT coefficients, then we obtain the mean of the absolute value for each 2D-DCT coefficient component in the mouth area (Fig. 3). The number of 2D-DCT coefficient components is 15. Therefore, we obtain 15 values as the elements of the feature vector. The facial expression intensity, defined as the norm of the difference vector between the feature vector of the neutral facial expression and that of the observed expression, can be used for analyzing a change of facial expression.

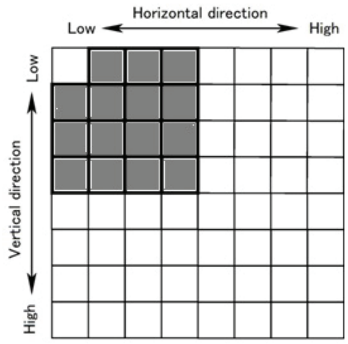


Fig. 2. Special frequency bands used for the analysis

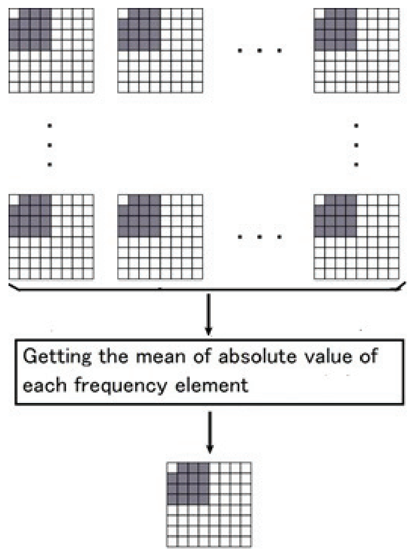


Fig. 3. Schematic diagram of the DCT feature parameter calculation in the mouth area [11]

4. Judgment of utterance

Combining the video signal obtained from Skype with the sound signal, we can distinguish the facial expression with speaking from that without speaking. Based on the method reported in [16]–[17], the sound data are smoothed and sampled to erase noise. The judgment of speaking is performed by using a threshold of the sound intensity. The threshold is determined by the average and the standard deviation of the sound intensity when the subject does not speak in the sound environment where Skype is used. The thresholds for the sound data values are set as $\bar{x}_s - 14\sigma_s$ and $\bar{x}_s + 14\sigma_s$, where $\bar{x}_s$ and σ_s express the average and the standard deviation, respectively, of the sound data value for one second under the condition of no utterance.

Then, all sampled data that fall within $\left[\bar{x}_s - 14\sigma_s, \bar{x}_s + 14\sigma_s\right]$ are considered to be the range of no utterance. When at least one sampled datum has a value outside $\left[\bar{x}_s - 14\sigma_s, \bar{x}_s + 14\sigma_s\right]$, our system judges that the sound data contain an utterance.

III. EXPERIMENT

1. Condition

The thermal image was produced by a thermal video system (Thermo Shot F30, NEC Avio Infrared Technologies Co.). Two males (subjects A in his 50s and B in his 20s) participated in the experiment. Using Skype, they held a conversation for approximately 100 seconds. The videos saved by VodBurner were transformed into AVI files, and WAV files were saved by Tapur. The AVI files were used for measuring the facial expression intensity. The WAV files were used for judgment of the utterance of both subjects A and B.

2. Results and discussion

Facial expression intensity change of Subject A during the conversation with Subject B were recorded (Fig. 4), and the subjects' utterances were classified (Fig. 5). In Fig. 5, images of the face and images of the mouth and the jaw show the characteristic timing positions for the facial expression intensity.

Subject A expressed constant changes of facial expression intensity during utterances (Fig. 5 (a), (b)), while he expressed drastic changes of facial expression intensity during no utterances (Fig. 5 (c), (d)), which corresponded to his smile. Therefore, the big changes in facial expression intensity shown in Fig. 4 are mainly expressed in the natural smile of Subject A.

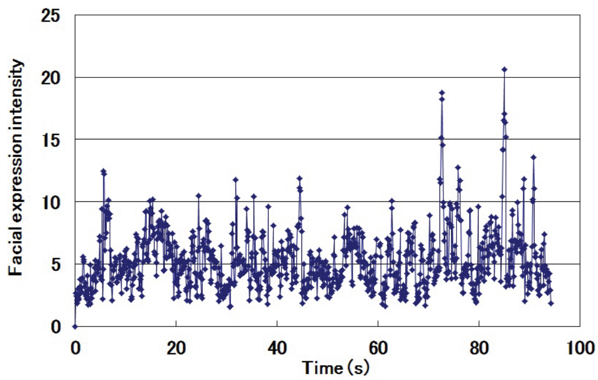


Fig. 4. Facial expression intensity change of Subject A during the conversation with Subject B

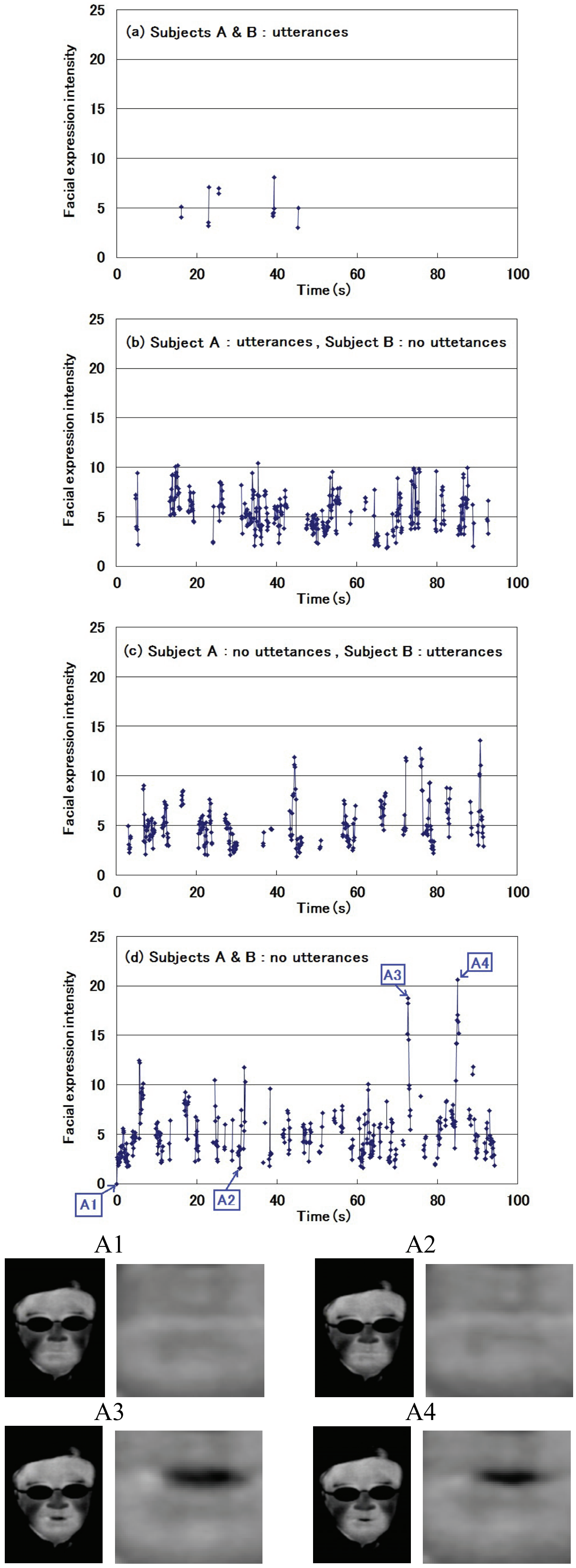


Fig. 5. Facial expression intensity changes (upper graphs), images of face and images of area of mouth and jaw (lower) of subject A during the conversation between subjects A and B

Several images of the face and images of the mouth and jaw at the characteristic timing points show that the proposed method can quantitatively express the facial expression (Fig. 5 lower images).

Based on the proposed method, we intend to develop a method for estimating the emotions or mental state of a patient.

IV. CONCLUSION

We proposed a method for analyzing the facial expressions of a person while speaking with a video phone system (Skype). The recorded video is analyzed by thermal image processing and the newly proposed feature vector of facial expression, which is extracted in the mouth area by applying 2D-DCT. The facial expression intensity, defined as the norm of the difference vector between the feature vector of the neutral facial expression and that of the observed expression, can be used to analyze the change of facial expression. The judgment of utterance is performed by using the intensity of the sound wave. The experimental results show the usefulness of the proposed method.

Acknowledgment

The present study was partially supported by KAKENHI (22300077).

REFERENCES

[1] Skype. http://www.skype.com/ Accessed 5 November 2013

[2] Yoshitomi Y, Kimura S, Hira E, et al (1996) Facial expression recognition using infrared rays image processing. In: Proceedings of the Annual Convention IPS Japan, Osaka, Japan, Sep 4-6, 1996, 2:339-340

[3] Yoshitomi Y, Miyawaki N, Tomita S, et al (1997) Facial expression recognition using thermal image processing and neural network. In: Proceedings of 6th IEEE International Workshop on Robot and Human Communication, Sendai, Japan, Sep 29-Oct 1, 1997, pp 380-385

[4] Yoshitomi Y, Kim SIll, Kawano T, et al (2000) Effect of sensor fusion for recognition of emotional states using voice, face image and thermal image of face. In: Proceedings of 6th IEEE International Workshop on Robot and Human Interactive Communication, Osaka, Japan, Sep 27-29, 2000, pp 178-183

[5] Ikezoe F, Ko R, Tanijiri T, et al (2004) Facial expression recognition for speaker using thermal image processing (in Japanese). Trans Human Interface Soc 6(1):19-27

[6] Nakano M, Ikezoe F, Tabuse M, et al (2009) A study on the efficient facial expression using thermal face image in speaking and the influence of individual variations on its performance (in Japanese). J IEEJ 38(2):156-163

[7] Koda Y, Yoshitomi Y, Nakano M, et al (2009) Facial expression recognition for a speaker of a phoneme of vowel using thermal image processing and a speech recognition system. In: Proceedings of 18th IEEE International Symposium on Robot and Human Interactive Communication, Toyama, Japan, Sep 29-Oct 1, 2009, pp 955-960

[8] Fujimura T, Yoshitomi Y, Asada T, et al (2011) Facial expression recognition of a speaker using front-view face judgment, vowel judgment, and thermal image processing. J Artif Life and Robotics 16(3):411-417

[9] Yoshitomi Y, Asada T, Shimada K, et al (2011) Facial expression recognition of a speaker using vowel judgment and thermal image processing. J Artif Life and Robotics 16(3):318–323

[10]Nakanishi Y, Yoshitomi Y, Asada T, et al (2013) Robust facial expression recognition of a speaker using thermal image processing and updating of fundamental training-data. J Artif Life and Robotics, 17(3):342-349

[11]Yoshitomi Y, Tabuse M, Asada T (2012) Facial expression recognition using thermal image processing. In: Carvalho VH (Ed) Image processing: methods, applications and challenges. Nova Science Publisher, New York, pp. 57-85

[12]Yoshitomi Y, Tabuse M, Asada T, (2011) Vowel judgment for facial expression recognition of a speaker. In: Ipšić I (Ed) Speech Technologies. InTech, Rijeka, pp. 405-424

[13]Nakanishi Y, Yoshitomi Y, Asada T, et al (2014) Facial expression recognition of a speaker using thermal image processing and reject criteria in feature vector space. J Artif Life and Robotics, in press

[14]VodBurner. http://www.vodburner.com/ Accessed 1 December 2013

[15]Tapur. http://www.tapur.com/jp/ Accessed 6 December 2013

[16]Ikezoe F, Nakano M, Yoshitomi Y, et al (2005) Facial expression recognition using thermal face image automatically acquired in speaking (in Japanese). In: Proceedings of Human Interface Symposium 2005, Fujisawa, Kanagawa, Japan, Se p.15-17, 2005, pp 7-12

[17]Nakano M, Yoshitomi Y, and Tabuse M (2006) Efficient facial expression recognition using thermal face image in speaking and its application to analysis of individual variations (in Japanese). In: Proceedings of Human Interface Symposium 2006, Kurashiki, Okayama, Japan, Sep.25-28, 2006, pp 1151-1156

Facial Expression Recognition Using Thermal Image Processing and Efficient Preparation of Training-data

Y. Nakanishi[1], Y. Yoshitomi[2], T. Asada[2], and M. Tabuse[2]

1: ITOKI CORPORATION, 1-4-12 Imafuku-higashi, Joto-ku, Osaka 536-0002, 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@kpu.ac.jp, t_asada@mei.kpu.ac.jp, tabuse@kpu.ac.jp

Abstract: In our previously developed method for facial expression recognition, we prepared training data for a pair of the first and last vowels pronounced by the speaker. The number of pairs is 25 for the Japanese language. We investigated the influence of training data on the facial expression accuracy using the training data of "taro," whose first and last vowels are /a/ and /o/, for the three intentional facial expressions of "angry," "sad," and "surprised," and the training data of 25 pairs of vowels for the two intentional facial expressions of "happy" and "neutral." Using the proposed method, the facial expressions of one subject were discriminable with accuracies of 100%, 70.0%, and 47.2% for "taro," "koji" (the first and last vowels of which are /o/ and /i/), and "tsubasa" (the first and last vowels of which are /u/ and /a/), respectively, for the three facial expressions of "happy," "neutral," and "other" when one of the five intentional facial expressions of "angry," "happy," "neutral," "sad," and "surprised" was exhibited. On the other hand, the facial expressions of one subject were discriminable with accuracies of 90.6% and 85.7% for "koji" and "tsubasa," respectively, when the training data were generated using "koji" and "tsubasa" as respective pronunciations.
Keywords: Facial expression recognition, Speech recognition, Vowel judgment, Thermal image processing, Efficient gathering of training data.

I. INTRODUCTION

The present study investigates the first stage of the development of a robot that has the ability to visually detect human feelings or mental states. Although the mechanism for recognizing facial expressions has received considerable attention in the field of computer vision research, it still falls far short of human capability, especially from the viewpoint of robustness under widely varying lighting conditions. One reason for this lack of robustness is that nuances of shade, reflection, and local darkness influence the accuracy of facial expression recognition through the inevitable change of gray levels. In order to avoid this problem and develop a method for facial expression recognition that is applicable under widely varying lighting conditions, we used images produced by infrared rays (IR), which reveal the thermal distribution of the face [1]-[12]. We adopted an utterance as the key to expressing human feelings or mental states because humans tend to express feelings vocally [3]-[12]. Although several studies on facial expression recognition using thermal image processing have been reported (see references [1]-[17]), only our research [3]-[12] has focused on the speaker. The present study also focuses on the speaker. In addition, we added a judgment function of a front-view face to the proposed method for facial expression recognition [7]. We developed a method for efficiently updating the training data because frequent updates are time-consuming [9]. Through experiments, we concluded that updating the training data corresponding to the facial expressions of "happy" and "neutral" is practical. These two facial expressions are not only very common in our daily lives, but are also easier to express than other facial expressions. Furthermore, the classifications of "neutral," "happy," and "other" are efficient for facial expression recognition under the condition that updating the training data of facial expressions is not performed frequently.

In our previously developed method for facial expression recognition of a speaker, we prepared the training data for a pair of the first and last vowels pronounced by the speaker [8]. The number of pairs is 25 for the Japanese language. It is time-consuming and difficult for a subject to express facial expressions for each pair of vowels.

In the present study, we investigate the influence of training data on facial expression accuracy using the training data of "taro," the first and last vowels of which are /a/ and /o/, for the three intentional facial expressions of "angry," "sad," and "surprised," and the training data of 25 pairs of vowels for the two intentional facial expressions of "happy" and "neutral."

II. IMAGE ACQUISITION

The principle behind thermal image generation is the Stefan-Boltzmann law, which is expressed as $W = \varepsilon\sigma T^4$, where ε is the emissivity, σ is the Stefan–Boltzmann constant (= 5.6705×10^{-12} W/cm^{2}K^4), and T

is the temperature (K). For human skin, ε is estimated to range from 0.98 to 0.99 [18], [19]. In the present study, the approximate value of 1 was used as ε for human skin because the value of ε for almost all substances is lower than that of human skin [18]. Consequently, the human face region is easily extracted from an image using the value of 1 for ε [1]-[12]. In principle, IR temperature measurements do not depend on skin color [19], darkness, or lighting condition, and so the face region and its characteristics are easily extracted from a thermal image.

III. METHOD FOR FACIAL EXPRESSION RECOGNITION

Figure 1 shows a flowchart of the proposed method, which consists of two modules. The first is a module for speech recognition and dynamic image analysis, and the second is a module for learning and recognition. In the module for learning and recognition, we embedded the module for front-view face judgment [7]. The proposed method is described in detail in our book [10].

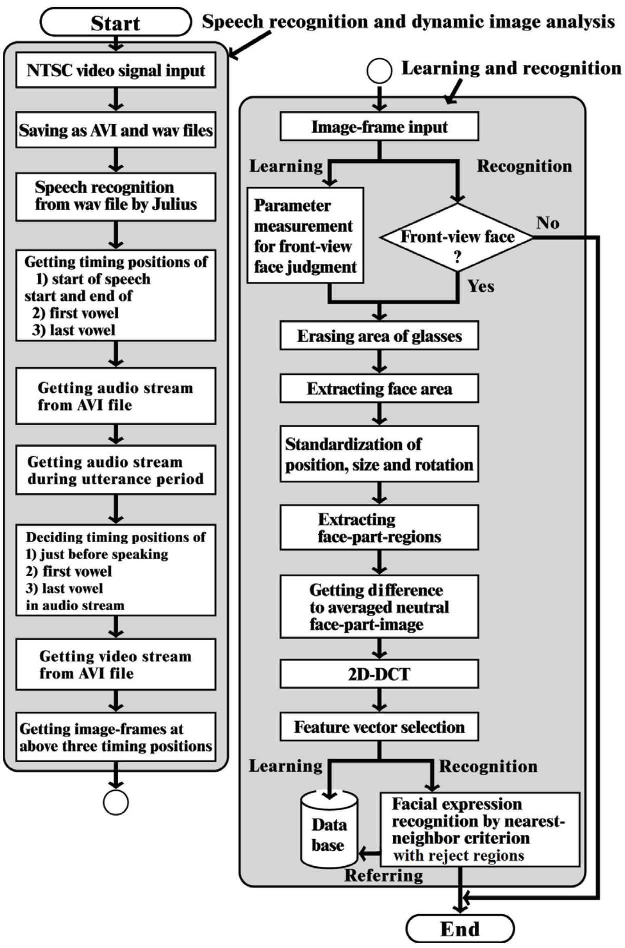


Fig. 1. Flowchart of the proposed method [12]

1. Speech recognition and dynamic image analysis

We use a speech recognition system called Julius [20] to obtain the timing positions of the start of speech and the first and last vowels in a WAV file [6]-[12]. Figure 2 shows an example of the waveform of the Japanese name "Taro." The timing position of the start of speech and the timing ranges of the first vowel (/a/) and the last vowel (/o/) are decided by Julius. Using the timing position of the start of speech and the timing ranges of the first and last vowels obtained from the WAV file, three image frames are extracted from an AVI file at the three timing positions. For the timing position just before speaking, we use the timing position of 84 ms before the start of speech, as determined in our previously study [5]. For the timing position of the first vowel, we use the position at which the absolute value of the amplitude of the waveform is the maximum while speaking the vowel. For the timing position of the last vowel, we apply the procedure used for the first vowel.

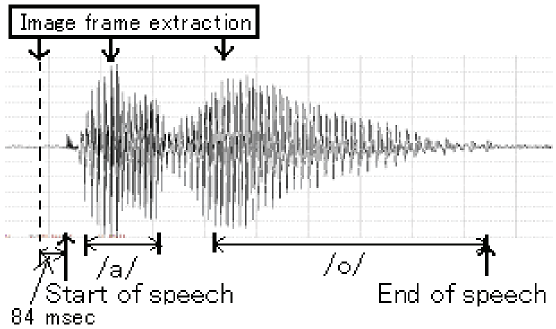


Fig. 2. Speech waveform of "taro" and timing positions for image frame extraction [6]

2. Learning and recognition

For static images obtained from the extracted image frames, the process of erasing the area of the glasses, extracting the face area, and standardizing the position, size, and rotation of the face are performed according to the method described in our previous study [5], [10]. Next, we generate difference images between the averaged neutral face image and the target face image in the extracted face areas in order to perform a 2D discrete cosine transform (2D-DCT). The feature vector is generated from the 2D-DCT coefficients according to a heuristic rule [4], [5]. The facial expression is recognized by the nearest-neighbor criterion in the feature vector space with the rejection domain using the training data just before speaking and that while speaking the phonemes of the first and last vowels [12].

VI. EXPERIMENTS

1. Condition

The thermal image produced by the thermal video system (Nippon Avionics TVS-700) and the sound captured from an Electret condenser microphone (Sony ECM-23F5), as amplified by a mixer (Audio-Technica AT-PMX5P), were transformed into a digital signal by an A/D converter (Thomson Canopus ADVC-300) and were input into a computer (DELL Optiplex 780, CPU: Intel Core 2 Duo E8400 3.00 GHz, main memory: 3.21 GB, and OS: Windows 7 Professional (Microsoft) with an IEEE1394 interface board (I·O Data Device 1394-PCI3/DV6). We used Visual C++ 6.0 (Microsoft) as the programming language. In order to generate a thermal image, we set the condition such that the thermal image had 256 gray levels for the detected temperature range. The temperature range for generating a thermal image was decided so as to easily extract the face area on the image. We saved the visual and audio information in the computer as a Type 2 DV-AVI file, in which the video frame had a spatial resolution of 720×480 pixels and 8-bit gray levels, and the sound was saved in a stereo PCM format, 48 kHz and 16-bit levels.

Subject A, a male with glasses, performed in alphabetic order each of the intentional facial expressions of “angry,” “happy,” “neutral,” “sad,” and “surprised,” while speaking the semantically neutral utterance of each of the Japanese first names listed in Table 1.

Table 1 Japanese first names used in the experiment [8]

		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

In the experiment, Subject A intentionally maintained a front-view in the AVI files, which were saved as both training and test data. We assembled 20 samples as training data and 10 or less samples as test data, in which all facial expressions of the subject were judged as front-view faces by our reported method [7]. The number of test data was decided as a result of the front-view face judgment. From one sample, we obtained three images at the timing positions of just before speaking and while speaking the phonemes of the first and last vowels.

In the present study, we have investigated the influence of training data on the facial expression recognition accuracy. The method for facial expression recognition using the training data of “taro” (the first and last vowels of which are /a/ and /o/), for the three intentional facial expressions of “angry,” “sad,” and “surprised,” and the training data of 25 pairs of vowels for the two intentional facial expressions of “happy” and “neutral” were selected. The method is hereinafter referred to as the efficient method. Then, as pronunciations, “taro,” “koji” (the first and last vowels of which are /o/ and /i/), and “tsubasa” (the first and last vowels of which are /u/ and /a/), were selected when applying the efficient method.

For comparison with the efficient method, the method for facial expression recognition using the training data of the same pronunciations as those of test data was also applied to the data for “koji” and “tsubasa.” This method is hereinafter referred to as the reference method. Figures 3 and 4 show examples of thermal images of the subject used for training by the efficient method and in the test, respectively.

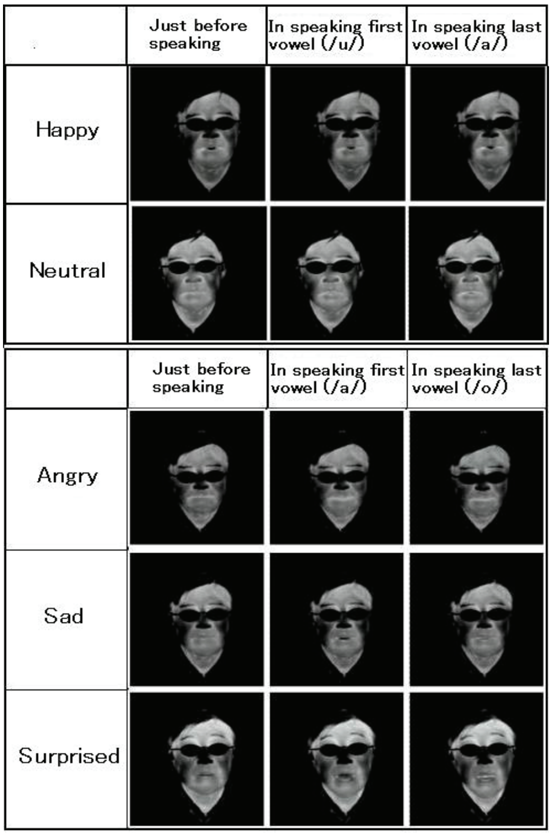


Fig. 3 Examples of thermal training images for speaking “tsubasa” in the efficient method

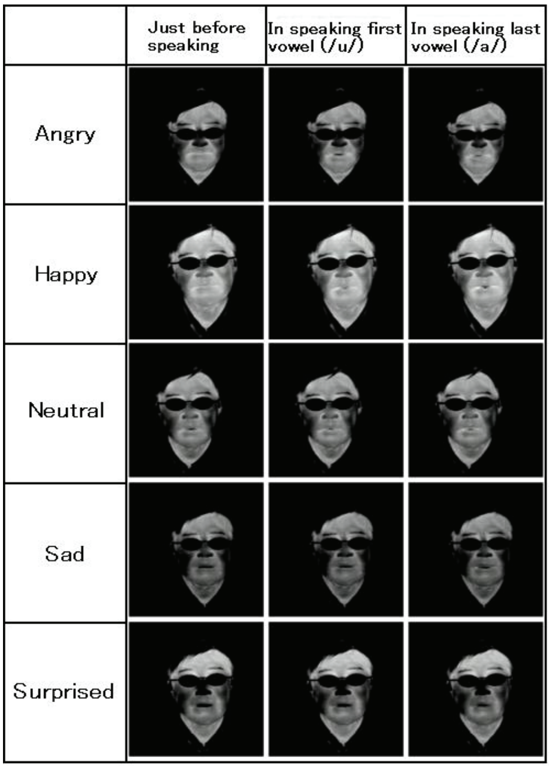


Fig. 4 Examples of thermal test images while speaking "tsubasa"

2. Results and discussion

Tables 2 and 3 show the values of the recognition accuracy of the facial expressions obtained using the efficient method and the reference method, respectively.

Using the efficient method, the average facial expression accuracies were 100%, 70.0%, and 47.2%, respectively, for "taro," "koji," and "tsubasa," for the three facial expressions of "happy," "neutral," and "other" when the speaker exhibited one of the five intentional facial expressions of "angry," "happy," "neutral," "sad," and "surprised." On the other hand, using the reference method, the average facial expression recognition accuracies were 90.6% and 85.7% for "koji" and "tsubasa," respectively. Accordingly, the influence of training data on the facial expression recognition accuracy might not be small.

V. CONCLUSION

We have investigated the influence of training data on facial expression accuracy using training data for "taro," the first and last vowels of which are /a/ and /o/, respectively, for the three intentional facial expressions of "angry," "sad," and "surprised," and the training data of 25 pairs of vowels for the two intentional facial expressions of "happy" and "neutral." Using the efficient

Table 2 Recognition accuracies for "taro," "koji," and "tsubasa" as obtained when the efficient method was used

taro		Input facial expression					
		Correct speech recognition			Poor speech recognition		
		Happy	Neutral	Others	Happy	Neutral	Others
Output	Happy	9/9			1/1		
	Neutral		10/10				
	Others			9/9			14/14
Rejected							7
Accuracy		28/28			15/15		
Total accuracy		43/43					

koji		Input facial expression					
		Correct speech recognition			Poor speech recognition		
		Happy	Neutral	Others	Happy	Neutral	Others
Output	Happy	0/3			1/4		3/14
	Neutral		9/9				
	Others	3/3			3/4		11/14
Rejected				12	3		1
Accuracy		9/12			12/18		
Total accuracy		21/30					

tsubasa		Input facial expression					
		Correct speech recognition			Poor speech recognition		
		Happy	Neutral	Others	Happy	Neutral	Others
Output	Happy	0/9		7/15			
	Neutral	9/9	6/9				
	Others		3/9	8/15			3/3
Rejected			1	8			4
Accuracy		14/33			3/3		
Total accuracy		17/36					

Table 3 Recognition accuracies for "koji" and "tsubasa" as obtained when the reference method was used

koji		Input facial expression					
		Correct speech recognition			Poor speech recognition		
		Happy	Neutral	Others	Happy	Neutral	Others
Output	Happy	2/3			4/5		1/7
	Neutral		9/9				
	Others	1/3		8/8	1/5		6/7
Rejected		1	1	2			6
Accuracy		19/20			10/12		
Total accuracy		29/32					

tsubasa		Input facial expression					
		Correct speech recognition			Poor speech recognition		
		Happy	Neutral	Others	Happy	Neutral	Others
Output	Happy	0/4		1/18			
	Neutral	4/4	9/9				
	Others			17/18			4/4
Rejected			1	5			3
Accuracy		26/31			4/4		
Total accuracy		30/35					

method, the facial expressions of one subject were discriminable with 100%, 70.0%, and 47.2% accuracy for "taro," "koji" (the first and last vowels of which are /o/ and /i/), and "tsubasa" (the first and last vowels of which are /u/ and /a/), respectively, for the three facial expressions of "happy," "neutral," and "other" when the speaker exhibited one of the five intentional facial expressions of "angry," "happy," "neutral," "sad," and "surprised." On the other hand, the facial expressions of one subject were discriminable with accuracies of 90.6% and 85.7% for "koji" and "tsubasa," respectively, when the reference method was used. In order to develop a practical method for facial expression recognition, we intend to investigate a method for reducing the influence of training data on facial expression accuracy.

Acknowledgments

We would like to thank Ms. C. Shiraishi of Kyoto Prefectural University for her valuable support during the data analysis. The present study was supported by KAKENHI (22300077).

REFERENCES

[1] Yoshitomi Y, Kimura S, Hira E, et al (1996) Facial expression recognition using infrared rays image processing. In: Proceedings of the Annual Convention IPS Japan, Osaka, Japan, Sep 4-6, 1996, 2:339-340

[2] Yoshitomi Y, Miyawaki N, Tomita S, et al (1997) Facial expression recognition using thermal image processing and neural network. In: Proceedings of 6th IEEE International Workshop on Robot and Human Communication, Sendai, Japan, Sep 29-Oct 1, 1997, pp 380-385

[3] Yoshitomi Y, Kim SIll, Kawano T, et al (2000) Effect of sensor fusion for recognition of emotional states using voice, face image and thermal image of face. In: Proceedings of 6th IEEE International Workshop on Robot and Human Interactive Communication, Osaka, Japan, Sep 27-29, 2000, pp 178-183

[4] Ikezoe F, Ko R, Tanijiri T, et al (2004) Facial expression recognition for speaker using thermal image processing (in Japanese). Trans Human Interface Soc 6(1):19-27

[5] Nakano M, Ikezoe F, Tabuse M, et al (2009) A study on the efficient facial expression using thermal face image in speaking and the influence of individual variations on its performance (in Japanese). J IEEJ 38(2):156-163

[6] Koda Y, Yoshitomi Y, Nakano M, et al (2009) Facial expression recognition for a speaker of a phoneme of vowel using thermal image processing and a speech recognition system. In: Proceedings of 18th IEEE International Symposium on Robot and Human Interactive Communication, Toyama, Japan, Sep 29-Oct 1, 2009, pp 955-960

[7] Fujimura T, Yoshitomi Y, Asada T, et al (2011) Facial expression recognition of a speaker using front-view face judgment, vowel judgment, and thermal image processing. J Artif Life and Robotics 16(3):411-417

[8] Yoshitomi Y, Asada T, Shimada K, et al (2011) Facial expression recognition of a speaker using vowel judgment and thermal image processing. J Artif Life and Robotics 16(3):318–323

[9] Nakanishi Y, Yoshitomi Y, Asada T, et al (2013) Robust facial expression recognition of a speaker using thermal image processing and updating of fundamental training-data. J Artif Life and Robotics, 17(3):342-349

[10]Yoshitomi Y, Tabuse M, Asada T (2012) Facial expression recognition using thermal image processing. In: Carvalho VH (Ed) Image processing: methods, applications and challenges. Nova Science Publisher, New York, pp. 57-85

[11]Yoshitomi Y, Tabuse M, Asada T, (2011) Vowel judgment for facial expression recognition of a speaker. In: Ipšić I (Ed) Speech Technologies. InTech, Rijeka, pp. 405-424

[12]Nakanishi Y, Yoshitomi Y, Asada T, et al (2014) Facial expression recognition of a speaker using thermal image processing and reject criteria in feature vector space. J Artif Life and Robotics, to appear

[13]Socolinsky D A, Selinger A (2002) A comparative analysis of face recognition performance with visible and thermal infrared imagery. In: Kasturi R, Laurendeau D, Suen C, et al (Eds.) Proceedings of 16th International Conference on Pattern Recognition, Quebec City, Canada ,Aug 11-15, 2002, 4, pp 217-222

[14]Khan M M, Ward R D, Ingleby M (2004) Automated classification and recognition of facial expressions using infrared thermal imaging. In: Proceedings of 2004 IEEE Conference on Cybernetics and Intelligent Systems, Singapore, Dec 1-3, 2004, 1, pp 202-206

[15]Khan M M, Ingleby M, Ward R D (2006) Automated facial expression classification and affect interpretation using infrared measurement of facial skin temperature variations. ACM Trans on Autonomous and Adaptive Systems, 1(1):91-113

[16]Jiang G, Song X, Zheng F, et al (2006) Facial expression recognition using thermal image. In: Proceedings of 27th Annual International Conference of the Engineering in Medicine and Biology Society, Shanghai , China, Jan 17-18, 2006, pp 631-633

[17]Hernández B, Olague G, Hammoud R, et al (2007) Visual learning of texture descriptors for facial expression recognition in thermal imagery. Computer Vis and Image Underst 16 (2-3):258-269

[18]Kuno H (1994) Infrared rays engineering (in Japanese). Tokyo, IEICE, p 22

[19]Kuno H (1994), Infrared rays engineering (in Japanese). Tokyo, IEICE, p 45

[20]Kawahara T, et al (2010) Open-Source Large Vocabulary CSR Engine Julius. Julius rev.4.1.5.1. http://julius.sourceforge.jp/ Accessed 12 March 2013

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Music Recommendation System through Internet for Improving Recognition Ability Using Collaborative Filtering and Impression Words

Y. Yoshitomi[1], T. Asada[1], R. Kato[1], Y. Yoshimitsu[1], M. Tabuse[1], N. Kuwahara[2], and J. Narumoto[3]
1: Graduate School of Life and Environmental Sciences Kyoto Prefectural University,
1-5 Nakaragi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan
yoshitomi@kpu.ac.jp, {t_asada, r_kato, y_yoshimitsu}@mei.kpu.ac.jp, tabuse@kpu.ac.jp,
2: Graduate School of Science and Technology, Kyoto Institute of Technology,
Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan, nkuwahar@kit.ac.jp
3: Graduate School of Medical Science, Kyoto Prefectural University of Medicine,
Kajii-cho, Kawaramachi-Hirokoji, Kamigyo-ku, Kyoto 602-8566, Japan, jnaru@koto.kpu-m.ac.jp

Abstract: Music therapy for improving recognition ability may be more effective if music that is liked by an individual is adopted. In our previously reported system, the recommendation process using collaborative filtering was terminated when no users in the reference list have the same preference for recommended music as that of a new user. Based on the scores for impression words, the second recommendation process finds the most similar music to the successfully recommended music among music that has not yet been recommended. In the present study, based on our previously reported system, we propose a system for music recommendation through the Internet using a videophone system (Skype). The effectiveness of the proposed system is verified. The recommendation accuracy of the proposed system was 90.6% for 10 trials by five elderly subjects. The average number of recommended pieces of music for the subjects by the proposed system was 14.9 per trial. In the future, we intend to increase the number of user evaluation scores in the database and apply the proposed system to older individuals and/or to individuals with cognitive impairment.
Keywords: Collaborative filtering, Music recommendation, Music therapy, Impression word, Internet, Skype.

I. INTRODUCTION

In Japan, the average age of the population has been increasing, and this trend is expected to continue. This trend is more remarkable in rural areas. Recently, music therapy has been used to improve the recognition ability of people, particularly older people. Music therapy may be more effective if music that is liked by an individual is adopted. We have been developing a music recommendation system aimed at improving recognition ability [1]-[2].

In the present study, based on our previously reported system [2], we propose a system for music recommendation through the Internet using a videophone system (Skype)[3], and verify the effectiveness of the proposed system.

II. MUSIC RECOMMENDATION METHOD USING IMPRESSION WORDS

We use ten pairs of impression words (Table 1) [4]. As an example, we show a user scores for one pair of impression words (quiet - busy), and the user scores the word pairs according to seven levels, which are then transformed to three levels, as shown in Table 2 [2]. A music database was created in which all songs in the database were assigned scores, i ($-3 \leq i \leq 3$), for each pair of impression words evaluated by the participants. Fig. 1 shows a flowchart of the music recommendation based on impression words. When music that has not been recommended to a user has the same scores, with the exception of “0”, as that of at least one recommended piece of music that has been evaluated highly by the user on the three-level scale for at least five impression words, the music is treated as having been positively evaluated by the user. In contrast, when music that has not been recommended to the user has the same scores, with the exception of “0”, as that of another piece of music that has been recommended to the user and has been negatively evaluated by the user on the three-level scale for at least seven impression words, the music is treated as having been negatively evaluated by the user. In Fig. 1, when none of the not recommended music receives a positive evaluation by the user, another recommendation is performed by using the subjective estimations of all users whose subjective estimations are stored in the database.

It is expressed by “with highest similarity” in Fig. 1 that the music has the highest proportion of the same three-level scores except “0” as that of other music recommended to the user and given a positive evaluation by the user among the music not yet recommended to the user. In Fig. 1, the “set of music with a similarity to the recommended music based on impression words” (MSRIW) is decided by using at least seven pairs of impression words in the case of a negative evaluation.

Table 1. Pairs of impression words [4]

quiet - busy
bracing - heavy
easy - uneasy
cheerful - gloomy
refreshing - depressing
happy - sad
comforting - harmful
calm - elevating
clean - dirty
magnificent - superficial

Table 2. Scores for pairs of impression words, quiet - busy [1]

Score	Three-level score	impression
3	-1	very busy
2		busy
1	0	slightly busy
0		neutral
-1		slightly quiet
-2	1	quiet
-3		very quiet

III. PROPOSED SYSTEM

Fig. 2 shows an overview of the proposed system. Fig. 3 shows a flowchart of our previously reported system, which combines collaborative filtering and our music recommendation process based on impression words.

In the proposed system, the recommendation process using collaborative filtering is terminated when the number of users is zero in the reference list of users showing exactly the same evaluation for the recommended music as that of the user up to that moment. Then, the recommendation process performs by finding out the most similar music, from the viewpoints of three-level scores except "0" on impression words, to that successfully recommended among music not yet recommended. The proposed system recommends music stored in the database to user u, as shown in Fig. 3. The recommendation process using the proposed system is terminated when the number of recommended songs reaches the upper limit K, which has been decided previously. Just before finishing the recommendation process, the database of users with subjective estimations of the music is updated by adding the subjective estimations of the user for whom the proposed system recommends music [2].

In the flowchart shown in Fig. 3, the estimation of user u for song m_R is set to 1 when the score of user u for song m_R is 4 or 5 ("slightly like" or "favorite"), and is set as 0 when the score is 1 to 3 ("dislike," "slightly dislike," or "neutral"). We used ten pairs of impression words (Table 1) [4]. As an example, we showed the user scores for one pair of impression words (quiet - busy). Users scored the word pairs according to seven levels, which are then transformed to three levels, as shown in Table 2 [1].

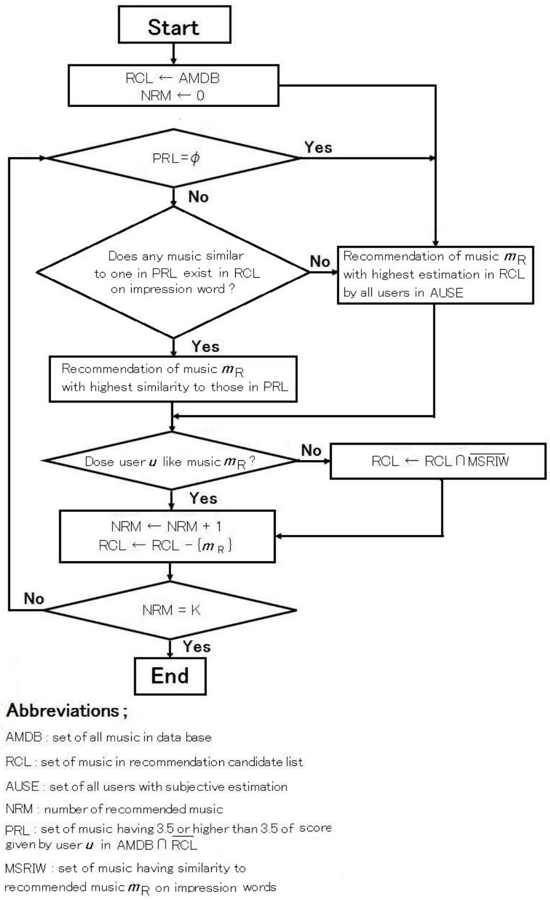


Fig. 1. Flowchart of music recommendation by using impression words [1]

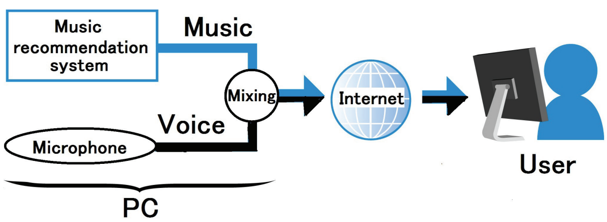


Fig.2. Overview of the proposed system

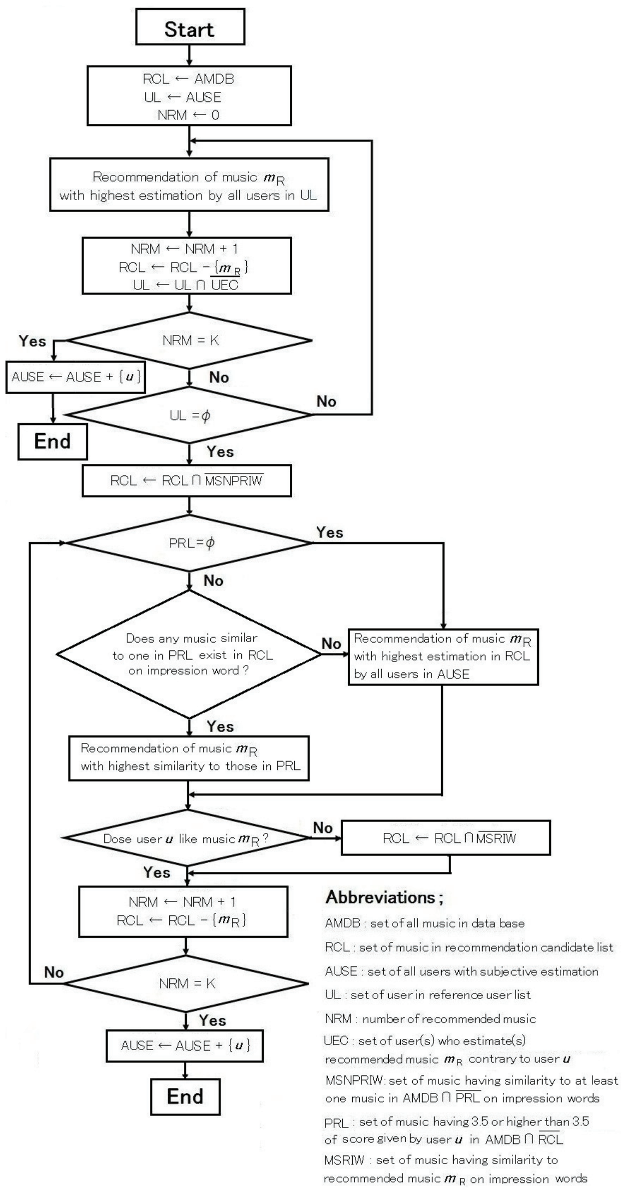


Fig. 3. Flowchart of music recommendation used in the proposed system [2]

IV. PERFORMANCE EVALUATION

1. Conditions

Since older people tend to prefer children's songs [5], we selected a CD described as an anthology of older songs enjoyed by older people with dementia [6] and selected 52 songs on the CD that were also included in a music textbook database for elementary schools [7]. The Japanese titles of the 52 songs are listed in Table 3. In order to evaluate the music recommendation method, all 52 of the selected songs in the database were assigned scores, s ($1 \leq s \leq 5$), by 12 subjects of different ages (13-19 years: one subject, 20-29 years: six subjects,

Table 3. List of songs used in the present study

Harugakita	Teruterubozu	Mushinokoe	Hiraitahiraita
Takibi	Natsuwakinu	Soranbushi	Zuizuizukkorobashi
Akaikutsu	Tanabata	Harunoogawa	Antagatadokosa
Usagi	Natsunoomoide	Muramatsuri	Usagitokame
Sakura	Warewauminoko	Kutsuganaru	Umi (Chinese character)
Hana	Umi (Hiragana)	Yuuhi	Urashimatarou
Koinobori	Kintarou	Kakashi	Momotarou
Seikurabe	Akatonbo	Furusato	Hanasakajiji
Fujinoyama	Yuuyakekoyake	Yuki	Ushiwakamaru
Chatsumi	Koujyounotsuki	Oshougatsu	Ureshiihinamatsuri
Kisha	Oborozukiyo	Katatsumuri	Nanatsunoko
Amefuri	Tsuki	Kagome	Ichigatsutsuitachi
Ame	Momiji	Touryanse	Dongurikorokoro

50-59 years: five subjects). We used 15 as the value for K in the evaluations of the proposed method. In addition, all 52 of the selected songs in the database were assigned scores, i ($-3 \leq i \leq 3$), for each pair of impression words by five subjects of different ages (20-29 years: three subjects, 40-49 years: one subject, 50-59 years: one subject). The subject in his fifties also assigned the scores, s. The average score i obtained from the five subjects for each pair of impression words was used as the score i for the performance evaluation. The 15 songs having three-level scores other than "0" for one impression word at most were not recommended in the recommendation process based on impression words because these songs did not have distinct characteristics from the viewpoint of impression words.

Updating of the AUSE, which is the set of all users with subjective estimations in Fig. 3, was effective in improving the accuracy of the recommendation and increasing the number of recommended songs [2]. Therefore, in order to improve the performance of the proposed system, we used a face-to-face system that updated the AUSE for 44 users before this experiment was performed.

For programming, we used Visual C++ 6.0 (Microsoft) on a PC (Dell Latitude E6599, CPU: Intel Core 2 Duo P8700 2.54 GHz, main memory: 4.00 GB, and OS: Windows 7, Microsoft) for the experiment. We used a videophone system (Skype) to send music to users over the Internet. The proposed system was set up at Kyoto Prefectural University in Kyoto City located in the southern part of Kyoto Prefecture in Japan. The participants for this experiment resided in Kyotango City in the northern part of Kyoto Prefecture, Japan. Using Skype over the Internet, 10 older users, referred to as users 1 through 10, of different ages (70-79 years: three subjects, 80-89 years: six subjects, 90-99 years: one subject) participated in the experiment on user-by-user updating for the AUSE. All of the users were female. Users 3, 4, and 8 were the same person and users 5, 6, 7, and 10 were also the same person.

2. Results and discussion

Fig. 4 shows the performance of the proposed system, with the updating of the AUSE for users 1 through 10. The mean value of the number of recommended songs for users 1 through 10 was 14.9 per trial. The mean value of the recommendation accuracy for 10 trials by the users 1 through 10 was 90.6%.

The impression words used in the present study were selected based on the linguistic method [4]. The questionnaire was completed by 100 subjects for 80 pieces of classical music and the impression words used in the present study were demonstrated to have very low correlation with each other [4]. For details on the selection of impression words, see Reference 4. The impression words used in the present study are not related to the acoustic features of music, such as tempo, tonality, rhythm, and harmony.

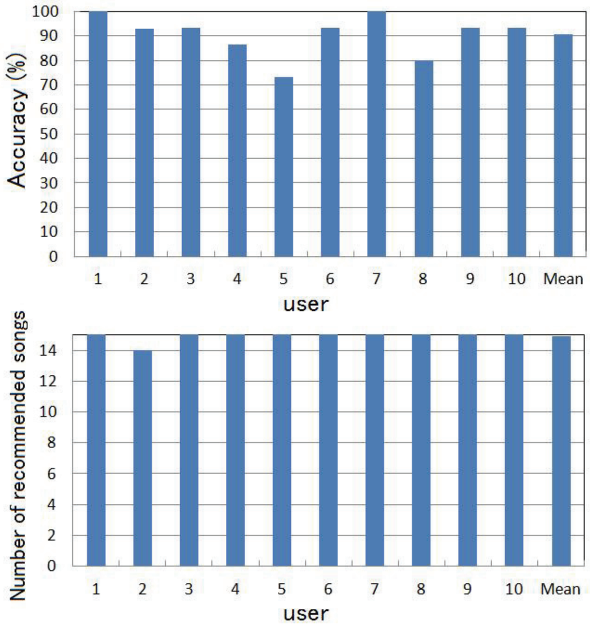


Fig. 4. Performance of the proposed system. Upper: recommendation accuracy, Lower: number of recommended songs

V. CONCLUSION

We have proposed a system for music recommendation over the Internet using a videophone system (Skype). The effect of this system was verified in the present study. The recommendation accuracy of the proposed system was 90.6% for 10 trials by five elderly subjects. The average number of recommended pieces of music for subjects by the proposed system was 14.9 per trial. In the future, we intend to increase the number of user evaluation scores in the database and apply the proposed system to older individuals and/or to individuals with cognitive impairment.

Acknowledgment

We would like to thank all of the subjects who participated in the experiments. The present study was supported in part by SCOPE (122307003) of the Ministry of Internal Affairs of Communications of Japan.

REFERENCES

[1]Koro C, Yoshitomi Y, Asada T, Yoshizaki S (2012), Music recommendation aimed at improving recognition ability using collaborative filtering and impression words. In:Sugisaka M (Ed), Proceedings of 17th International Symposium on Artificial Life and Robotics (AROB17th), Beppu, Oita Japan, Jan 19-21, 2012, pp 222-225

[2]Yoshizaki S, Yoshitomi Y, Koro C, Asada T (2013), Music recommendation hybrid system for improving recognition ability using collaborative filtering and impression words. J Artif Life and Robotics, in press.

[3]Skype. http://www.skype.com/ Accessed 5 November 2013

[4]Kumamoto T, Ohta K (2002) Design of scales to represent user's impressions of a music piece for use in a music-retrieval system (in Japanese). IPSJ SIG Notes, 2001-NL-147(6):35-40

[5]Takahashi T (1997) Research report on songs familiar to people advanced in years (in Japanese). J Japanese music therapy associate 15(1):68-75

[6]Akahoshi T (2009) Good old anthology enjoyable for people advanced in years and troubled with dementia (in Japanese). Kirara shobo, Tokyo

[7]Music textbook database for elementary school by Kanagawa prefectural education center (in Japanese). http://kjd.edu-ctr.pref.kanagawa.jp/daizai_music/ Accessed 11 March 2013

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Kansei Engineering based Evaluation for an e-Learning System with IP-based Network Design and Animation

Yoshiro Imai[1], Chiaki Kawanishi[12], Tetsuo Hattori[1]

[1]Graduate School of Engineering, Kagawa University, 2217-20 Hayashi-cho, Takamatsu 761-0396 Japan
[2]Miss Kawanishi was a student of Graduate School of Engineering(Kagawa University) from April of 2011 to March of 2013.
(Tel : +81-87-864-2244)
E-mail: [1]{imai, hattori}@eng.kagawa-u.ac.jp

Abstract: An e-Learning system has been developed, which is designed for network education with facilities of specification of IP-based network topology and demonstration of packet-transferring animation. This system can execute on the major browsers by means of accessing a specific Web server. It is frequently used in some lectures of university from beginners of network to the students of information related course. Kansei Engineering approach is utilized and applied in order to improve and evaluate the e-Learning system.

Keywords: e-Learning tool for network education, Visualization of network topology and IP routing, Kansei Engineering-based Evaluation with questionnaire

1. INTRODUCTION

As Internet and its usage become popular and popular, students of universities, especially students belonging to faculties of sciences and technologies, must study network-related subjects such as computer networks, Internet structures and organization, architectures of Internet protocols, and so on. There are some problems, however, that even such students suffer from difficulty of practical understanding and graphical comprehension, because of invisibility of behaviour and structure of Internet[1][2].

From our experiences, it is effective for students to understand the target subject by means of e-Learning tool with visualization facility[3][4][5]. It is also useful for instructors to educate some invisible subjects, namely ones with abstractive concepts, with simplified and visualized models in relatively short period.

So we have decided to design and implement some kind of e-Learning tool prototype for rapid understanding of Internet introductory mechanism. The educational tool is designed and implemented as Web-based system for user convenience with easy manipulation and smart version-up. It is very much important for portability to implement it as Web-application which will work in many kinds of Web browsers, especially major browsers such as Internet Explorer, Google Chrome , FireFox and so on.

This paper presents the following sections for design, implementation and trial evaluation of our learner-centric graphical educational tool for network study, which has been developed and utilized to understand network-related subjects effectively and efficiently.

The next (= second) section introduces the related works of Japan as well as other countries. The third one describes design concepts for our educational tool and illustrates the detail configuration of our graphical educational tool from internal procedures to user interface, namely structure and behaviour of our tool. It also compares new facilities with their old versions for clear explanation of our system.

The fourth one reports trial evaluations of our tool based on some questionnaires in our university and shows significance of our system for understanding of network- related subjects. And finally the last (= fifth) one concludes some summaries and future plans for our tool prototype.

2. AN EDUCATIONAL TOOL FOR NETWORK STUDY

This section presents an educational tool for network study in Information Processing-related department of university. It explains an outline of the tool, characteristic facily of it, and animation-based learning for network behaviour.

2.1 Outline of our Educational Tool

It is important to recognize network topology and data flow between nodes in order to understand network-related subjects in universities. As you know, for the learners, a concept of ‘Packet’, tasks of each component and relation between many components (address or connectivity) must be understood approximately at the same time. Almost all the instructors should prepare some suitable materials or environment for learners to obtain not only structure but also behaviour of network mechanism. The former means 2- or 3-dimensional relationship between network components, while the later specifies time-series packet transfer from one to another.

At the view point of IP-based network structure and behaviour, ‘IP Routing’ can play the important role to specify the network structure and define packet transfer from one to another. We consider that IP routing is one of the most useful layers to learn the computer network. That is very much centered in the network-related subjects. And it is very much suitable to start learning structure and behaviour of Network. We had focused this IP routing mechanism and began to design an educational tool to understand IP routing in a short period.

Figure1 shows graphical user interface of our developed educational tool for network learning. The tool is invoked on user's browser as Web application. Its user interface is organized to provide two major sub-windows for users. At the left-hand sub-window of this tool, users can specify the network topology by means of locating IP routers, source/destination nodes and mutual relation between them with drawing directly connected lines. As we discussed later, more detail structure of user-specified network topologies will be illustrated with from Figure2 to Figure4.

At the right-hand sub window, the system automatically generates an according table to define IP routing rules, namely IP-based packet transfer between IP routers and sources/destinations which have been specified by users (learners). Learner, especially beginners of network-related subjects, can recognize IP routing detail from the system in the above way. In other words, learners can understand behaviour of IP routing without detail specification of the relevant IP routing rules.

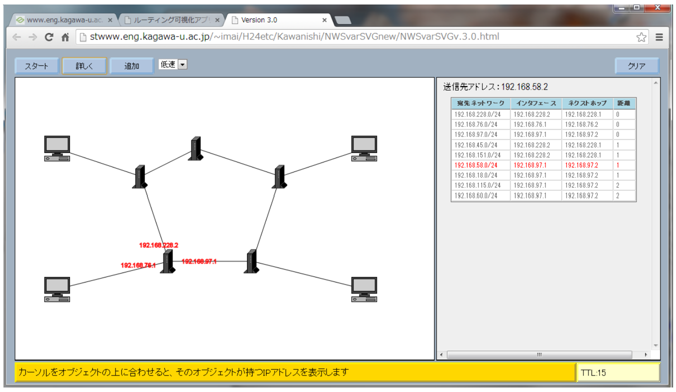


Fig. 1 Graphical user interface of Educational tool for Network study.

2.2 Visualization Facilities

A major facility of our educational tool is visualization of internal structure and behaviour of network. Visualization facility is designed with the following functions;

1. Users can select a kind of node as network component, locate it on the left-hand sub-window and configure the target network by means of repeat of the above manipulations.
2. Users can specify some node, i.e. PC or Server, as a source of packet and also specify other node as a destination. Figure2 shows PC icon surrounded by red circle as a source of packet.
3. Users can instruct the system to animate the flow of packet transfer from a node specified as a source to another node also specified as a destination through or within the network designed by users themselves.
4. Users can obtain time-series visualized image for practical demonstration of packet transfer in the IP networking level with suitable animating speed and other useful information about IP routing mechanism.

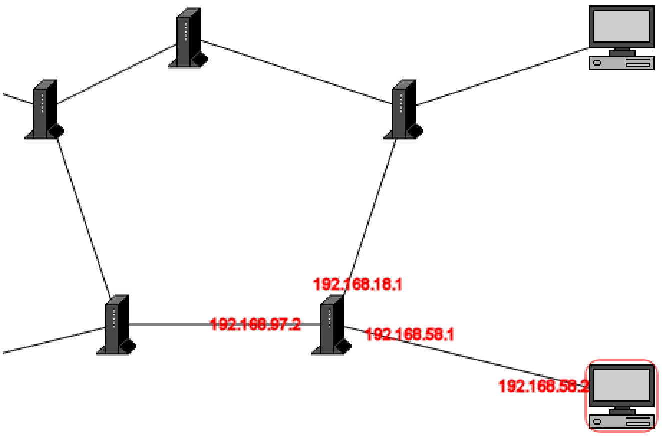


Fig. 2 Expression of IP for Router and PC.

An animation function of our educational tool is to demonstrate data flow in the network in order to show packet transfer from source component specified by user to destination one. Every packet starts from the source node specified by user where a node with packet is focused with surrounded by red circle in Figure2 and propagation of packet is illustrated with red colored line between two nodes. That means that the relevant packet is now transferred from one node to another. Figure2 also illustrates the packet transfer from an IP router with three connections at the right-bottom to a PC surrounded with red circle at the right-most.

With simple network allocation function by user and demonstrative animation one for packet transfer, our educational tool can provide visualization facility to learners of network-related subjects (= the users of our system). Based on visualization facility, therefore, our tool can play an important role to allow learners to obtain suitable image of network with its practical structure and behaviour in a relatively short period.

2.3 Animation-based Learning for Network Behaviour

The educational tool performs animated packet transfer based on user-defined network configuration and source/ destination. Users can recognize how to realize IP routing for packet transfer with IP routing information described in Figure3. When some packet transfer is being animated, IP routing information table can focus the relevant expression by means of red-colored changing. The second expression from bottom has been applied for IP routing in Figure3.

When users want to know what IP address each node has, they can only place their mouse icon on the relevant node. The tool will provide a view of the relevant IP addresses each node has in the target network configuration. Figure4

送信先アドレス：192.168.58.2

宛先ネットワーク	インタフェース	ネクストホップ	距離
192.168.97.0/24	192.168.97.2	192.168.97.1	0
192.168.58.0/24	192.168.58.1	192.168.58.2	0
192.168.18.0/24	192.168.18.1	192.168.18.2	0
192.168.228.0/24	192.168.97.2	192.168.97.1	1
192.168.76.0/24	192.168.97.2	192.168.97.1	1
192.168.45.0/24	192.168.97.2	192.168.97.1	2
192.168.151.0/24	192.168.97.2	192.168.97.1	2
192.168.115.0/24	192.168.18.1	192.168.18.2	1
192.168.60.0/24	192.168.18.1	192.168.18.2	1

Fig. 3 IP Routing Information Table showing some Specification in it to be applied with red-colored changing.

shows that a mouse-icon placing node has three IP addresses, namely, for example, 192.168.18.1 and 192.168.58.1, and 192.168.97.2 in red-colored expression. Because such a node is IP routers and it can connect two different sub network with two different interfaces. With the above functions, our educational tool can capture users' attention on the animation performing. So these functions are ones of functional contributions to visualization facility and support users' learning relatively in a short period.

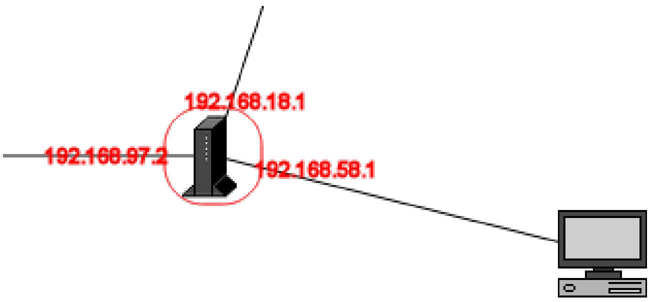


Fig. 4 Mouse-icon Placing Node(Router) with Red-Circled Expression.

3. TRIAL EVALUATION OF OUR EDUCATIONAL TOOL

This section reports a trial evaluation of our educational tool. We have prepared a few questionnaires of user's expression for our educational tool. And we have applied some kinds of Kansei Engineeringapproach to results of the questionnaires

3.1 Content Definition of Questionnaire of Learner's Impression for the Tool

The items of contents in the questionnaire have been defined in order to evaluate whether our tool is suitable for learners to understand network-related subjects. They are as follows:

Q1: Whether our tool can support your understanding packet transfer from some node to another or not ?
Q2: Whether our tool can support your recognition of relation between IP routing information and behaviour of packet transfer or not ?
Q3: Whether our tool can promote you to define user-specified network topology or not ?
Q4: Whether our tool's animation functions can contribute you to learn network structure and behaviour or not ?
Q5: Is user manipulation about network definition suitable or difficult ? (suitable: good, difficult: bad)
Q6: Is user manipulation about preparation for animation suitable or difficult ?

Especially, the last 2 items for the questionnaire, namely *Q5* and *Q6*, are adopted to confirm user-customizing facilities to configure a whole network. User manipulation must be useful and sometimes dominate GUI services so that we want to obtain user impression about user manipulation of our educational tool.

3.2 Kansei Engineering-based Evaluation

We have carried three-times questionnaires into execution for different groups in order to obtain various results of questionnaire. These would be very much useful and suggestive for system's improvement. At first, we have conduct our prototype questionnaire for members of our laboratory because of expectation to acquire some hints to brush up the contents of questionnaire. Strictly speaking, our laboratory has already contributed to define our questionnaire, so this conduction has no additional effects to modify our questionnaire prototype consequently.

Table1 shows a result of 1st trial questionnaire. And we have carried other two-times of questionnaire into execution without modification even after performing the first questionnaire.

Table 1 A Result of 1st Trial questionnaire of Learners' Impression for our Educational Tool (in our laboratory).

Question No	Excellent	Good	Poor	Bad	Total No
Q1	2	3	1	0	6
Q2	1	4	1	0	6
Q3	4	1	0	1	6
Q4	4	1	1	0	6
Q5	5	1	0	0	6
Q6	5	0	0	0	6

The second time of questionnaire is to carry the same content of the previous questionnaire into execution for students of information engineering department in the second year. They have already begun to learn computer network, so that they are familiar with network-related subjects. The result of questionnaire is very good and their evaluation for our educational tool is the best of all the three times questionnaires.

It is written in Table2.

Table 2 A Result of 2nd official questionnaire of Learners' Impression for our Educational Tool (in a real classroom).

Question No	Excellent	Good	Poor	Bad	Total No
Q1	11	11	1	0	23
Q2	9	13	1	0	23
Q3	11	10	2	0	23
Q4	13	6	4	0	23
Q5	12	7	3	1	23
Q6	14	5	2	2	23

The last one is to carry the questionnaire into execution for students in the first year but they do not include members of engineering faculty. Such students are some sort of beginners of Internet and their knowledge and experience about computer network and Internet is not so much but they are interesting in Internet through their daily digital lives. The result of their questionnaire is in Table3. We donot recognize why the last 2 items for the questionnaire, namely *Q5* and *Q6*, are not answered by 2 students.

Table 3 A Result of 3rd official questionnaire of Learners' Impression for our Educational Tool (in a real classroom).

Question No	Excellent	Good	Poor	Bad	Total No
Q1	8	10	1	0	19
Q2	6	11	2	0	19
Q3	4	13	2	0	19
Q4	3	14	2	0	19
Q5	1	8	6	2	17
Q6	1	7	7	2	17

It is confirmed that the average of three results of questionnaires are good and evaluations from the relevant learners (i.e. users of our educational tool) are almost excellent or good in Figure5. Of course, our educational tool must be improved more and more in the future, but impression of users sounds like good and favorably-disposed toward our educational tool.

4. CONCLUSION

This paper describes summaries of several network- related researches, configuration of our educational tool for network study and its evaluation through the three-times questionnaires in our university. The tool has been designed and implemented as Web application and useful for users to learn several network-related subjects in university. We have obtained following evaluation from the above execution of questionnaires. Our conclusion is summarized as follows:

(1) Our educational tool has been developed as one of Web applications and useful for users to learn network-related subjects in university.

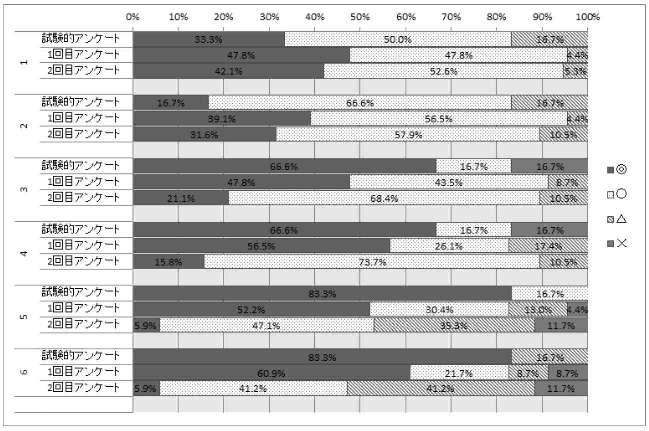


Fig. 5 Integration of the three results of questionnaires

(2) Our tool is implemented with two-coupled modules to provide user defining network with some kinds of components such as node and connectivity and then perform demonstrative animation for visualization of internal structure and behaviour of the target network.

(3) It is confirmed that user impression about our tool is good and users evaluate our tool very well through the three-times questionnaires for many students with different expertise.

Acknowledgement

This research can be performed by kind supports and interests from all the students who have participated in the classroom lectures.

REFERENCES

[1] Arai, M., Tamura, N., Watanabe, H., Ogiso, C. Takei, S.: Development and Evaluation of TCP/IP Protocol Learning Tools (in Japanese). In: Journal of Information Processing Society Japan (JIPSJ), Vol.44, No.12, pp.3242–3251, 2003.

[2] Tateiwa, Y. Yasuda, T., Yokoi, S.: Development of a System to Visualize Computer Network Behavior for Learning to Associate LAN Construction Skills with TCP/IP Theory, Based on Virtual Environment Software (2007). In: Journal of Information Processing Society Japan (JIPSJ), Vol.48, No.4, pp.1684–1694, 2007.

[3] Stafford, T., F.: Understanding Motivations for Internet Use in Distance Education. In: IEEE Transactions on Education, Vol.48, No.2, pp.301–306, 2005.

[4] Sivakumar, S. C., Robertson, W., Artimy, M., Aslam, N.: A Web-Based Remote Interactive Laboratory for Internetworking Education. In: IEEE Transactions on Education, Vol.48, No.4, pp.586–598, 2005.

[5] Snow, C., Pullen, J.M., McAndrews, P: Network Education Ware: An Open-Source Web-Based System for Synchronous Distance Education. In: IEEE Transactions on Education, Vol.48, No.4, pp.705–712, 2005.

Model Introduced SPRT for Structural Change Detection of Time Series (I)

*Yoshihide KOYAMA, *Tetsuo HATTORI, **Hiromichi KAWANO

**Graduate School of Engineering, Kagawa University / 2217-20 Hayashi, Takamatsu City, Kagawa 761-0396, Japan*
***NTT Advanced Technology/ Musashino-shi Nakamachi 19-18, Tokyo 180-0006, Japan*
(hattori@eng.kagawa-u.ac.jp)

Abstract: Previously, we have proposed a method applying Sequential Probability Ratio Test (SPRT) to structural change detection problem of ongoing time series data. In this paper, we introduce a structural change model with Poisson process into the system that outputs a set of time series data, moment by moment. The model can be considered as a kind of Hidden Markov Model. Then we can calculate the value of P(Z |H1) that denotes the probability of observing the data Z under the condition that H1(alternative hypothesis) is true. Also, we concretely show the theory of change detection of time series by the model introduced SPRT.

Keywords: Time series, Change detection, SPRT(Sequential Probability Ratio Test), Hidden Markov Model

I. INTRODUCTION

To make a prediction of ongoing time series data, we have three stages in general ([1], [2]). First, we have to find a prediction model that adequately represents the characteristics of the early time series data. Second, we have to detect the structural change of the time series data, as quickly and correctly as possible, when the estimated prediction model does not meet the data any more as shown in Fig.1 ([3],[4]). Third, we have to reconstruct the next prediction model as soon as possible after the change detection.

For the second problem, we have already proposed an application of SPRT (sequential probability ratio test) that has been mainly used in the field of quality control [5], [6]. And we presented the experimental results in comparison with Chow Test that is well-known standard method for such structural change detection of time series data ([6], [7]).

However, in the SPRT method that we previously proposed for the structural change detection problem, the probability of structural change occurrence is not known beforehand. In addition, in the method, only the ratio of the conditional probability is used. Then, the formulation of SPRT includes the notion that the structure may be able to recover even after the change occurrence. Then, there is a possibility that error happens in the change detection using the ratio of the conditional probability and so the threshold of the SPRT for the detection is decided based on the error probability.

In this paper, we introduce a Hidden Markov Model with Poisson process on the structural change one and propose a new probability ratio method for the change detection problem.

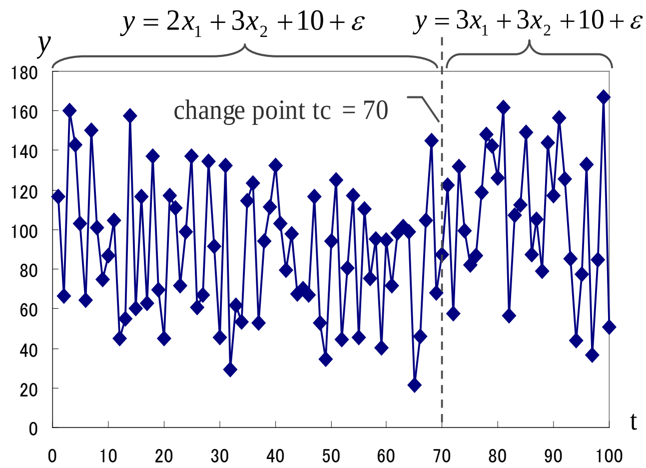


Fig.1. Example of time series data where the true change point tc=70.

II. SPRT

1. SPRT ([8])

The Sequential Probability Ratio Test (SPRT) is used for testing a null hypothesis H_0 (e.g. the quality is under pre-specified limit 1%) against hypothesis H_1 (e.g. the quality is over pre-specified limit 1%). And it is defined as follows:

Let Z_1, Z_2, $\cdots Z_i$ be respectively observed time series data at each stage of successive events, the probability ratio λ_i is computed as follows.

$$\lambda_i = \frac{P(Z_1 \mid \mathbf{H}_1)\cdot P(Z_2 \mid \mathbf{H}_1)\cdots P(Z_i \mid \mathbf{H}_1)}{P(Z_1 \mid \mathbf{H}_0)\cdot P(Z_2 \mid \mathbf{H}_0)\cdots P(Z_i \mid \mathbf{H}_0)} \quad (1)$$

where $P(Z \mid H_0)$ denotes the distribution of Z if H_0 is true, and similarly, $P(Z \mid H_1)$ denotes the distribution of Z if H_1 is true.

Two positive constants C_1 and C_2 ($C_1 < C_2$) are chosen. If $C_1 < \lambda_i < C_2$, the experiment is continued by taking an additional observation. If $C_2 < \lambda_i$, the process is terminated with the rejection of H_0 (acceptance of H_1). If $\lambda_i < C_1$, then terminate this process with the acceptance of H_0.

$$C_1 = \frac{\beta}{1-\alpha}, \quad C_2 = \frac{1-\beta}{\alpha} \quad (2)$$

where α means type I error (reject a true null hypothesi s), and β means type II error (accept a null hypothesis as true one when it is actually false).

2. Procedure of SPRT ([8])

The concrete procedure of applying the SPRT method to the structural change detection problem is described using the notation of Fig. 2 as follows.

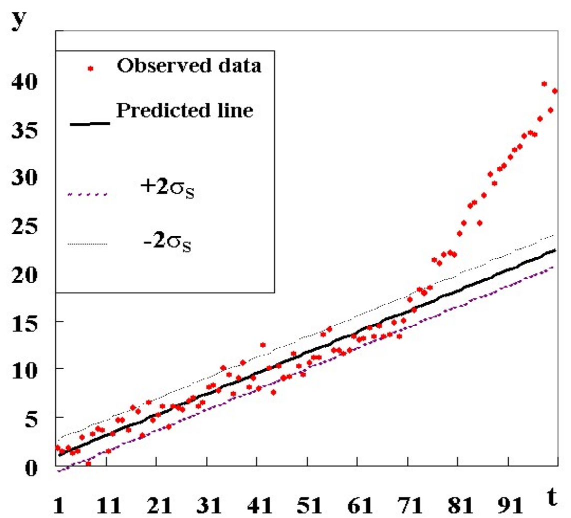


Fig.2. Example of time series data and single regression line, where the true change point tc =70 ([6]).

Step1: Make a prediction expression and set the tolerance band (***a***) (e.g. ***a***=$2\sigma_S$) that means permissible error margin between the predicted data and the observed one. (σ_S denotes a standard deviation in learning sample data at early stage.)

Step2 : Set up the null hypothesis H_0 and alternative hypothesis H_1.

H_0 : Change has not occurred yet.

H_1 : Change has occurred.

Set the values α, β and compute C_1 and C_2, according to Equation (2). Initialize i = 0, $\lambda_0 = 1$.

Step3: Incrementing i (i = i+1), observe the following data y_i. Evaluate the error $|\varepsilon_i|$ between the data y_i and the predicted value from the aforementioned prediction expression.

Step4: Judge as to whether the data y_i goes in the tolerance band or not, i.e., the ε_i is less than (or equal to) the permissible error margin or not. If it is Yes, then set $\lambda_i = 1$ and return to Step3. Otherwise, advance to Step5.

Step5: Calculate the probability ratio λ_i, using the following Equation (3) that is equivalent to Equation (1).

$$\lambda_i = \lambda_{i-1}\frac{P(\varepsilon_i \mid \mathbf{H}_1)}{P(\varepsilon_i \mid \mathbf{H}_0)} \quad (3)$$

where, if the data y_i goes out the tolerance band, $(P(\varepsilon_i|H_0), P(\varepsilon_i|H_1)) = (\theta_0, \theta_1)$, otherwise, $(P(\varepsilon_i|H_0), P(\varepsilon_i|H_1)) = ((1-\theta_0), (1-\theta_1))$.

Step6: Execution of testing.

(i) If the ratio λ_i is greater than C_2 (= $(1-\beta)/\alpha$), dismiss the null hypothesis H_0, and adopt the alternative hypothesis H_1, and then End.

(ii) Otherwise, if the ratio λ_i is less than C_1 (= $\beta/(1-\alpha)$), adopt the null hypothesis H_0, and dismiss the alternative hypothesis H_1, and then set $\lambda_i = 1$ and return to Step3.

(iii) Otherwise (in the case where $C_1 \le \lambda_i \le C_2$), advance to Step7.

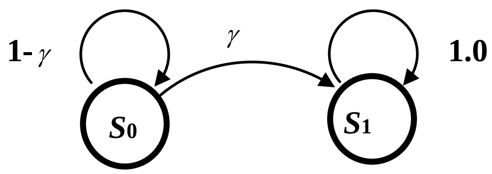
Step7: Observe the following data y_i incrementing i. Evaluate the error $|\varepsilon_i|$ and judge whether the data y_i goes in the tolerance band, or not. Then, return to Step5 (calculation of the ratio λ_i).

III. STRUCTURAL CHANGE MODEL INTRODUCED

1. Hidden Markov Model

We present a Hidden Markov Model where the structural change is Poisson occurrence of average γ. In this model, once the change has occurred during the observing period, the structure does not go back to the previous one. The reason why we set such a model is that we focus on the detection of the first structural change in the sequential processing. The concept of the structural change model is shown in Fig.3.

Moreover, we introduce a more detailed model. Let ***R*** be the probability of the prediction failure that the data value goes out the tolerance zone when the structure is unchanged. Let ***Rc*** be the probability of the unfitting when the structure change occurred. We can consider that ***Rc*** is greater than ***R***, i.e., ***Rc*>*R***. The detailed model is illustrated as a probabilistic finite state automaton in Fig.4.



***S*0 :** State that the structure is unchanged.
***S*1 :** State that the structural change has occurred.
γ **:** Probability of the structural change occurrence. (Poisson Process.)

Fig.3. Structural change model.

2. New Sequential Probability Ratio

Let $a_1 a_2,\ldots,a_i,\ldots a_n \quad a_i \in \{\mathrm{IN},\mathrm{OUT}\}$ be a string (or symbol sequence) obtained from the observed data.

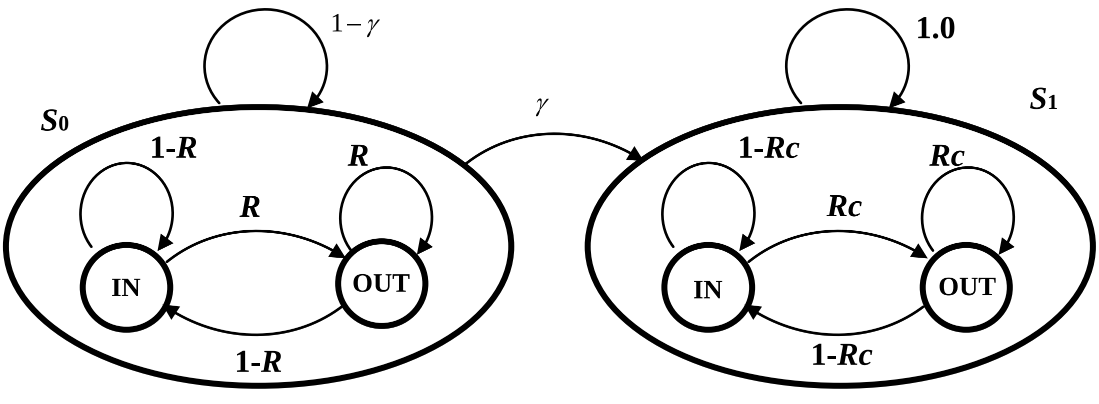
Let θ_i and $\tilde{\theta}_i$ be the conditional probability that outputs the observed data (or above symbol sequence, $C_n = a_1 a_2 \ldots a_n$ in the state ***S*0** and ***S*1**, respectively. That is, it means that $\theta_i \in \{R, 1-R\}$ and $\tilde{\theta}_i \in \{R_c, 1-R_c\}$, respectively.

And let $P(a_1 \ldots a_n, \mathrm{H}_0)$ and $P(a_1 \ldots a_n, \mathrm{H}_1)$ be the joint probability of the symbol sequence C_n happens with the event H_0 (the structural change is not occurred) and H_1 (the change is occurred), respectively.

Then, the following equations hold.

$$P(a_1 \ldots a_n, \mathrm{H}_0) = P(C_n, \mathrm{H}_0) = (1-\gamma)^n \theta_1 \ldots \theta_n = (1-\gamma)^n \prod_{i=1}^{n} \theta_i \tag{4}$$

$$\begin{aligned} P(a_1 \ldots a_n, \mathrm{H}_1) &= P(C_n, \mathrm{H}_1) \\ &= \gamma \prod_{i=1}^{n} \tilde{\theta}_i + ((1-\gamma)\theta_1)(\gamma \prod_{i=2}^{n} \tilde{\theta}_i) \\ &\quad + ((1-\gamma)^2 \theta_1 \theta_2)(\gamma \prod_{i=3}^{n} \tilde{\theta}_i) + \ldots \\ &= \sum_{k=1}^{n} ((1-\gamma)^{k-1} \cdot \prod_{j=0}^{k-1} \theta_j)(\gamma \prod_{i=k}^{n} \tilde{\theta}_i) \end{aligned} \tag{5}$$



R : Probability of prediction failure (**OUT**) when the structure is unchanged.
Rc : Probability of prediction failure (**OUT**) when the structural change has been occurred.
IN: The state that observed data goes into the tolerance zone.
OUT: The state that observed data goes out of the tolerance zone.

Fig.4. Proposed Hidden Markov Model.

$$P(\mathrm{H}_0 \mid a_1 \ldots a_n) = P(\mathrm{H}_0 \mid C_n) = \frac{P(\mathrm{H}_0 \mid C_n) \cdot P(C_n)}{P(C_n)}$$
$$= \frac{P(C_n, \mathrm{H}_0)}{P(C_n)} = \frac{P(C_n, \mathrm{H}_0)}{P(C_n, \mathrm{H}_0) + P(C_n, \mathrm{H}_1)} \tag{6}$$

$$P(\mathrm{H}_1 \mid a_1 \ldots a_n) = P(\mathrm{H}_1 \mid C_n) = \frac{P(\mathrm{H}_1 \mid C_n) \cdot P(C_n)}{P(C_n)}$$
$$= \frac{P(C_n, \mathrm{H}_1)}{P(C_n)} = \frac{P(C_n, \mathrm{H}_1)}{P(C_n, \mathrm{H}_0) + P(C_n, \mathrm{H}_1)} \tag{7}$$

New Sequential Probability Ratio (NSPR) that we propose is represented using the aforementioned equations as follows.

$$\mathrm{NSPR} = \frac{P(\mathrm{H}_1 \mid a_1 \ldots a_n)}{P(\mathrm{H}_0 \mid a_1 \ldots a_n)} = \frac{P(\mathrm{H}_1 \mid C_n)}{P(\mathrm{H}_0 \mid C_n)} = \frac{P(C_n, \mathrm{H}_1)}{P(C_n, \mathrm{H}_0)}$$
$$= \boxed{\frac{\sum_{k=1}^{n} ((1-\gamma)^{k-1} \cdot \prod_{j=0}^{k-1} \theta_j)(\gamma \prod_{i=k}^{n} \tilde{\theta}_i)}{(1-\gamma)^n \theta_1 \ldots \theta_n = (1-\gamma)^n \prod_{i=1}^{n} \theta_i}} \tag{8}$$

When the NSPR is greater than 1.0, we regard that the structural change has been occurred before the present time.

IV. CONCLUSION

We have presented a Hidden Markov Model for the structural change detection problem, and have proposed a New Sequential Probability Ratio method for the detection using the occurrence probability based on the proposed model.

We consider that this method will be more promising than the SPRT one that we previously applied it to the change detection problem. We will describe the experimental results at another opportunity.

REFERENCES

[1] C.G.E.P.Box and G.M.Jenkins (1976), Time Series Analysis:Forecasting and Control, Prentice Hall.

[2] Peter J. Brockwell and Richard A. Davis (2003), Introduction to Time Series and Forecasting, Springer; 2nd edition.

[3] C.Han, P.k.Willet and D.A.Abraham (1999), Some methods to evaluate the performance of Page's test as used to detect transient signals, IEEE Trans. Signal processing, Vol.47, No.8, pp.2112-2127.

[4] S.D.Blostein (1991), Quickest detection of a time-varying change in distribution, IEEE Trans. Information Theory, Vol.37, No.4, pp.1116-1122.

[5] A. Wald (1947), Sequential Analysis, John Wiley & Sons.

[6] Hiromichi Kawano, Tetsuo Hattori, Ken Nishimatsu (2008), Structural Change Point Detection Method of Time Series Using Sequential Probability Ratio Test－Comparison with Chow Test in the ability of early detection－(in Japanese), IEEJ Trans. EIS, Vol.128, No.4, pp.583-592.

[7] Chow, G. C. (1960), Tests of Equality Between Sets of Coefficients in Two Linear Regressions, Econometrica, Vol.28, No.3, pp.591-605.

[8] Katsunori Takeda, Tetsuo Hattori, Izumi Tetsuya, Hiromichi Kawano (2010), Extended SPRT for Structural Change Detection of Time Series Based on Multiple Regression Mode, Proceedings of the 15th International Symposium on Artificial Life and Robotics (AROB 15th '10), pp.755-758, ISBN: 978-4-9902880-4-4, Oita, Japan, Feb.

[9] Katsunori Takeda, Tetsuo Hattori, Tetsuya Izumi, Hiromichi Kawano (2010), "Extended SPRT for Structural Change Detection of Time Series Based on Multiple Regression Model," International Journal of Artificial Life and Robotics, Springer, ISSN:1433-5298, Vo.15, No.4, pp.417-420, 2010.

[10] Tetsuo HATTORI, Hiromichi KAWANO (2011), Change Detection Method of Time Series as an Optimal Stopping Problem -- Constructive Proof of Optimal Solution Theorem --, Proc. of ICBAKE 2011 (2011 International Conference on Biometrics and Kansei Engineering), IEEE Computer Society, ISBN 978-0-7695-4512-7, pp.100-105, Kagawa, Japan, 2011.

Model Introduced SPRT for Structural Change Detection of Time Series (II)

--- Kansei Channel and Bayes' Estimation ---

*Yoshihide KOYAMA, *Tetsuo HATTORI, **Hiromichi KAWANO

**Graduate School of Engineering, Kagawa University / 2217-20 Hayashi, Takamatsu City, Kagawa 761-0396, Japan*
***NTT Advanced Technology/ Musashino-shi Nakamachi 19-18, Tokyo 180-0006, Japan*
(hattori@eng.kagawa-u.ac.jp)

Abstract: In this paper, using the notion of a binary Channel Matrix what we call "Kansei Channel" as well known in Information Theory, we present an equivalent relation between the SPRT and Bayes' Updating. Moreover, we show the relationship between the SPRT (Sequential Probability Ratio Test) and the Sequential Probability Ratio of two corresponding conditional probabilities when a Hidden Markov Model with Poisson process is introduced as the structural change model.

Keywords: Change detection, SPRT, Hidden Markov Model, Information Theory, Kansei Channel, Bayes' Updating

I. INTRODUCTION

There are three problems in dealing with ongoing time series data ([1]-[5]). First, we have to make a prediction model from the early data. Second, we have to detect the structural change of the time series data, as quickly and correctly as possible, when the prediction model does not meet the observing data any more. Third, we have to remake the prediction model as soon as possible after the detected change point. For the second change detection problem, we previously proposed an application of SPRT (Sequential Probability Ratio Test) that has been mainly used in the field of quality control ([5]-[9]).

In this paper, using the notion of a binary Channel Matrix what we call "Kansei Channel" as well known in Information Theory, we present an equivalent relation between the SPRT and Bayes' Updating. Moreover, we also explain the relation between the SPRT and the Sequential Probability Ratio of two conditional probabilities when a Hidden Markov Model with Poisson process is introduced as the never returning structural change model [10].

II. SPRT AND BAYES' UPDATING

1. SPRT ([8], [9])

The Sequential Probability Ratio Test (SPRT) is used for testing a null hypothesis H_0 (e.g. the quality is under pre-specified limit 1%) against hypothesis H_1 (e.g. the quality is over pre-specified limit 1%). And it is defined as follows:

Let Z_1, Z_2, $\cdots Z_i$ be respectively observed time series data at each stage of successive events, the probability ratio λ_i is computed as follows.

$$\lambda_i = \frac{P(Z_1 \mid \mathbf{H}_1) \cdot P(Z_2 \mid \mathbf{H}_1) \cdots P(Z_i \mid \mathbf{H}_1)}{P(Z_1 \mid \mathbf{H}_0) \cdot P(Z_2 \mid \mathbf{H}_0) \cdots P(Z_i \mid \mathbf{H}_0)} \quad (1)$$

where $P(Z \mid H_0)$ denotes the distribution of Z if H_0 is true, and similarly, $P(Z \mid H_1)$ denotes the distribution of Z if H_1 is true.

Two positive constants C_1 and C_2 ($C_1 < C_2$) are chosen. If $C_1 < \lambda_i < C_2$, the experiment is continued by taking an additional observation. If $C_2 < \lambda_i$, the process is terminated with the rejection of H_0 (acceptance of H_1). If $\lambda_i < C_1$, then terminate this process with the acceptance of H_0.

$$C_1 = \frac{\beta}{1-\alpha}, \quad C_2 = \frac{1-\beta}{\alpha} \quad (2)$$

where α means type I error (reject a true null hypothesi s), and β means type II error (accept a null hypothesis as true one when it is actually false).

2. Bayes' Updating Based on Binary Channel

We regard that the observation and SPRT process for the aforementioned structural change point detection of time series data is essentially a kind of Bayes' Updating by a binary channel what we call "Kansei

Channel" composed of a set of subjective probability as shown in Fig.1, where input event $X=(x_1, x_2)$ and output event $Y=(y_1, y_2)$.

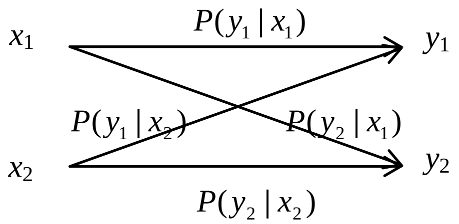


Fig.1. Binary channel (Kansei Channel).

Using the conditional probability, the transition matrix (or channel matrix) $\boldsymbol{P}_B$ is represented as follows.

$$\boldsymbol{P}_B = \begin{bmatrix} P(y_1 | x_1) & P(y_2 | x_1) \\ P(y_1 | x_2) & P(y_2 | x_2) \end{bmatrix} \tag{3}$$

Let the prior probability distribution of input event X be $\{P_0(x_1), P_0(x_2)\}$ as initial values. Let the observed output sequence $\{ y_{j_1}, y_{j_2}, \cdots, y_{j_n} \}$ ($y_{j_k} = y_1$ or y_2 (k=1, 2,…, n)). Then, by the first observation of output, the posterior probability is given by Bayes' theorem.

$$p(x_1 | y_{j_1}) = \frac{P_0(x_1)P(y_{j_1} | x_1)}{P(y_{j_1})}$$

$$p(x_2 | y_{j_1}) = \frac{P_0(x_2)P(y_{j_1} | x_2)}{P(y_{j_1})}$$

$$\text{where, } P(y_{j_1}) = \sum_i P_0(x_i)P(y_{j_1} | x_i) \tag{4}$$

According to the Bayes' Updating rule, we have the following updated prior probability $\{P_1(x_1), P_1(x_2)\}$.

$$P_1(x_1) = p(x_1 | y_{j_1}) = \frac{P_0(x_1)p(y_{j_1} | x_1)}{P(y_{j_1})}$$

$$P_1(x_2) = p(x_2 | y_{j_1}) = \frac{P_0(x_2)p(y_{j_1} | x_2)}{P(y_{j_1})} \tag{5}$$

Similarly, by the second observation, we have

$$P_2(x_1) = p(x_1 | y_{j_2}) = \frac{P_1(x_1)p(y_{j_2} | x_1)}{P(y_{j_2})}$$

$$P_2(x_2) = p(x_2 | y_{j_2}) = \frac{P_1(x_2)p(y_{j_2} | x_2)}{P(y_{j_2})}$$

$$\text{where, } \quad P(y_{j_2}) = \sum_i P_1(x_i)P(y_{j_2} | x_i) \tag{6}$$

Thus, by the nth observation of output, we have

$$P_n(x_i) = p(x_i | y_{j_n}) = \frac{P_{n-1}(x_i)P(y_{j_n} | x_i)}{P(y_{j_n})}$$

$$= \frac{P_0(x_i)\prod_{k=1}^{n} P(y_{j_k} | x_i)}{\prod_{k=1}^{n} P(y_{j_k})} \qquad (i = 1,2) \tag{7}$$

Then the probability ratio λ_n is as follows.

$$\lambda_n = \frac{P_n(x_2)}{P_n(x_1)} = \frac{P_0(x_2)\prod_{k=1}^{n} P(y_{j_k} | x_2)}{P_0(x_1)\prod_{k=1}^{n} P(y_{j_k} | x_1)} \tag{8}$$

If we assume that

$$P_0(x_1) = P_0(x_2) = \frac{1}{2}$$

Then, we obtain the same ratio as that of SPRT. Therefore, we consider that the Bayes' Updating based on binary channel is a kind of generalization of SPRT.

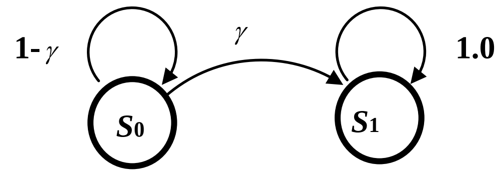
As for the conditional probability, we can introduce our subjective probability. In addition, we can adopt not only the discrete probability, but also continuous density function.

III. CHANGE MODEL AND CONDITIONAL PROBABILITY RATIO

1. Hidden Markov Model

We present a Hidden Markov Model where the structural change is Poisson occurrence of average γ . In this model, once the change has occurred during the observing period, the structure does not go back to the previous one. The reason why we set such a model is that we focus on the detection of the first structural change in the sequential processing. The concept of the structural change model is shown in Fig.2-4.

Moreover, we introduce a more detailed model. Let ***R*** be the probability of the prediction failure that the data value goes out the tolerance zone when the structure is unchanged. Let ***Rc*** be the probability of the failure when the structure change has been occurred. We can consider that ***Rc*** is greater than ***R***, i.e., ***Rc*>*R***. The detailed model is illustrated as a probabilistic finite state automaton in Fig.2-4.



S0 : State that the structure is unchanged.
S1 : State that the structural change has occurred.
γ **:** Probability of the structural change occurrence. (Poisson Process.)

Fig.2. Structural change model.

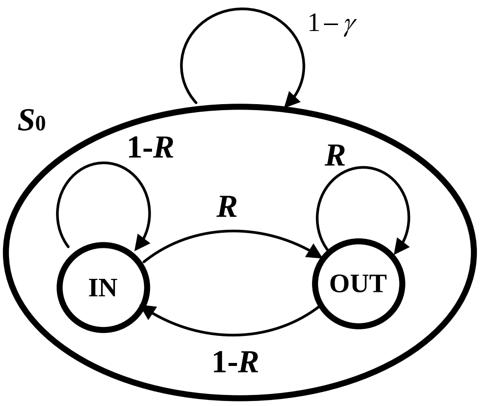


Fig.3. Internal model of the State ***S0***.

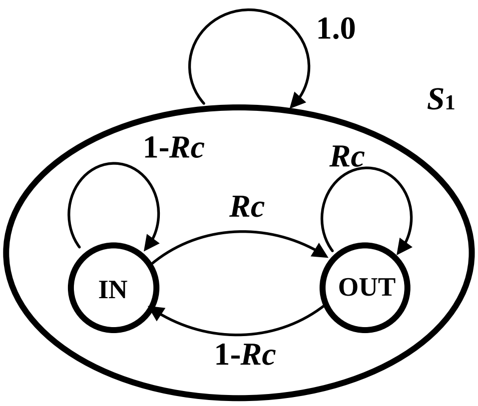


Fig.4. Internal model of the State ***S1***.

2. Sequential Probability Ratio

Let $a_1 a_2, \ldots, a_i, \ldots a_n \quad a_i \in \{\text{IN}, \text{OUT}\}$ be a string (or symbol sequence) obtained from the observed data.

Let θ_i and $\tilde{\theta}_i$ be the conditional probability that outputs the observed data (or above symbol sequence, $C_n = a_1 a_2 \ldots a_n$ in the state ***S0*** and ***S1***, respectively. That is, it means that $\theta_i \in \{R, 1-R\}$ and $\tilde{\theta}_i \in \{R_c, 1-R_c\}$, respectively.

And let $P(a_1 \ldots a_n, \mathrm{H}_0)$ and $P(a_1 \ldots a_n, \mathrm{H}_1)$ be the joint probability of the symbol sequence C_n happens with the event H_0 (the structural change is not occurred) and H_1 (the change is occurred), respectively.

Then, the following equations hold.

$$P(a_1 \ldots a_n, \mathrm{H}_0) = P(C_n, \mathrm{H}_0) = (1-\gamma)^n \theta_1 \ldots \theta_n = (1-\gamma)^n \prod_{i=1}^{n} \theta_i \tag{9}$$

$$\begin{aligned} P(a_1 \ldots a_n, \mathrm{H}_1) &= P(C_n, \mathrm{H}_1) \\ &= \gamma \prod_{i=1}^{n} \tilde{\theta}_i + ((1-\gamma)\theta_1)(\gamma \prod_{i=2}^{n} \tilde{\theta}_i) \\ &\quad + ((1-\gamma)^2 \theta_1 \theta_2)(\gamma \prod_{i=3}^{n} \tilde{\theta}_i) + \ldots \\ &= \sum_{k=1}^{n} ((1-\gamma)^{k-1} \cdot \prod_{j=0}^{k-1} \theta_j)(\gamma \prod_{i=k}^{n} \tilde{\theta}_i) \end{aligned} \tag{10}$$

$$\begin{aligned} P(\mathrm{H}_0 \mid a_1 \ldots a_n) &= P(\mathrm{H}_0 \mid C_n) = \frac{P(\mathrm{H}_0 \mid C_n) \cdot P(C_n)}{P(C_n)} \\ &= \frac{P(C_n, \mathrm{H}_0)}{P(C_n)} = \frac{P(C_n, \mathrm{H}_0)}{P(C_n, \mathrm{H}_0) + P(C_n, \mathrm{H}_1)} \end{aligned} \tag{11}$$

$$\begin{aligned} P(\mathrm{H}_1 \mid a_1 \ldots a_n) &= P(\mathrm{H}_1 \mid C_n) = \frac{P(\mathrm{H}_1 \mid C_n) \cdot P(C_n)}{P(C_n)} \\ &= \frac{P(C_n, \mathrm{H}_1)}{P(C_n)} = \frac{P(C_n, \mathrm{H}_1)}{P(C_n, \mathrm{H}_0) + P(C_n, \mathrm{H}_1)} \end{aligned} \tag{12}$$

A New Sequential Probability Ratio (NSPR) of two conditional probabilities is represented as follows, using the aforementioned equations.

$$\begin{aligned} \mathrm{NSPR} &= \frac{P(\mathrm{H}_1 \mid a_1 \ldots a_n)}{P(\mathrm{H}_0 \mid a_1 \ldots a_n)} = \frac{P(\mathrm{H}_1 \mid C_n)}{P(\mathrm{H}_0 \mid C_n)} = \frac{P(C_n, \mathrm{H}_1)}{P(C_n, \mathrm{H}_0)} \\ &= \frac{\sum_{k=1}^{n} ((1-\gamma)^{k-1} \cdot \prod_{j=0}^{k-1} \theta_j)(\gamma \prod_{i=k}^{n} \tilde{\theta}_i)}{(1-\gamma)^n \prod_{i=1}^{n} \theta_i} \end{aligned} \tag{13}$$

On the contrary, there is another type of conditional probability ratio such as used in the SPRT. We call it "Semi Sequential Probability Ratio Test (SSPRT)".

That is,

$$\mathrm{SSPRT} = \frac{P(a_1 \ldots a_n \mid \mathrm{H}_1)}{P(a_1 \ldots a_n \mid \mathrm{H}_0)} = \frac{P(C_n \mid \mathrm{H}_1)}{P(C_n \mid \mathrm{H}_0)}$$
$$= \frac{P(C_n, \mathrm{H}_1)/P(\mathrm{H}_1)}{P(C_n, \mathrm{H}_0)/P(\mathrm{H}_0)}$$
$$= \left(\frac{P(\mathrm{H}_0)}{P(\mathrm{H}_1)}\right) \cdot \left(\frac{P(C_n, \mathrm{H}_1)}{P(C_n, \mathrm{H}_0)}\right) = \left(\frac{P(\mathrm{H}_0)}{P(\mathrm{H}_1)}\right) \cdot \mathrm{NSPR} \tag{14}$$

Let a notation $S_0{}^k \cdot S_1{}^l$ mean an event sequence where, after the state S_0 continuously occurs k times, the state S_1 continuously occurs l times. Then, the $P(\mathrm{H}_0)$ and $P(\mathrm{H}_1)$ are represented as follows.

$$P(\mathrm{H}_0) = P(S_0{}^n) = (1-\gamma)^n \tag{15}$$

$$P(\mathrm{H}_1) = P(S_1{}^n) + P(S_0 \cdot S_1{}^{n-1}) + P(S_0{}^2 \cdot S_1{}^{n-2})$$
$$+ P(S_0{}^k \cdot S_1{}^{n-k}) + \ldots + P(S_0{}^{n-1} \cdot S_1)$$
$$= \sum_{k=o}^{n-1} P(S_0{}^k \cdot S_1{}^{n-k}) = \sum_{k=o}^{n-1} (1-\gamma)^k \cdot \gamma$$
$$= 1 - P(\mathrm{H}_0) = 1 - (1-\gamma)^n \tag{16}$$

Then we have the relation between the NSPR and Semi SPRT as follows.

$$\mathrm{NSPR} = \left(\frac{P(\mathrm{H}_1)}{P(\mathrm{H}_0)}\right) \cdot \mathrm{SSPRT} = \left(\frac{1}{(1-\gamma)^n} - 1\right) \cdot \mathrm{SSPRT} \tag{17}$$

When the NSPR is greater than 1.0, we regard that the structural change has been occurred before the present time.

IV. CONCLUSION

Using the notion of a binary Channel as well known in Information Theory, we have shown that the SPRT is equivalent to the Bayes' Updating Rule. Moreover, we have explained the relation between the SPRT and the Sequential Probability Ratio of two conditional probabilities when a Hidden Markov Model with Poisson process is introduced as a structural change model.

REFERENCES

[1] C.G.E.P.Box and G.M.Jenkins (1976), Time Series Analysis:Forecasting and Control, Prentice Hall.

[2] Peter J. Brockwell and Richard A. Davis (2003), Introduction to Time Series and Forecasting, Springer; 2nd edition.

[3] C.Han, P.k.Willet and D.A.Abraham (1999), Some methods to evaluate the performance of Page's test as used to detect transient signals, IEEE Trans. Signal processing, Vol.47, No.8, pp.2112-2127.

[4] S.D.Blostein (1991), Quickest detection of a time-varying change in distribution, IEEE Trans. Information Theory, Vol.37, No.4, pp.1116-1122.

[5] A. Wald (1947), Sequential Analysis, John Wiley & Sons.

[6] Hiromichi Kawano, Tetsuo Hattori, Ken Nishimatsu (2008), Structural Change Point Detection Method of Time Series Using Sequential Probability Ratio Test－Comparison with Chow Test in the ability of early detection－(in Japanese), IEEJ Trans. EIS, Vol.128, No.4, pp.583-592.

[7] Chow, G. C. (1960), Tests of Equality Between Sets of Coefficients in Two Linear Regressions, Econometrica, Vol.28, No.3, pp.591-605.

[8] Katsunori Takeda, Tetsuo Hattori, Izumi Tetsuya, Hiromichi Kawano (2010), Extended SPRT for Structural Change Detection of Time Series Based on Multiple Regression Mode, Proceedings of the 15th International Symposium on Artificial Life and Robotics (AROB 15th '10), pp.755-758, ISBN: 978-4-9902880-4-4, Oita, Japan, Feb.

[9] Katsunori Takeda, Tetsuo Hattori, Tetsuya Izumi, Hiromichi Kawano (2010), "Extended SPRT for Structural Change Detection of Time Series Based on Multiple Regression Model," International Journal of Artificial Life and Robotics, Springer, ISSN:1433-5298, Vo.15, No.4, pp.417-420, 2010.

[10] Tetsuo HATTORI, Hiromichi KAWANO (2011), Change Detection Method of Time Series as an Optimal Stopping Problem -- Constructive Proof of Optimal Solution Theorem --, Proc. of ICBAKE 2011 (2011 International Conference on Biometrics and Kansei Engineering), IEEE Computer Society, ISBN 978-0-7695-4512-7, pp.100-105, Kagawa, Japan, 2011.

Automated Color Image Arrangement Method and Kansei Impression

Y. Kawakami[1], T. Hattori[1], Y. Imai[1], H. Matsushita[1], H. Kawano[2] and R.P.C.J. Rajapakse[3]

[1]*Kagawa University, 2217-20, Hayashi, Takamatsu, Kagawa, 761-0396 Japan*
Tel : 81-87-864-2000; Fax : 81-87-864-2032
s11d621@stmail.eng.kagawa-u.ac.jp, {hattori, imai, haruna}@eng.kagawa-u.ac.jp
[2]*NTT Advanced Technology Corporation, 19-18, Nakamachi, Musashino, Tokyo 180-0006, Japan*
hiromichi.kawano@ntt-at.co.jp
[3]*Tainan National University of the Arts, 66, Daci, Guantian, Tainan, 72045, Taiwan*

Abstract: This paper proposes a new color image arrangement method using an elastic transform on some kinds of axes. In this paper, we present the principle of our method using HMGD (Histogram Matching based on Gaussian Distribution). And we describe that the automated method applies the HMGD to input color image only when the image has single-peakdness in its histogram on the focused axis. We also show that the method gives a good Kansei effect in the case of applying the HMGD onto Lightness axis.

Keywords: Automated color arrangement, elastic transform, histogram matching, HMGD, Curvature

I. INTRODUCTION

Automated image processing for enhancement and/or arrangement of color images has been more familiar to us according to the spreading of Digital Camera, Smart Phone, DVD, etc. [1]-[3]. However, we consider that the research on the automated arrangement method that brings about good sensibility effect (or Kansei effect) is still on the way to practical use.

In this paper, we propose a novel color image arrangement method using an elastic transform based on histogram on some kinds of axes [4]-[6]. As for the axes, there are Lightness axis and principal component one that can be obtained by Principal Component Analysis (PCA) in the RGB three-dimensional vector space that is an attribute space of color image. Especially, we explain the main principle of the proposed method in the case of using HMGD (Histogram Matching based on Gaussian Distribution). And we illustrate that the method using the HMGD works very well for input image that has single peak histogram on the Lightness axis. Also in this paper, we show that a curvature computation for the cumulative histogram of original image will be effective, in order to automatically detect whether the original image has single peak histogram or not.

II. ELASTIC TRANSFORM BASED ON HISTOGRAM MATCHING

1. Principle

We describe the principle of histogram based elastic transform in the following. Let $f(x)$ and $f(y)$ be two probabilistic density functions on real variables x and y, respectively. The probabilistic density function (PDF) is corresponding to histogram of gray level image. However, the histogram is defined on discrete variable. In addition, let $y=\phi(x)$ be a continuous and monotonous increase function between variables x and y as shown in Fig.1.

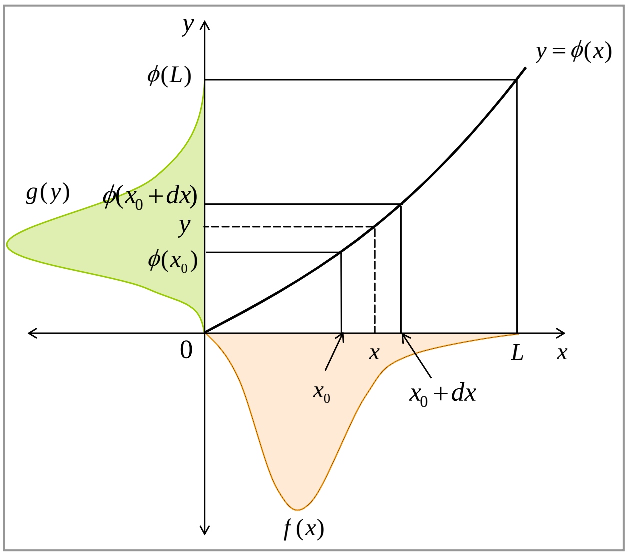


Fig.1. Continuous and monotonous increase function $y=\phi(x)$ and probabilistic density functions $f(x)$ and $g(y)$

In addition, let value of x be the range from 0 to L. Accordingly, variable y ranges from 0 to $\phi(L)$. Let P mean the probability. From the above definitions and Fig.1, we have Equation (1) ~ (3).

$$P(0 \le x \le L) = \int_{x=0}^{x=L} f(x)dx = 1 \qquad (1)$$

$$P(0 \le y \le \phi(L)) = \int_{y=0}^{y=\phi(L)} g(y)dy = 1 \qquad (2)$$

$$\begin{aligned} & f(x)dx \\ & = P(x_0 \le x \le x_0 + dx) \\ & = P(\phi(x_0) \le y \le \phi(x_0 + dx)) \\ & = P(y_0 \le y \le y_0 + dy) = g(y)dy \end{aligned} \qquad (3)$$

From Equation (3), we obtain Equation (4) and (5) because $y_0 = f(x_0)$ and $y_0 + dx = \phi(x_0 + dx)$.

$$f(x)dx = g(y)dy = g(y)\phi'(x)dx \qquad (4)$$

$$f(x) = g(y)\phi'(x) \qquad (5)$$

Thus, if we know the $y=\phi(x)$ and $g(y)$, then we have the $f(x)$. Using the above equations, we derive the principle of HE. Let $\phi(x)$ be defined by Equation (6).

$$\phi(x) = L\int_0^x f(x)dx \qquad (6)$$

Since $\phi'(x) = L\phi(x)$, according to Equation (5), we derive following Equation (7).

$$f(x) = g(y)Lf(x) \qquad (7)$$

Therefore, we understand that if we take the transform function as Equation (6), $g(y)$ becomes uniform distribution. It corresponds to the HE processing, which means that function defined by cumulative histogram transforms the original histogram into the uniform one.

Inversely, if we define the transform function $\phi(x)$ as an integral of desired PDF $f(x)$ for example, Gaussian distribution, we can obtain the desired PDF using the $\phi(x)$ and the uniform distribution such as Equation (8) (Fig.2).

The abovementioned theory means that, if we combine the both transform, we can obtain the transformation from an original distribution (PDF) to a desired one. This means that an image with original histogram can be transformed into another image with desired histogram. We consider that it is the principle of the Histogram Matching (HM) [4].

2. Elastic Transform on Axis

We can choose the abovementioned transform such as HE (Histogram Equalization) and HM (Histogram Matching) on arbitrary axis (for example, principal component axis) in the color attribute (RGB) space as

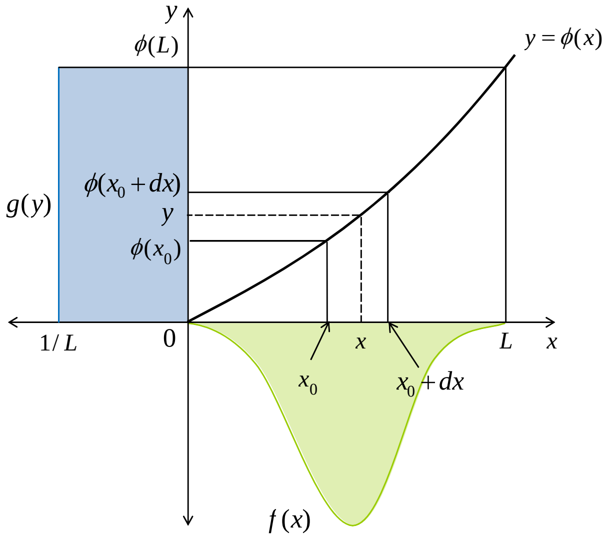


Fig.2. Conceptual image of the transform from unif orm distribution (PDF) to the desired one (PDF)

shown in Fig.3. For example, in the case where the HE processing on the Lightness axis is applied, the HE brings about image enhancement by contrast stretching. In addition, in the case of HE on a principal component axis in RGB space, we guess that the contrast stretching will be done along to a certain tone of color. Fig.4 shows examples of the HE on Lightness axis and PC (Principal Component) axis. Fig.5 shows resultant examples of the HE and HMGD (Histogram Matching based on Gaussian Distribution) on Lightness axis. The HMGD processing shows a more moderate effect than HE.

In this paper, we show that the HMGD is useful for the automated color image arrangement.

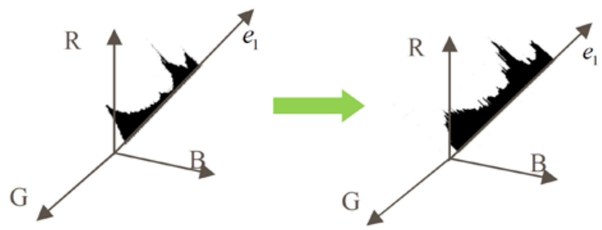


Fig.3. Conceptual image of histogram based elastic transform on arbitrary axis in RGB space.

(a) Original (b) Lightness axis (c) PC axis

Fig.4. Examples of HE on Lightness and PC axis

(a) Original (b) HE (c) HMGD

Fig.5. Example of the results by HE and HMGD

3. Detection of Histogram Peakedness by using Curvature Computation

In this section, we describe how we detect the histogram peakedness by using curvature. Let *y* be a function with respect to *x*, the definition of the curvature *R* is given by Equation (9).

$$R = \frac{\left(\frac{d^2y}{dx^2}\right)}{\left\{1+\left(\frac{dy}{dx}\right)^2\right\}^{\frac{3}{2}}} \tag{9}$$

Let $g(x)$ be a Gauss density function with variance σ^2 and average a. And also let $g(x)$ be representing a histogram of input image whose pixel values from 0 to L. That is,

$$g(x) = \frac{K}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}} \tag{10}$$

In Equation (10), K means a coefficient that satisfies following Equation (11).

$$\frac{K}{\sigma\sqrt{2\pi}} \int_0^L e^{-\frac{(u-a)^2}{2\sigma^2}} du = 1 \tag{11}$$

Let y=f(x) be a function representing the cumulative histogram. Then $f(x)$ can be represented Equation (12). Since $y = f(x) = \int_0^x g(u)du$, it can be described as Equation (13).

$$f(x) = \int_0^x g(u)du = \frac{1}{\sigma\sqrt{2\pi}} \int_0^x e^{-\frac{(u-a)^2}{2\sigma^2}} du \tag{12}$$

$$\frac{dy}{dx} = g(x) = \frac{K}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}} \tag{13}$$

By the same way,

$$\frac{d^2y}{dx^2} = \frac{dg(x)}{dx} = \frac{K}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}} \cdot \left(-\frac{1}{2\sigma^2}\right)\{2(x-a)\}$$

$$= \left(\frac{K}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}}\right) \cdot \left(-\frac{(x-a)}{\sigma^2}\right) = g(x) \cdot \left(\frac{(a-x)}{\sigma^2}\right) \tag{14}$$

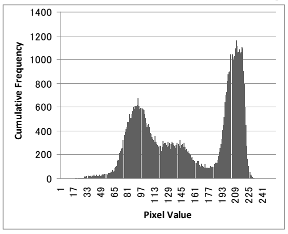
Hence, we obtain the following equation (15) and we can approximate to Equation (9).

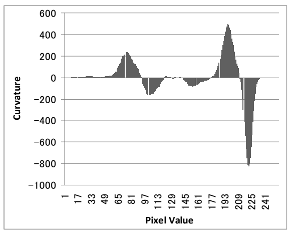
$$R = \frac{\frac{(a-x)}{\sigma^2} g(x)}{\left\{1+g(x)^2\right\}^{\frac{3}{2}}} \cong \frac{(a-x)}{\sigma^2} g(x) \tag{15}$$

The curvature R varies the sign according to the value of x; (x < a) R>0, (x=a) R=0, (x>a) R<0. So we consider that a histogram with single peak shows the same characteristic regarding the curvature *R* of the cumulative histogram.

Fig.6 shows an example of the curvature computation. The real curvature computation is performed on the discrete values. We can find that there are largely changing points in the curvature from plus value to minus value at more than one point. So, we easily understand that the histogram of the original image is not a single peak type.

(a) Original image





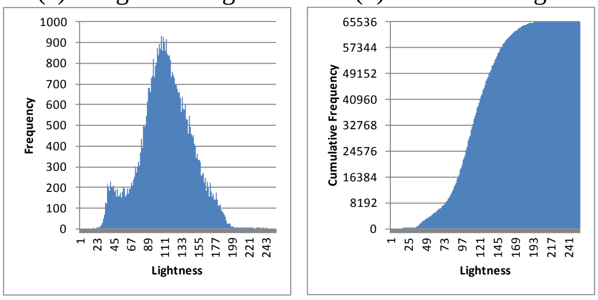
(b) Histogram of image (c) Curvature of image

Fig.6. Example of the curvature computation

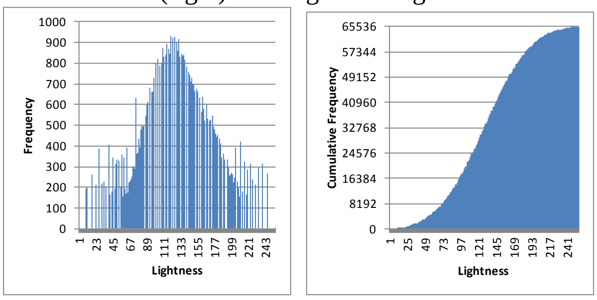
III. EXPERIMENTATION

Fig.7(a), (b) shows an example of results by HMGD processing on Lightness axis, where original image has almost single peak histogram. The set of histogram and cumulative histogram (original image and HMGD image) are shown in Fig.7(c), (d). And the curvature of cumulative histogram of input image is shown as well in Fig.7(e). Similarly, Fig.8 also shows that input image has single peak histogram and that the cumulative histo-

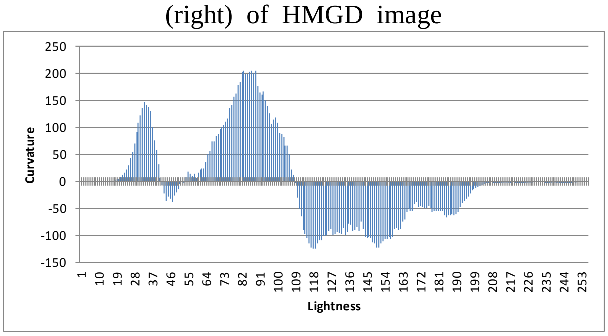
(a) Original image (b) HMGD image



(c) Histogram (left) and cumulative histogram (right) of original image



(d) Histogram (left) and cumulative histogram (right) of HMGD image



(e) Curvature of cumulative histogram of original image

Fig.7. Example of the results by HMGD

(a) Original image (b) HMGD image

Fig.8. Example of the results by HMGD

gram of the output image clearly approaches to the cumulative Gaussian distribution. For these cases where the histogram of input image has almost single peak, the HMGD works well for resulting a good feeling impression (or Kansei effect).

IV. CONCLUSION

Aiming at automated affective color image arrangement, we have proposed a concept of Elastic Transform (ET) method based on histogram and/or histogram matching, which can be applied to some kinds of axis such as Lightness axis. And we have suggested that HMGD (Histogram Matching based on Gaussian Distribution) is regarded as one of the ET method.

As for the concrete method for automated color image arrangement, we have proposed a method that applies HMGD processing to input image after detecting whether the input image has single peak histogram or not. In this paper, we have also provided the curvature computation for the peak detection in histogram. And also we have shown some experimental results. We consider that HMGD processing method after histogram peaks detection based on the number of curvature change points will be useful and promising.

REFERENCES

[1] Gonzalez CR, Woods ER, Digital Image Processing (1993). Addison-Wesley Publishing Company

[2] Jahne B, Digital Image Processing --Concepts, Algorithms, and Scientific Applications -- 4th Edition (1995). Springer

[3] Umbaugh ES, Computer Vision and Image Processing: A Practical Approach Using CVIPtools (1998). Prentice Hall PTR

[4] Burger W, Burge JM, Principles of Digital Image Processing: Fundamental Techniques (2009). Springer

[5] Izumi T, Hattori T, Sugimono S, Takashima T (2009), Color Image Arrangement Using Elastic Transform on Principal Component Axis (in Japanese). Journal of Japan Society of Kansei Engineering, Vol.8, No.3:667-674

[6] Izumi T, Hattori T, Fujita E, Sugimono S, Kawano H (2010), Feeling Impression and Quantities Accompanying Calculation of Fluctuation in Color Image (in Japanese). Journal of Japan Society of Kansei Engineering, Vol.9, No.2:243-250

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Novel Transistor by Opt-coupling of LED and PD and its Application

Junichi Fujita, Daisuke Sato, Kensho Okamoto, Tetsuo Hattori

Department of Electronics and Information Engineering, Kagawa University, 2217-20 Hayashi-cho, Takamatsu, Kagawa, Japan
Tel : 81-87-864-2292; Fax : 81-87-864-2262
s12d451@stmail.eng.kagawa-u.ac.jp, {oka, hattori}@eng.kagawa-u.ac.jp

Abstract: This paper presents a simple transistor-like amplification function by the optical coupling of LED (Light Emitting Diode) and Photodiode where the Photodiode current by photo electromotive force is feed backed into the LED, in the same way as the case of common emitter circuit of the conventional Bipolar Junction Transistor. Further-more, we will introduce the fact that the novel transistor becomes thyristor-like switching device under the co ndition.
Keywords: Transistor, Thyristor, Distar, Diristar, Opt-coupling, LED, Photodiode

I. INTRODUCTION

It is well known that the bipolar transistor is a semiconductor device capable of amplifying a small electric signal to a significant output level, e.g. of volt-order, in the case of analogue application. However, as a matter of facts, the transistor itself generates neither a high voltage nor large current. This might be misunderstand in other way. Practically, the bipolar transistor behaves as a variable resistance element whose resistance value between the collector and the emitter can be controlled widely from near zero ohm to almost infinity ohm by a small input current, namely the base current I_B, passed from the base to the emitter. In an common emitter transistor circuit, the collector is connected to a resistor R_C, and then a constant DC voltage V_{CC}, more than several volts, is applied between R_C and the emitter. Under such a condition, if I_B is passed from the base to the emitter then R_{CE}, which is the internal resistance between the collector and the emitter, will changes drastically. This causes the changes of the collector current I_C and also V_{CE}, the output voltage between C and E.

Prof. Kensho Okamoto noticed the behavior of the bipolar junction transistor as an electronically controllable resistor and therefore he thought that it might be possible to realize a completely new amplifier device by a system combining LED and silicon photodiode, or silicon solar cell, with positive feedback manner. He named this device "distar".

Further-more, we will introduce the fact that the novel transistor becomes thyristor-like switching device under the condition. Okamoto named this device "diristar".

According to our literature search, there are only two similar researches in the world other than us ([1][2]). However no practical devices were developed in their works.

II. TRANSISTOR BY PHOTO-COUPLING

The bipolar junction transistor was invented in 1948. There are two types for the bipolar transistor, i.e. PNP and NPN transistors. Since the birth of transistor, the operation of such bipolar transistors have been explained by the behavior of electrons and positive holes in the emitter(E), base(B) and collector(C) regions as illustrated in Fig. 1 (a). Okamoto found that the well-known common base transistor circuit shown in Fig. 1 (a) could be realized by putting together an infrared light-emitting diode (IR-LED) and a Si photodiode (Si-PD) as shown in Fig. 1 (b).

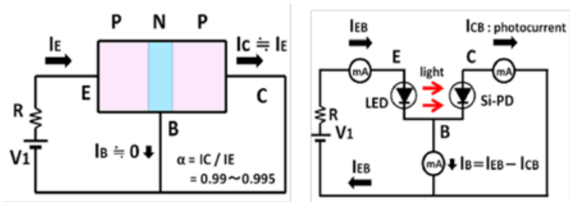
Then Okamoto made the quasi emitter common transistor circuit using IR-LED and Si-PD as shown in Fig. 2 (b) in order to confirm whether the circuit operated as the real transistor amplifier shown in Fig. 2 (a). Fig. 3 (a) shows an example of the device which consists of an IR-LED and Si-PD. The novel device was named "distar". Fig. 3 (b) shows picture of actual distar module.

Amplification factor α, which is defined as I_C/I_E shown in Fig. 3(b), our first distar was about 0.34. However, later we obtained α value more than 0.9 in a best combination of LED and Si-PD. This is a surprising matter because $\alpha \fallingdotseq 0.99$ is conventional transistor. According to the transistor theory, if α=0.9 then it is possible to build an emitter-common transistor with β=9. This means the LED-photodiode system can be used as an amplifier similar to the bipolar transistor.

The result is shown in Fig. 4. It is known from Fig. 4 that the quasi emitter common circuit using the distar device works as a practical voltage amplifier. We can obtain a voltage gain more than 40dB by adjusting each value of RC, RB and V. However it is difficult to manufacture a power or current amplifier using the distar. Because, a normal photodiode does not generate photocurrent as large as e.g. a 100 mA, which is required to drive an 8 ohm speaker.

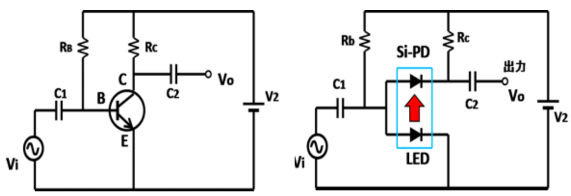
So, we used a Si solar cell module instead of a Si photodiode. Thus, we succeeded in the development of a distar power (or main) amplifier.

Fig. 5 is an audio pre-main amplifier circuit using two sets of distars, and Fig. 6 is a hand-made practical distar audio amplifier. We can enjoy beautiful and noise-free music sound using the audio system [3].



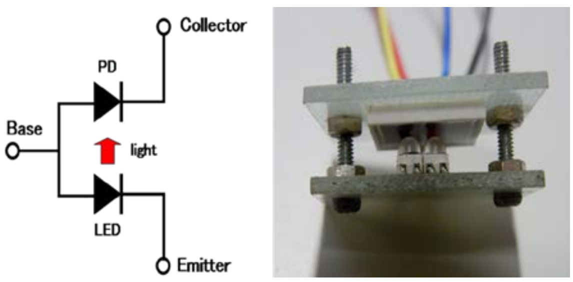
(a) Common base transistor (b) its simulation by LED and Si photodiode.

Fig.1. Description of amplification behavior.



(a) Common emitter circuit (b) quasi common emitter circuit

Fig.2. Amplification circuits.



(a) Composition of distar (b) Exterior view of distar

Fig.3. The distar consist of Infrared LED and Si-PD.

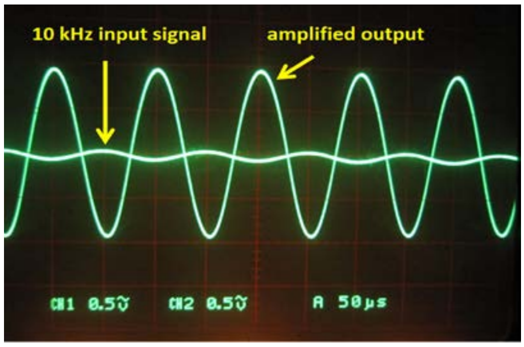


Fig.4.Amplification waveform using distar in Fig.3..

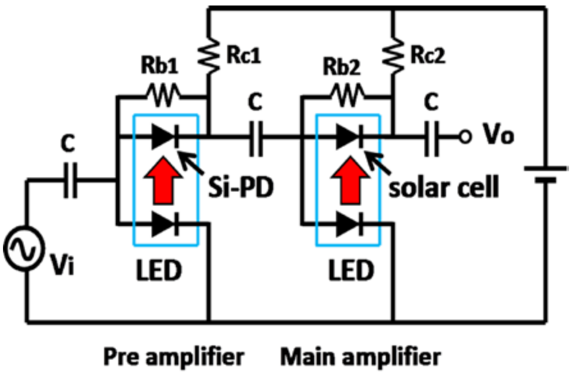


Fig.5. Description of amplification behavior.

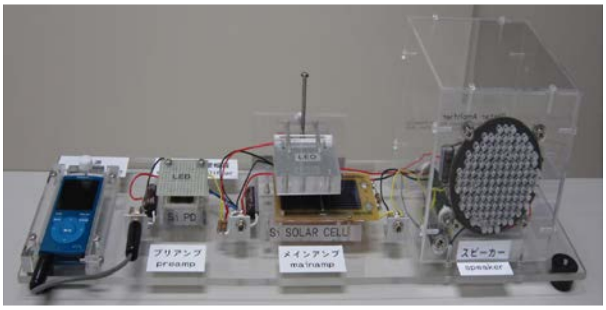


Fig.6. Description of amplification behavior.

III. AMPLIFICATION FUNCTION

We will explain about the amplification principle of the distar in this section. In the circuit of Fig. 7, IC does not flow because PD is connected in the reverse direction against voltage source V. therefore,

$$I_C = I_E = 0 \quad \text{(when IB = 0)} \qquad (1)$$

As shown in Fig. 8, Base-current IB flows through the LED connected between B to E when the switch SW2 is closed. Hence

$$I_E = I_B \quad \text{(when IB flows)} \tag{2}$$

Then, the LED emits light and the LED light enters into the PD, because the LED and the PD is optically faced each other. Consequently, photovoltaic effect occurs in PD and the photocurrent IC generates and flows through RC and the PD. The α is defined as the ratio of the current flowing through the PD and the current owing through the LED as shown in Eq. (3).

$$\alpha = \frac{I_C}{I_E} \tag{3}$$

Therefore, IC is expressed as Eq. (4).

$$I_C = \alpha \times I_E \tag{4}$$

At this time, IE = IB as shown in Eq. (2), Eq. (4) can be expressed as Eq. (5).

$$I_C = \alpha \times I_B \tag{5}$$

At the next stage, the current flows through LED, i.e. IE will increase, because IE becomes IE = IB +α IC = IB + α IB increased IB. And IC increases. Accordingly, to as Eq.(6), the light emission intensity of the LED also increases.

$$I_C = \alpha \cdot I_B + \alpha(\alpha \cdot I_B) \tag{6}$$

Each value of IE and IC are enhanced in chain reacting manner and finally IC becomes

$$I_C = \alpha \cdot I_B + \alpha(\alpha \cdot I_B) + \alpha(\alpha^2 \cdot I_B) + \cdots \tag{7}$$

$$I_C = I_B\left(\alpha + \alpha^2 + \alpha^3 + \cdots\right) \tag{8}$$

In equation (8), $(\alpha+\alpha^2+\alpha^3+\cdots)$ is a well-known the infinite progression and it sum will be convert to Eq. (9) when α is $0 \leqq \alpha < 1$,

$$\left(\alpha + \alpha^2 + \alpha^3 + \cdots\right) = \frac{\alpha}{(1-\alpha)} \tag{9}$$

Eventually, IC becomes as follows.

$$I_C = \frac{\alpha I_B}{(1-\alpha)} \tag{10}$$

Ratio of IB and IC is as shown in Eq. (11).

$$\beta = \frac{I_C}{I_B} = \frac{\alpha}{(1-\alpha)} \tag{11}$$

In Eq. (10), if α = 0.5 then β = 0.5. However, if α = 0.9 then β = 9. As α exceeds 0.9, β enhances greatly e.g. α = 0.99, β = 99.

On other hand, in case of $\alpha \geq 1$,

$$I_C = +\infty \tag{12}$$

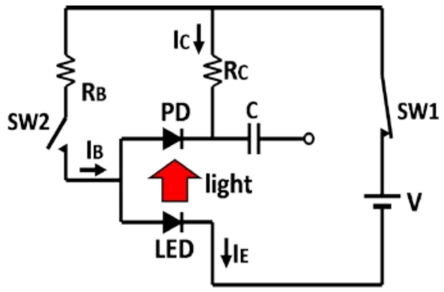


Fig.7. Description of amplification principle 1.

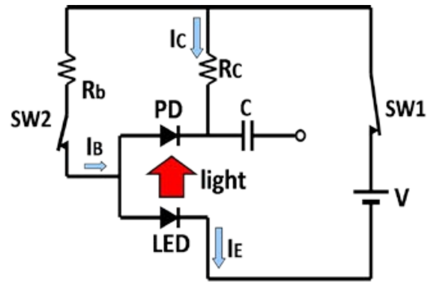


Fig.8. Description of amplification principle 2.

IV. VARIATION

A. Other Combination

In Fig. 9, the measurement circuit of quasi common base circuit using distar. The light intensity of LED and photocurrent of PD are very important in distar and diristar. For this reason, we have been looking for the best conditions only infrared LED. Because, light emission wavelength of the LED (940nm) and light reception of the PD (960nm) is almost identical as shown in Fig. 10.

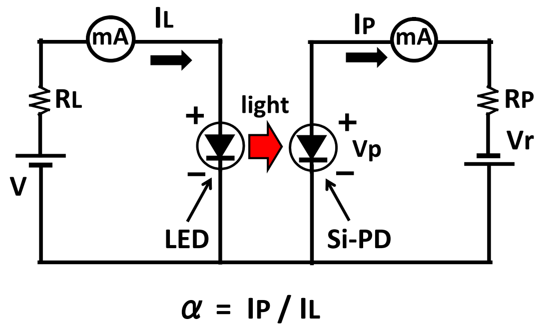


Fig.9. quasi common base circuit using distar.

However, we made an distar using high luminous efficiency red LED as shown in Fig. 11. Then, the electrical characteristics of this distar was better than we thought.

The amplification factor α using distar consists of IR-LED and PD as shown in Fig. 12. The α significantly varies depending on the LED current ILED when ILED less than about 200mA.

Fig. 13 shows the α using distar consists of red-LED and PD. It is stable in larger than a few 10 mA.

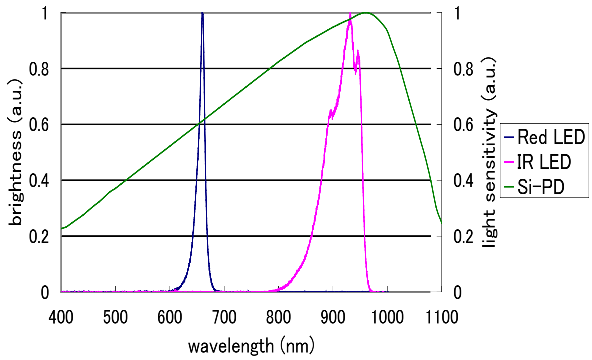


Fig.10. Spectrum of light emission and light reception.

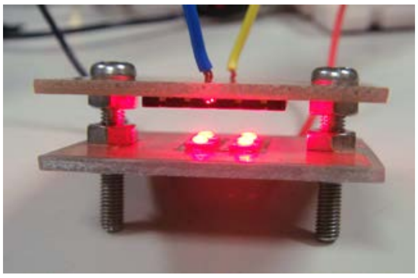


Fig.11. The distar that consists of Red LED and Si-PD.

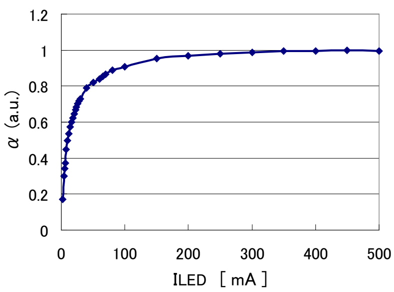


Fig.12. The α using distar consist of IR-LED and Si-PD.

B. Operational Amplifier

Now, we are challenging to make operational amplifier using distar module as shown in Fig.14 ([4]). And, we have confirmed that the Amplifier works as well as the conventional operational amplifier.

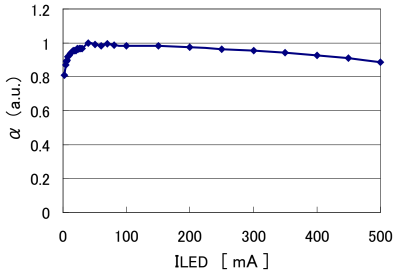


Fig.13. The α using distar consist of Red LED and Si-PD.

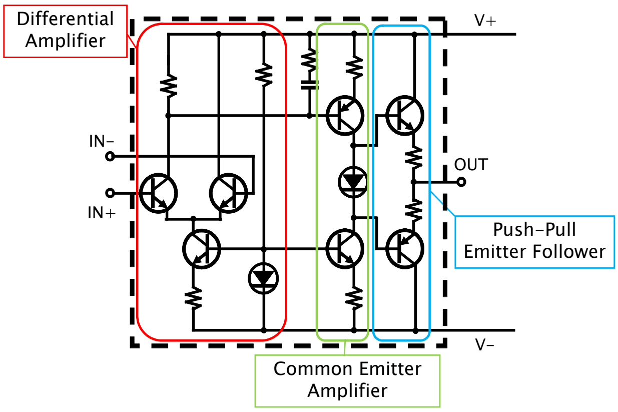


Fig.14. Equivalent circuit of the Operational Amplifier.

V. CONCLUSION

We have presented a novel transistor by opt-coupling of LED and PD. There are still rooms for improvement in realizing efficient transistors by the coupling. However, very practical and high power transistor such as thyristor can be built by the combination of large scale LED arrays and solar cell. We consider that, not only the power transistors but also switching devices based on the opt-coupling of LED and PD, are very promising in the future.

REFERENCES

[1] Van Zeghbroeck B.J. (1989), Photon transport transistor: Electron Devices Meeting, pp.543-546.
[2] W.N. Cheung (1996), IEEE Journal of Quantum Electronics{32} No.3, pp.502-506.
[3] J. Fujita (2013), Novel Audio Amplifier System Using Opt-coupling of LED and PD, Proceedings of 2013 International Conference on Biometrics and Kansei Engineering (ICBAKE2013). pp.301-304, Tokyo, Japan.
[4] M. Suzuki (1991), Authentic book Design of transistor circuit , JAN9784789830485, P294.

Study on the pose under complex multiple targets environment for the ind ustrial robot based on machine vision

Jiwu Wang [1], Xianwen Zhang [1], Huazhe Dou [1], Sugisaka Masanori [2]

[1] Department of Mechanical and Electrical Engineering, Beijing Jiaotong University, Beijing, China, 100044
[2] Nippon Bunri University, Oita 870-0397, Japan

[1]jwwang@bjtu.edu.cn
[2] ms@alife-robotics.co.jp

Abstract: Effectively obtaining multiple targets pose is the key element to improve flexible applications of industrial robots under complicated working environments. Machine vision technology is the main method to solve this problem. In order to effectively reduce influence of changing conditions (such as the illumination, angle and scale etc.) on target pattern recognition, the fusion algorithms of scale invariant feature transform (SIFT) and geometric moment, is developed in this paper. To achieve accurate identification of multi-targets' pose when their position and orientation change in the two-dimensional space, the optical-section method is used to determine the depth information of some feature points of the target. Moreover, that data is used to correct the pose recognized by the traditional pattern technology. Finally, using SCARA-type industrial robots as a platform, the relative algorithms are verified with corresponding experiments.

Keywords: machine vision; industrial robot; target recognition; image processing**.**

I. INTRODUCTION

Intelligent industrial robot is concentration and depth of integration of advanced manufacturing technology, information technology and smart technology, is the forefront of high-end equipment and manufacturing bases in the future, but also important symbol of a country's industrial level[1] Industry of the future will require a lot of robots and automated equipment to support, robots will become a huge industry, it makes fundamental changes in traditional patterns of industrial production, have a profound impact on the development of human society.

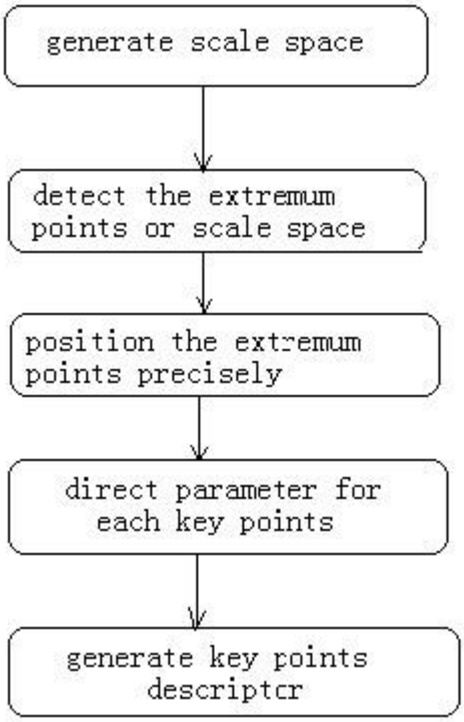
At present, most of the traditional industrial robots pick up and process according to fixed path, through prior programming.The initial pose and end pose of an object is predetermined, robots operate only according to the regular procedure.For industrial robots, work on the automated production line or a flexible manufacturing system[2] the largest group of action is "grab and place".In order to accomplish this type of operation ,robots require to accurate positioning information of manipulated objects: 1) pose before being manipulated must be known, to ensure the robot to accurately grab; 2) pose after being manipulated must be known, to ensure the robots to complete task accurately. If robot want to work stably and effectively ,the premise is keeping this two pose stable and not changed. But in actual work conditions, especially in line of occasions, it is very difficult to guarantee the pose of workpiece not changed. Inevitably, deviation between actual target object and ideal target object exists, so robot with flexible to meet this deviation is required.

Increasing industrial robot's flexible and improving its repeatability is one of the problem to be solved urgently for industrial robot development[3] Having a certain understanding of their environment is the premise for robots possess the necessary autonomy, so additional sensors to improve the perception on the environment is required. In this context, vision, proximity, touch, and force plays a major role, in which machine vision is considered to be most important sensory abilities. Object recognition and position by introducing vision can help robot grab and place according with wishes, improve production flexibility.

II. Study on the fusion algorithm of SIFT and moment invariants

SIFT algorithm (Scale-invariant feature transform) can determine the multi-target effectively, but the reference point generated in image processing is uncertain sometimes during the target recognition[4]. Although does not affect the target recognition, it is easy to generate unstable results. So SIFT algorithm alone can not effectively determine the target position and orientation.

SIFT algorithm steps as the following figure 1 shows.



Moment invariants algorithm is mainly aimed at the geometric characteristics, the selected reference point of the target is established before recognition[5]Moment invariants algorithm entirely rely on the target gray value ,which is obviously affected by the external light and other factors.

The moment invariant definition of image f(x,y) is as formula(1) shows. From formula (1) we can achieve the formula of the formula for invariant image displacement, as formula (2) shows to us.

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x,y) dxdy \quad (1)$$

$$\mu_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q \quad (2)$$

In order to fully utilize the advantages of the both algorithms ,firstly using SIFT algorithm to determine the identified target, then determine the area of geometric moment characteristics calculation in the captured images directly, according to the geometrical characteristics of the target[6] Determining the target pose accurately an effectively by improving conditions to meet the geometric moment calculation.

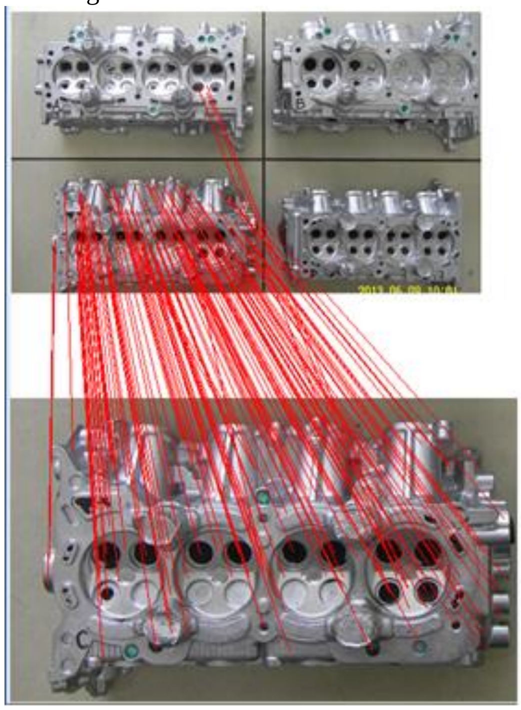


Fig.2. Target determination by using SIFT

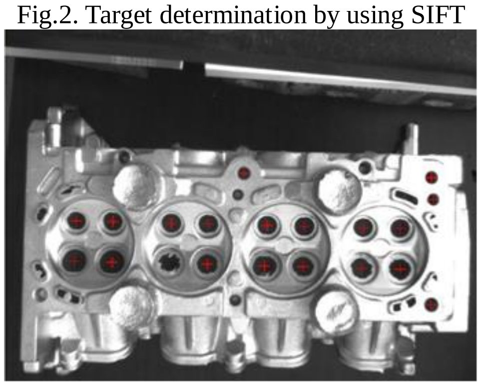


Fig. 3. Circle center recognition by using moment invariants

Figure 1 is an application example of SIFT algorithm. By using SIFT algorithm, we can determine the recognized workpiece is class C from (A、B、C、D) the four classes workpieces. Figure 2 using moment invariants algorithm to determine the circle center of the workpiece. The results show that we can achieve multi-target recognition and pose calculation effectively by using this method.

III. Study on the fusion algorithm of optical section and pattern recognition

In the process of pattern recognition, in many cases, it is very difficult to achieve judgment of the target effectively just rely on the characteristics of two-dimensional information, so the target depth information is required. At the same time, target depth information combine with two dimensional geometry, improve the accuracy of target distinction.

Considering the change of the target pose in three-dimensional space, only rely on geometric moment method to determine the key points pose, the result is not accurate, mainly due to changes in the height direction could affect the computation of the pose. If we are to achieve the precise description of the position and orientation, height information change is essential to be reflected in the calculation of the original pose.

Concrete solution is determining the angle change in height direction through optical-section method[7] then use that angle to correct the geometric calculated pose. The illustration of optical section as the following figure 3 shows, the optical band reflect the depth information of the target.

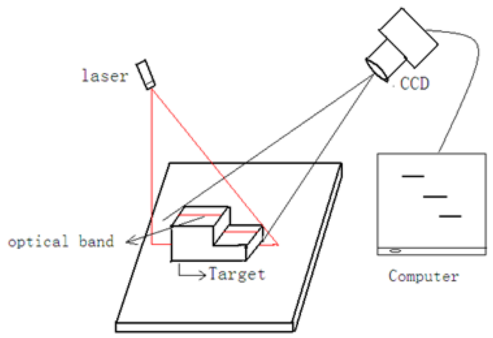


Fig. 4. Illustration of optical section

IV. CONCLUSION

In this paper, the machine vision system of industrial robot is established, which provides the foundation insurance for follow-up recognition work. The recognition software system as the following figure 4 shows, which is programmed with VC++6.0 .

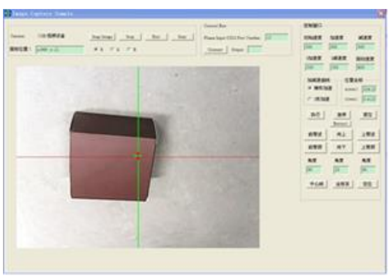


Fig.5. The industrial robot recognition system

In this paper, for pose change in the two-dimensional space, establish fusion algorithm of SIFT and Moment invariants, overcome 1) target rotation, scaling, translation; 2)image affine/projection transformation;3) illumination effect; 4) target occlusion; 5)clutter scenes;6) noise and other factors effect on the accuracy of target recognition[8] The integration of the both algorithms, meet the calculation condition of moment invariants effectively, improving the effectiveness of local features by using invariants, improving calculation robustness of the target pose.

iii. For pose change in three dimensional space, it is the first time to determine part depth information of captured target pictures through introducing optical section method, based on the traditional pattern recognition. Effectively complement the limit of only using two dimension data in traditional two dimension pattern recognition, enhance dimension and reliability of pattern recognition. Moreover, amending the corresponding angle for traditional calculation results.

REFERENCES

[1] Hui Qian,Zhiyong Yang,Tian Huang(2005),Key Technology to the Vision SoftWare Development of High-speed and Parallel Manipulator .10:34-37

[2] K. L. Zhuang, J. Z. Wang, and J. Zhou(2011), “Application of Machine Vision in Angle Inspection ,” Equipment Manufacturing Technology,vol. 4, pp. 4-10,

[3] Y. B. Wang, Y. Z. Hu, and M. Y. Lu(2006),Siemens machine vision system and its application in automobile engine assembly line, Automation systems engineering , vol. 1, pp. 20-22

[4] Fanyan Bu,Jieqing Tan(2011),Image matching approach using SIFT descriptor and image interpolation.47(16)156-158

[5] Xuhao He,Yueming Hu,Qifeng Qi(2008), Research of Geometric Moments Invariants in Identifying Chip.29(1)8-11

[6] Richard Szeliski(2012),Computer vision:algorithms and applications

[7] Xu,Min Tan,Yuan Li(2011),Visiual measurement an d control for robots

[8] Jianchang Zhang,Jianzhang Ma,Chengjun Ding (2008),Object recognition based on robot vision in complex environment.11(24)222-223

Research and Development of an Intelligent Robot of Medical Assistance Based on Embedded RTOS

Yuan Li, Fengzhi Dai, Junhong Xi, Binyao Li, Guodong You

College of Electronic Information and Automation, Tianjin University of Science and Technology, China
yuanli12139@gmail.com

Abstract: For practical purpose, this paper focuses on design and implementation of a wheeled intelligent robot that could be applied in hospitals or nursing homes for medical assistance. The hardware part includes two microcontrollers, MC9S12XS128 and STC89C52, assisted by a laptop as the control center of the system, various sensors and other robotic pieces. The software part is basically achieved on CodeWarrior 5.0, Keil μVision4 and Visual C++ 6.0 with OpenCV library, porting μC/OS-II real-time operating system to MC9S12XS128 in order for improving stability and flexibility of the system. The robot has functions of infrared remote control, fuzzy rule-based path tracking, obstacle avoidance, real-time LCD display and voice broadcast, color recognition and robot arm grasping.

Keywords: Robotics; Microcontroller; Embedded System; μC/OS-Ⅱ

I. INTRODUCTION

Artificial intelligence is a direction of the development of intelligent science [1]; robot and its correlation techniques has obtained a significant progress in recent years, especially in the field of Integrated Circuits and Sensors. In software aspects, Embedded Real-Time Operating Systems (RTOS) have been applied into plenty of intelligent devices such as mobile phone because of superior portability, stability, and efficiency. However, the development of intelligent robots for medical assistance based on these new technologies is relatively slow. In spite of many types of electric scooters or wheelchairs have been sold in market and used in hospitals and nursing homes, most of them has simple and single functions without adopting advanced intelligent control approaches [2, 3]. For this reason introducing Embedded RTOS to solve this problem is a necessary work, and a versatile intelligent robot for future medical assistance shown in Fig.1 was implemented in the research of this paper. The robot is designed a series of functions including manual remote control, independent path tracking with obstacle avoidance, and fundamental pattern recognition. The function of elevating under remote-control mode greatly increases the range the robot arm can touch [4]. But most importantly, in the software aspect of the robot, the application of μC/OS-ⅡEmbedded RTOS allows all the functions of the robot being coordinated in a more flexible and reasonable way. It deserves to believe that this new type of robot will make great change and improvement for our way and quality of life as long as it can be applied in hospitals, nursing homes or other civilian fields.

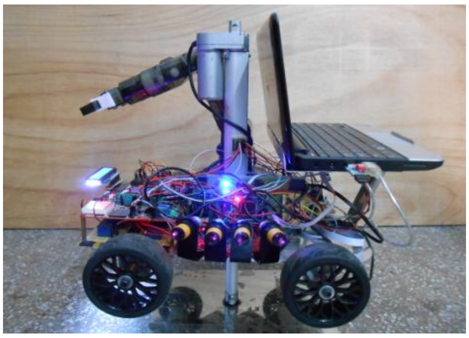


Fig. 1. Appearance of the robot

II. HARDWARES AND PRINCIPAL FUNCTIONS

Microcontroller MC9S12XS128 and STC89C52 cooperated with Dell Inspiron Mini10 laptop are adopted for accomplishing all the functions of the robot. XS128 ported μC/OS-Ⅱ Embedded RTOS is in charge of driving wheels, elevating, path tracking, obstacle avoidance, robot arm; C52 decodes infrared signal of remote control through external interruption, and controls LCD display by parallel I/O ports and voice broadcast by universal asynchronous receiver /transmitter; an interactive interface on the laptop based on MFC and OpenCV library is designed for real-time image capture, processing and color recognition. Two microcontrollers are linked by I/O ports of each. The laptop sends the result of color recognition by serial port communication to XS128.

The motors of wheels and elevating platform are driven by two MC33886 and one L298N respectively;

infrared remote control is achieved by HS0038 infrared receiver linked to C52 so as to control the motion of robot and switching automatic/manual mode; a gray scale sensor with seven LED of high brightness detects black path on the ground for path tracking function; five obstacle-avoiding sensors installed on the head and left edge enable the robot to avoid obstacles while tracking path; The robot arm owns five degree of freedom driven by two ULN2003A and ten relays; the LCD1602 installed on the head of the robot and the speech synthesis module based on SYN6288 can display and broadcast the robot's running status in real time.

III. PORTING μC/OS-Ⅱ TO MC9S12XS128

The main program code of XS128 is compiled and debugged upon CodeWarrior 5.0 IDE. The robot's functions of remote control, path tracking, obstacle avoidance and color recognition can be managed flexibly and efficiently under the environment of μC/OS-Ⅱ Embedded RTOS. But before that, μC/OS-Ⅱ should be successfully ported to microcontroller XS128.

CPU12X has a set of registers including a 16-bit program counter (PC), a 16-bit stack pointer register (SP), a 16-bit accumulator D, 16-bit index registers X and Y, a 16-bit condition code register, a global page (GPAGE) register and specific resource page (RPAGE, EPAGE and PPAGE) registers. As responding interruptions, the hardware of XS128 will automatically push the core registers onto the stack whose address grows downward [5].

Among μC/OS-Ⅱ files, define above hardware information including data types, stack data type, stack's growth direction, tasks' scheduling function OS_TASK_SW( ) and macro definition of critical section of code in the OS_CPU.H file; porting four assembly functions including OSStartHighRdy ( ), OSCtxSw ( ), OSIntCtxSw ( ) and OSTickISR ( ) in the OS_CPU_A.ASM file and six functions including OSTaskStkInit ( ), OSTaskCreateHook ( ), OSTaskDelHook ( ), OSTaskSwHook ( ), OSTaskStatHook ( ) and OSTaskTickHook ( ) in the OS_CPU_C.C file [6].

After porting three files mentioned above, make a system clock for the RTOS by using XS128's RTI module. For one thing, set up the system bus frequency to 32MHz.Then set related registers of RTI and add the time-tick interrupt vector to the file TBDML_linker.prm. Finally, include all header files into INCLUDES.H and configure and tailor the file OS_CFG.H based on needs, retaining the functions about task and message mailbox.

So far, the porting of μC/OS-Ⅱ is accomplished and the system is ready for programming task functions in the file MAIN.C.

IV. CONTROL PROGRAM BASED ON μC/OS-Ⅱ

After introducing Embedded RTOS that can run more than one task simultaneously to the robot software system, each program of the robot's function can be regarded as a task programming in the main file MAIN.C. So there are four tasks that need run in the μC/OS-Ⅱ environment. According to their priorities from high to low, these four tasks are: Task of Color Recognition whose priority is 6, Task of Obstacle Avoidance whose priority is 7, Task of Remote Control whose priority is 8 and Task of Path Tracking whose priority is 9. Fig. 2 clearly shows all the robot's tasks above and the way of switch between any two tasks of them.

1. Task of Color Recognition

Red/Blue color recognition is achieved by upper computer of color recognition shown in Fig. 3. The upper computer is programmed on Visual C++ 6.0 environment with MFC and OpenCV library. The digital camera collects HSV value of the objects within its field of view and the computer calculates the area of the objects that match the HSV range of red or blue. The result of recognition will be sent to XS128 by serial port for further processing.

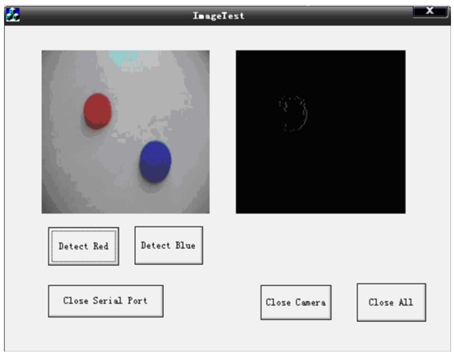


Fig. 3. Upper computer for color recognition

This task, when it first begins, will call the function msg2=(INT8U*)OSMboxPend(ExMbox2,0,&err2) to wait message from mailbox "ExMbox2". Then the system determines whether or not the message is what the task needs before executing following program; if the message suggests that the object is red or blue, the color is chosen by user, the system will order the robot arm to grasp the target object. Therefore, the task of color recognition is in suspended state as waiting message and it will not be woken up until the upper computer has detected the object that has certain color the user wants to look for.

2. Task of Obstacle Avoidance

The robot is designed to automatically bypass obstacles under path tracking mode[4]. When this task begins, it is suspended by calling function msg1=(INT8U*)OSMboxPend(ExMbox1,0,&err1) to wait the message from mailbox "ExMbox1" of path tracking task. Therefore, the task of obstacle avoidance is also in suspended state as waiting message and the task of remote control or path tracking is running. It will not be woken up until there is obstacle detected in front of the robot.

3. Task of Remote Control

This function is achieved by communication protocol NEC, namely detecting duration of high level at receiving ends and using the time interval between pulses to differentiate logic "0" and "1". An external interrupt service subroutine is designed in C52 for decoding of infrared signal. Each time after a decoding, the subroutine will return an array storing the result of decoding. Then C52 will give different orders according to different value in the array to relative modules to implement certain functions such as LCD display and voice broadcast.

C52 is also able to send certain remote control commands to XS128 through I/O ports of the two microcontrollers. Once XS128 receives the order of switching operational mode from C52 as the robot is running under remote control mode, the system will call function OSTaskSuspend(OS_PRIO_SELF) to suspend the current task and make a task switching, thereby the Task of Path Tracking whose priority is highest in ready state will acquire the access to CPU and start running. So the robot switches to path tracking mode.

4. Task of Path Tracking

The seven-channel gray scale sensor of the robot can output analog voltage signals ranging from 1.2V to 4.0V in terms of different grayscale values on the ground. Through the ATD channel of XS128, these analog signals are transformed to digital signals. Compare each digital value with predefined threshold in the robot system. If the digital value is greater than the threshold, it means that the corresponding channel of gray scale sensor has detected black path. In order to move more smoothly, Fuzzy Control Algorithm is adopted in the design of this robot.

Similar to the remote control task, the system can switch back to remote control mode once receiving the switching command from C52.But the different part is that the current path tracking task is re-suspended by calling function OSTaskResume(Task3Prio). Here, Task3Prio = 8, which signifies the priority of remote control task. Since at this time the suspended task of remote control resumes into ready state and its priority is higher than task of path tracking, the path tracking task is instantly deprived of its access to CPU, making the running mode of the robot switch back to remote control mode.

On top of that, the path tracking task can fulfill duty of sending message to start the task of obstacle avoidance and the task of color recognition by mailbox function of μC/OS-Ⅱ. Once the front obstacle-avoiding sensor detects out obstacle or XS128's SCI module has received color information from the laptop, it will call the function OSMboxPost(ExMbox1,(void*)&msg) or the function OSMboxPost(ExMbox2,(void*)&ch) to send messages.

V. CONCLUSION

This paper combines traditional robotic technology with Embedded RTOS, designing and implementing a new type of intelligent robot in the field of medical assistance. After practical debugging, the robot can implement many functions like remote control, path tracking, obstacle avoidance, color recognition, and real-time LCD display and voice broadcast, all of which can be switched flexibly by the application of μC/OS-ⅡEmbedded RTOS. Fig. 4 and Fig. 5 show the robot tracking the black path and avoiding obstacle. As Fig. 6 shows, the robot also can help user grasp the object of certain color by using digital camera and robot arm. This kind of robot could be applied into hospitals or nursing homes in future to take care of old people or patients with their life and lower the burden of nurses.

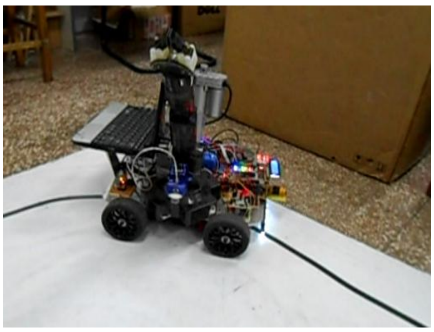


Fig. 4.The robot runs under path tracking mode

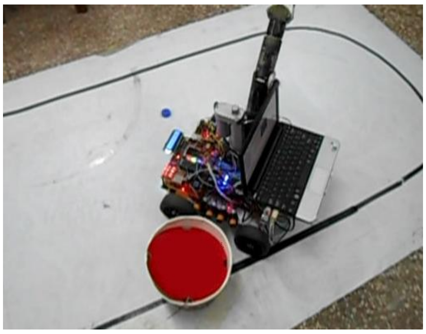


Fig. 5.The robot bypasses obstacle

Fig. 6.The robot recognizes and grasps target object

ACKNOWLEDGMENT

This paper was supported in part by the Tianjin Natural Science Foundation of China (No.13JCZDJC29100).

REFERENCES

[1] Shoshana L. Hardt, William J. Rapaport (1986), Recent and Current Artificial Intelligence, The AI Magazine Summer, pp. 91-100

[2] Zhaoxia Wang, Yongxin Liu (2010), Design of road tracing navigation control for smart car use CCD sensor, E-Health Networking, 2010 International Conference on Digital Ecosystems and Technologies,pp. 345-348

[3] Mochizuki Tatsuya, HayashiEiji (2007), Navigation system for a mobile robot using an omni-directional camera, 12th International Symposium on Artificial Life and Robotics,pp. 450-453

[4] Yuan Li, Chunyu Yu, Fengzhi Dai, et al (2012),Implementation of a Versatile Intelligent Vehicle with Tracking and Remote Control for Hospital or Nursing Homes, 3rdIEEE International Conference on Software Engineering and Service Sciences, pp. 505-508

[5] MC9S12XS256Reference Manual, Rev. 1.13 (2012), Freescale SemiconductorInc

[6] Jean J. Labrosse (2003), MicroC/OS-II: The Real-Time Kernel Second Edition, CMP Books

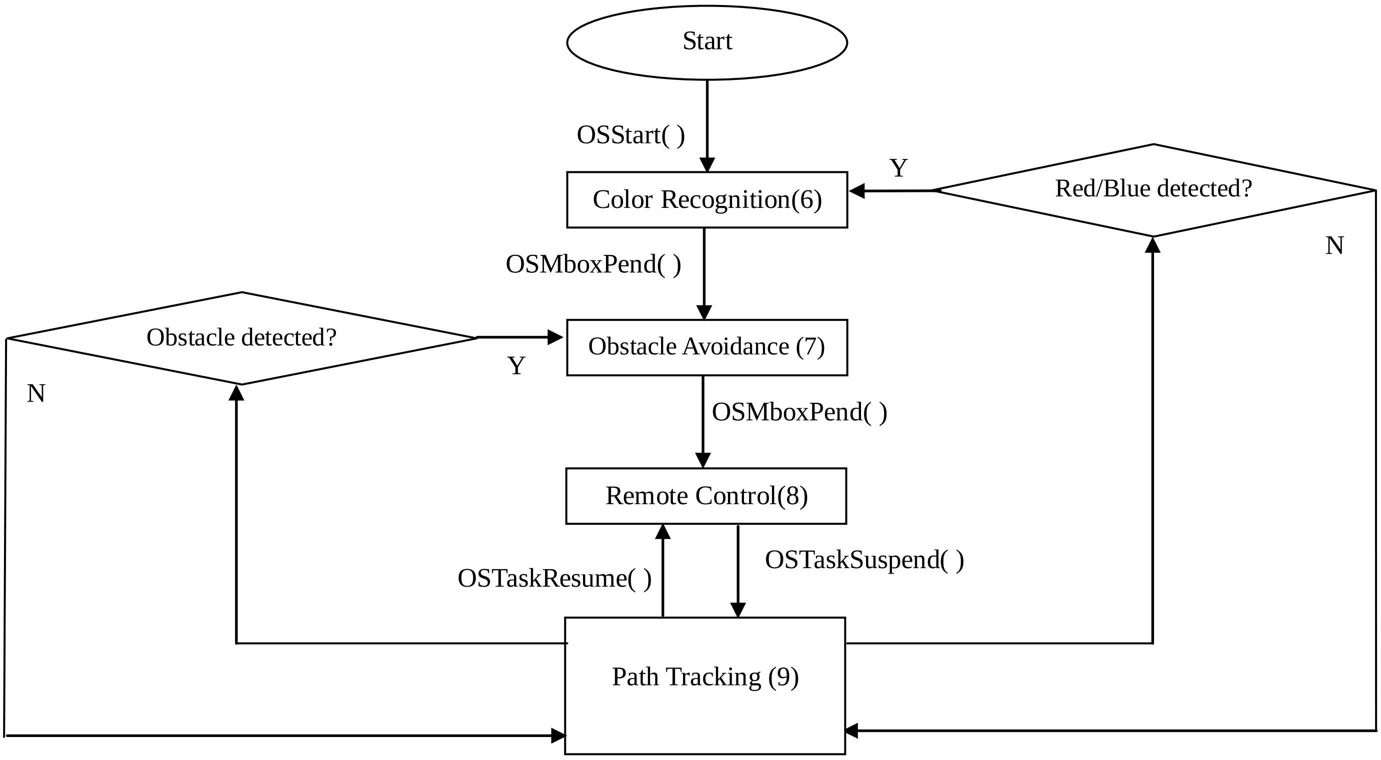


Fig. 2. The state transition diagram of the robot's tasks

High Acceleration Robotic Arm for Dynamic and Dexterous Manipulation of Deformable Object

Hiroaki Seki[1], Hiroki Shibata[1], Yoshitsugu Kamiya[1], Masatoshi Hikizu[1], Khairul Salleh Mohamed Sahari[2]

[1] *Faculty of Mechanical Engineering, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa, Japan*
[2] *Centre for Advanced Mechatronics and Robotics, College of Engineering, Universiti Tenaga Nasional, Jalan IKRAM-UNITEN, 43000, Kajang, Selangor, Malaysia*
Tel : 81-76-2344705 Fax : 81-76-2344706
{hseki, kamiya, hikizu}@t.kanazawa-u.ac.jp, khairuls@uniten.edu.my

Abstract: Deformable object manipulation is very important for home service robots. This paper discusses the development of a high acceleration robotic arm capable of performing dexterous manipulation of deformable object such as string. A mass-spring model of the string is first developed to study the design requirement of the robot as well as to determine the required motion of the robot to smartly manipulate the string. The developed system comprising of the string model and robotic arm is then used to manipulate unknown string. Based on some simple basic routine, the parameters of the string can be approximated based on the string model. Once the parameters are known, the string can be smartly manipulated by the robotic arm.

Keywords: high-acceleration robot, wire-driven robot, dexterous manipulation, home service robot, mass-spring system.

I. INTRODUCTION

Currently, development of home service robots has attracted a lot of attention from robotics researchers [1]-[6]. These robots are required to handle not only rigid objects, but soft and deformable objects as well. However, there are still a lot of research areas to be explored in the field of deformable object manipulation by robot. Some manipulation examples involving deformable objects include making a bed and clothes folding. In clothes manipulation, when a person manipulates clothes for folding, he or she will somehow perform dynamic manipulation of the clothes such as swinging the object to spread it etc. This paper discusses the development of a robotic arm with high acceleration that will enable more dexterous manipulation of clothes.

II. REQUIRED ACCELERATION

In order to design and develop a suitable manipulator, the big question is how much acceleration is required from the robot. A simple experiment is conducted to roughly estimate the acceleration required from the robotic manipulator. A human tosses an object of 18g in weight vertically into the air while a high-speed camera captures images of a marker placed at the person's wrist at 0.01 sec intervals (Figure 1 (left)). The distance d is then obtained based on the release point and also the highest point achieved by the object.

The experiment is repeated 10 times. Figure 1 (right) show the most distinguish results related to the path of the mark and the variation of the acceleration. In this case, the object was released at 0.19 sec and the height d was approximately 210 mm, with the highest acceleration of 22.3 m/s^2, recorded during the moment when the object is released.

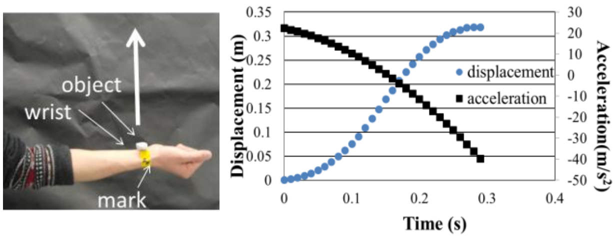


Figure 1: Finding approximate acceleration requirement

III. WIRE-DRIVEN MANIPULATOR

Most commercial robotic manipulators have their actuators on the links itself, which contributes significantly to the increase in weight. This is a major disadvantage in getting a high acceleration from the manipulators. Thus, in this paper, a robotic manipulator using wire driven concept is suggested to obtain high acceleration. Instead of having the manipulators on the links, the manipulators are placed outside the robot. as shown in Figure 2.

Based on this concept, the following robotic manipulator shown in Figure 3 has been designed and developed. The total weight of the robot (link 1 to link

3) is 1.3kg. The motor used are Maxon brushless DC motors, with potentiometer installed on all the rotational axes. Comparing our developed prototype with a Mitsubishi RV-2SQ, which is very similar in the length, it is lighter by more than 17kg. Table 1 shows comparison between our developed prototype with RV-2SQ in terms of the maximum angular velocity. Table 2 shows the detail specification of our developed prototype.

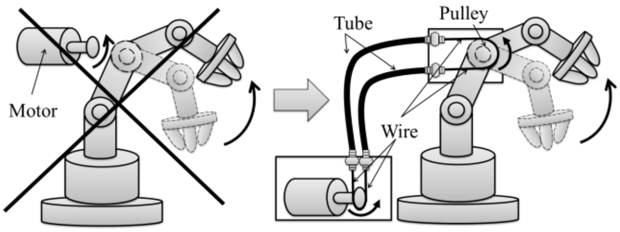


Figure 2: Concept of wire-driven manipulator

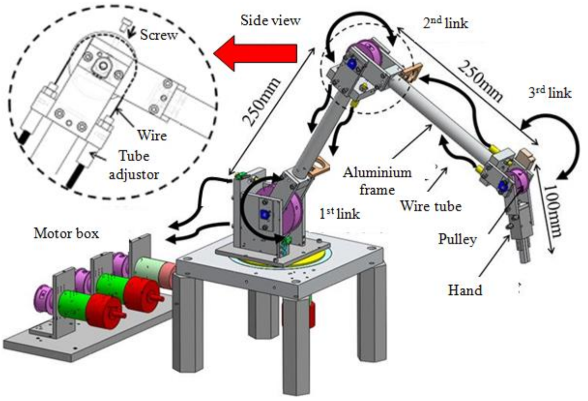


Figure 3: Developed High Acceleration Manipulator

Table 1: Maximum angular velocity comparison

	Our prototype [deg/sec]	RV-2SQ [deg/sec]
1st link	286	150
2nd link	458	275
3rd link	687	450

Table 2: Specifications of developed prototype

	1st link	2nd link	3rd link	Base
Limit [deg]	-20~+120	±120	±120	±160
Max ang vel [rad/s]	5	8	12	-
Max ang acc [rad/s^2]	20	50	50	-
M. of inertia [gmm^2]	6.82x10^7	1.29x10^7	1.78x10^5	6.92x10^7
Torque [mNm]	2587	1760	383	1963
Motor	EC-i40	EC-i40	EC-32flat	EC-i40
Motor driver	EPOS2 24/5	EPOS2 24/5	EPOS2 24/2	EPOS2 24/5
Gearhead	GP32C(66)	GP32C(66)	GP32A(23)	GP32C(23)
Speed ratio	97.02	66	23	73.6

The model of the prototype is shown in Figure 4. This is basically important to be embedded inside the simulator which also includes the mass-spring representation of the deformable object, which for the time being is only limited to string.

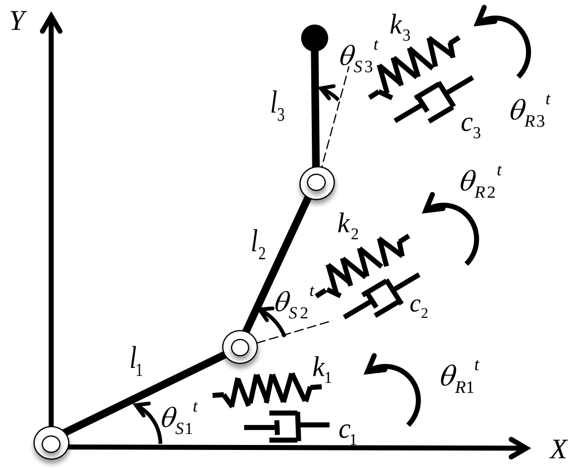


Figure 4: Model of robot prototype

From Figure 4, the following equations can be determined or established. The moment of inertia for each link can be determined as follows:

$$I_{gi} = \frac{m_i l_i^2}{12} \tag{1}$$

Subsequently,

$$\begin{aligned} J_1{}^t &= (I_{g1} + m_1 l^t{}_{c1}{}^2) + (I_{g2} + m_2 l^t{}_{c2}{}^2) + (I_{g3} + m_3 l^t{}_{c3}{}^2) \\ J_2{}^t &= (I_{g2} + m_2 l^t{}_{c2}{}^2) + (I_{g3} + m_3 l^t{}_{c3}{}^2) \\ J_3{}^t &= (I_{g3} + m_3 l^t{}_{c3}{}^2) \end{aligned} \tag{2}$$

Based on Newton's law, the torque can be determined as follows:

$$J_i{}^t \ddot{\theta}_{Ri} = -k_i(\theta_{Ri}{}^t - \theta_{Si}{}^t) - c_i(\dot{\theta}_{Ri}{}^t - \dot{\theta}_{Si}{}^t) \tag{3}$$

Subsequently,

$$\ddot{\theta}_{Ri} = \frac{\dot{\theta}_{Ri}{}^{t+1} - \dot{\theta}_{Ri}{}^t}{dt} \tag{4}$$

$$\dot{\theta}_{Ri}{}^{t+1} = \frac{\theta_{Ri}{}^{t+1} - \theta_{Ri}{}^t}{dt} \tag{5}$$

$$\dot{\theta}_{Ri}{}^{t+1} = \dot{\theta}_{Ri}{}^t + \left(-\frac{k_i}{J_i{}^t}(\theta_{Ri}{}^t - \theta_{Si}{}^t) - \frac{c_i}{J_i{}^t}(\dot{\theta}_{Ri}{}^t - \dot{\theta}_{Si}{}^t)\right) \cdot dt \tag{6}$$

$$\theta_{Ri}{}^{t+1} = \theta_{Ri}{}^t + \left(\dot{\theta}_{Ri}{}^t + \left(-\frac{k_i}{J_i{}^t}(\theta_{Ri}{}^t - \theta_{Si}{}^t) - \frac{c_i}{J_i{}^t}(\dot{\theta}_{Ri}{}^t - \dot{\theta}_{Si}{}^t)\right) \cdot dt\right) \cdot dt \tag{7}$$

IV. MASS-SPRING MODEL OF STRING

To validate the performance of the developed manipulator, and to enable dynamic manipulation on any unknown deformable object, a mass-spring system model of a simple string is developed. String is chosen due to its simplicity. Figure 5 shows the mass-spring system.

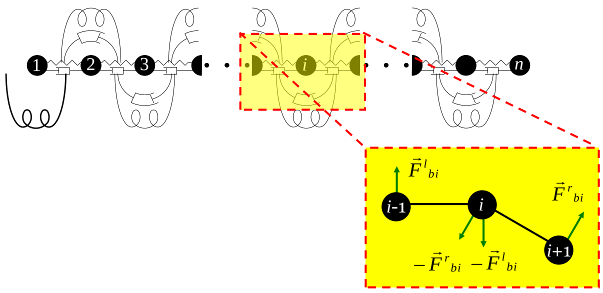


Figure 5: Mass-spring model of string

The system also considers air drag (F_{ci}) and gravitational force (F_g), besides the elastic force (F_{si}), damping force (F_{di}), elastic bending force (F_{bi}) and elastic damping force (F_{hi}) between the mass points. The equations related to the mass-spring model are shown as follow:

$$\frac{\vec{F}_{si}}{m} = k_s(\ell_i - L)\cdot\vec{d}_i \tag{8}$$

$$\frac{\vec{F}_{di}}{m} = \zeta_s\sqrt{2k_s}\left\{\left(\dot{\vec{r}}_{i+1} - \dot{\vec{r}}_i\right)\bullet\vec{d}_i\right\}\vec{d}_i \tag{9}$$

$$\frac{\vec{F}^{r}_{bi}}{m} = \frac{k_h\cdot\theta_i\cdot\left\{\vec{d}_i\times\left(\vec{d}_{i-1}\times\vec{d}_i\right)\right\}}{\ell_i} \tag{10}$$

$$\frac{\vec{F}^{l}_{bi}}{m} = \frac{k_h\cdot\theta_i\cdot\left\{\vec{d}_{i-1}\times\left(\vec{d}_{i-1}\times\vec{d}_i\right)\right\}}{\ell_{i-1}} \tag{11}$$

$$\frac{\vec{F}^{r}_{hi}}{m} = \frac{\zeta_h\cdot L\cdot\sqrt{k_h}\cdot\dot{\theta}_i\cdot\left\{\vec{d}_i\times\left(\vec{d}_{i-1}\times\vec{d}_i\right)\right\}}{\ell_i} \tag{12}$$

$$\frac{\vec{F}^{l}_{hi}}{m} = \frac{\zeta_h\cdot L\cdot\sqrt{k_h}\cdot\dot{\theta}_i\cdot\left\{\vec{d}_{i-1}\times\left(\vec{d}_{i-1}\times\vec{d}_i\right)\right\}}{\ell_{i-1}} \tag{13}$$

$$\frac{\vec{F}_{ci}}{m} = -C_{c1}\cdot\dot{\vec{r}}_i - C_{c2}\cdot\left|\dot{\vec{r}}_i\right|\cdot\dot{\vec{r}}_i \tag{14}$$

By applying Euler Explicit Method, the velocity and subsequently the position of the mass points can then be determined from the sum of the forces above.

$$\dot{\vec{r}}_i^{\,t_{h+1}} = \dot{\vec{r}}_i^{\,t_h} + \ddot{\vec{r}}_i^{\,t_h}\cdot dt \tag{15}$$

$$\vec{r}_i^{\,t_{h+1}} = \vec{r}_i^{\,t_h} + \left(\dot{\vec{r}}_i^{\,t_h} + \ddot{\vec{r}}_i^{\,t_h}\cdot dt\right)\cdot dt \tag{16}$$

where t_h is the current time step, and t_{h+1} is the next time step.

V. EXPERIMENTAL RESULTS

Experiments were conducted to check the performance of the developed manipulator. An unknown string with a 5mm diameter is selected as shown in Figure 6.

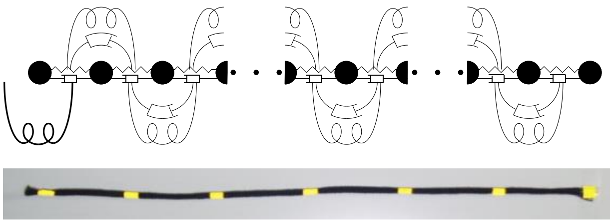


Figure 6: Unknown string used in experiments

Since this paper is intended to report on the development of the robot manipulator, experimental results regarding parameters determination of the unknown string is omitted.

Chosen experiment is to program the robot to make the string form a u-shape as shown in Figure 7. Assuming that point 0 is held by the robot gripper, the angle limitations for all mass points (70°≤θ_1≤110°, -40°≤θ_2≤-10°, -60°≤θ_3≤-30°, -60°≤θ_4≤-30°, -40°≤θ_5≤-10°, -110°≤θ_6≤-70°) plus distance between points 1 and 6 (x) are determined. 2 seconds after the robot manipulator stops its manipulation, and if the u-shape is achieved, the experiment is considered successful. The result is shown in Figures 8 to 10.

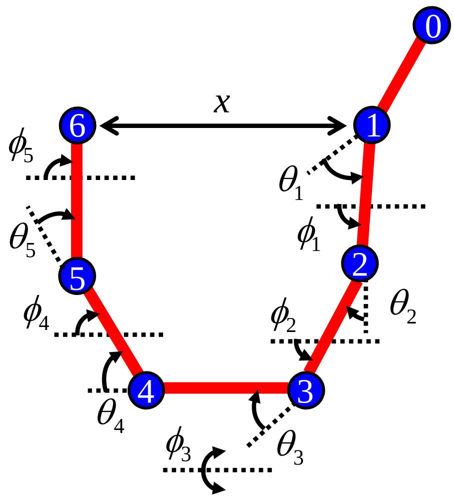


Figure 7: Forming string to become u-shape

The manipulator is trained by reinforced learning on how to smartly make a u-shape from the string. By 4th attempt, the manipulator can smartly perform the u-shape. Figure 8 shows the simulated Bezier curve projected for each of the link. Figure 9 shows the simulated and actual result of the string manipulation. Please note that camera EX-F1 with 60 fps capability (Casio) was used to capture the images of the string in real time. Figure 10 shows the result for the wrist movement of the manipulator. Figure 11 shows the scenes during the 3rd attempt.

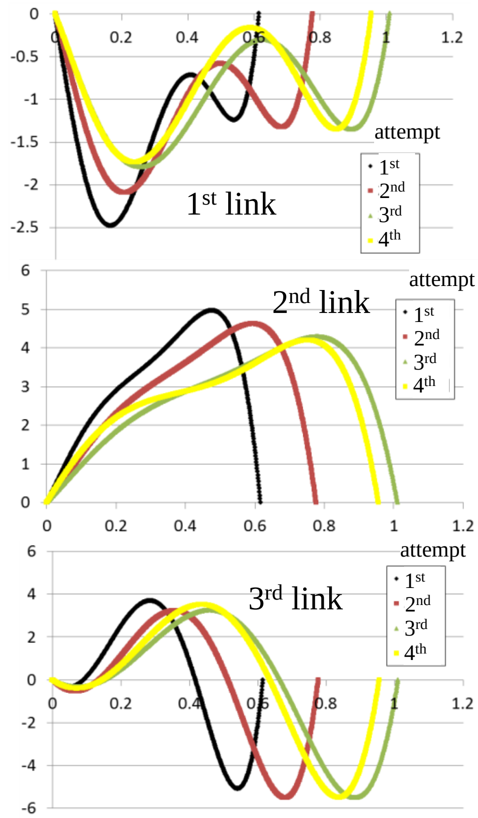
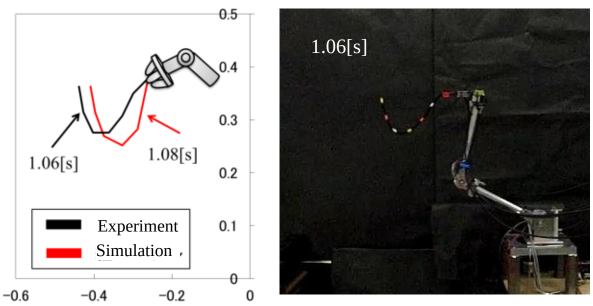


Figure 8: Projected Bezier curve; ω[rad/s] versus T[s]



Trajectory comparison [m]

When u-shape is formed

Figure 9: Results after 3rd attempt

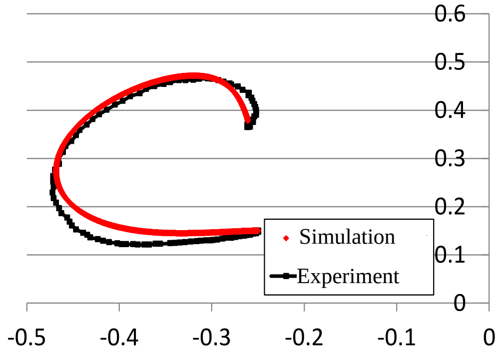


Figure 10: Trajectory of wrist of manipulator

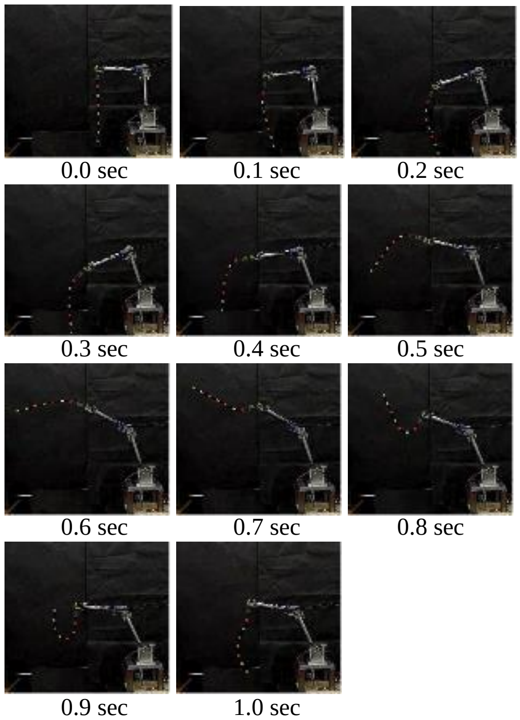


Figure 11: Scenes during string manipulation

VI. CONCLUSION

A wire-driven 3-link robotic manipulator is proposed and developed. The proposed robotic manipulator is lighter and gives more acceleration which cannot be obtained from normal robotic arms in the market, making it more dexterous and flexibility. On top of that, the combination between the robot and our developed simulator enables the robot to plan its movement based on the prediction or solution provided by the simulator, which is intelligent and cost effective.

Experimental results have confirmed the success of the proposed manipulator to a certain extent. However, the manipulator still requires improvement, especially to further improve the transmission of power from the motor through the wires. Further works may also include application to clothes to check the robustness of the manipulator as well as the simulator.

ACKNOWLEDGMENT

The authors would like to thank Ministry of Science, Technology and Innovation, Malaysia and Kanazawa University, Japan for funding this work.

REFERENCES

[1] J.G. Ryu, B.C. Yoo and T. Nishmura, "Service Robot Operated by CDMA Networks for security Guard at Home", Service Robot Applications, InTech, pp. 184-194, 2008.

[2] T. Taipalus and K. Kosuge, "Development of Service Robot for Fetching Objects in Home Environment", Proc. of IEEE Int. Conf. on Computational Intelligence in Robotics and Automation, pp. 451-456, 2005.

[3] A. Bicchi, "Hand for Dextrous Manipulation and Robot Grasping: A Difficult Road Toward Simplicity", IEEE Transaction on Robotics and Automation, vol. 1, no. 16, pp. 622-652, 2000.

[4] K. Yamazaki and M. Inaba, "A Cloth Detection Method Based On Image Wrinkle Feature for Daily Assistive Robot", IAPR Conference on Machine Vision Applications, pp. 366-369, 2009.

[5] S. Hata, T. Hiroyasu, J. Hayash, H. Hojoh, and T. Hamada, "Flexible handling robot system for cloth," Proc. of IEEE Int. Conf. on Mechatronics and Automation, pp. 49–54, 2009.

[6] K.S.M. Sahari, H. Seki, Y. Kamiya and M. Hikizu, "Clothes Manipulation by Robot Grippers with Roller Fingertips", Advanced Robotics, vol. 24, no. 1-2, pp. 139-158, 2010.

Switching based controller algorithm design for uncertainties rejection in smooth takeoff / landing and of quad-rotor

M. Hassan Tanveer, D. Hazry, S. Faiz Ahmed, M. Kamran Joyo, Faizan. A. Warsi, A.T. Hussain
Centre of Excellence For Unmanned Aerial Systems (COEUAS) Universiti Malaysia Perlis (UniMAP), Seriab, 01000, Kangar, Perlis, Malaysia.
hassantanveer@live.com, , hazry@unimap.edu.my, syedfaiz@unimap.edu.my, sky_kamran12@hotmail.com, f_warsi128@hotmail.com, abadal@unimap.edu.my

Abstract: The proposed control technique for smooth takeoff / landing controlling of Quad-rotor type (UAV) unmanned Aerial Vehicles under different uncertainty conditions is evaluated in this article. Takeoff / landing controlling of Quad-rotor UAV under high wind burst and noises has been a challenging research domain for the researchers. In this paper a switching based controller is proposed for better controlling of quad-rotor through NMPC and PID controller. This switching controller works in such a way that noises problem are handled by NMPC controller and in the occurrence of external disturbances, auto-tuned PID controller is activated. Simulation Results shows the effectiveness of the proposed switching based method.

Keywords: Quad-rotor, Takeoff / Landing Control, Robust PID, Non Linear Model Predictive Control (NMPC).
.

I. INTRODUCTION

The demands of UAVs have influenced the researchers to design UAVs that are capable of operating in rough surroundings and effectively operational in complex missions. During the last decade the UAVs capable with vertical takeoff and landing configuration have intensively been researched as their demands in civilian applications have increased. Copters and rotorcrafts have proved to be the great solution to the unmanned aviation industry. The quad-rotor have become the most prominent amongst the researchers as they are capable of maneuvering efficiently and their VTOL capability makes them part of helicopter type UAVs. Its advantage over other copters and rotorcrafts is its easy structural assembling. However this type of rotorcraft UAV carries severe issues regarding its aerodynamic controls which is responsible of its rough flight performance and requires an effective controller. In this article altitude controlling for quad-rotor is focused. An innovative hybrid control technique is proposed based on PID and MPC controllers. The PID tackles the disturbances and MPC tackles the noisy impact of the sensors.

In the recent years VTOL and altitude control of quad-rotor has remained an issue due to the kinematic constraints and unstable dynamics. However some of the control techniques have been developed in this area. In 2013 Adaptive Hybrid Control Algorithm Design for Quad-rotor UAV by (D. Hazry, 2013) , In 2006, Ozguner used sliding mode and PID control for the system(R. Xu, 2006). In 2007, Samir Bouabdullah and Ronald Siegwart used a Backstepping method for altitude, attitude and position control of a quad-rotor, the results of this technique showed a flexible control structure. Moreover, showed that quad-rotor was able to perform autonomous hovering with altitude control and autonomous take-off and landing (Samir Bouabdullah, 2007). In 2008, Bouadi presented stabilizing control laws by sliding mode and backstepping approach (H. Bouadi, 2008). This paper focuses on proposed hybrid control technique that is based on MPC and PID for VTOL and altitude control. The proposed algorithm is simulated on (MATLAB).
The organization of the paper is structured as the section 2 refers to the related work. Section 3 follows the Kinematics and Dynamics of quad-rotor. The proposed control technique is discussed in section 4. Experimental results and simulations have been provided in section 5.

II. QUADROTOR SYSTEM

These equations of motion are extracted from the Newton Euler method. The above-mentioned method is discussed comprehensively in the article by (M. Kamran Joyo, 2013).

$$\begin{cases} \ddot{X} = (\sin\psi\sin\phi + \cos\psi\sin\theta\cos\phi)\frac{U_1}{m} \\ \ddot{Y} = (-\cos\psi\sin\phi + \sin\psi\sin\theta\cos\phi)\frac{U_1}{m} \\ \ddot{Z} = -g + (\cos\theta\cos\phi)\frac{U_1}{m} \\ \dot{\phi} = \frac{I_{YY}-I_{ZZ}}{I_{XX}}\theta\psi - \frac{I_{TP}}{I_{XX}}\theta\omega + \frac{U_2}{I_{XX}} \\ \dot{\theta} = \frac{I_{ZZ}-I_{XX}}{I_{YY}}\theta\psi - \frac{I_{TP}}{I_{YY}}\theta\omega + \frac{U_3}{I_{YY}} \\ \dot{\psi} = \frac{I_{XX}-I_{YY}}{I_{ZZ}}\theta\psi - \frac{U_4}{I_{ZZ}} \end{cases}$$
(1)

$$\begin{cases} U_1 = b(\omega_1{}^2 + \omega_2{}^2 + \omega_3{}^2 + \omega_4{}^2) \\ U_2 = lb\,(-\omega_2{}^2 + \omega_4{}^2) \\ U_3 = lb(\omega_1{}^2 + \omega_3{}^2) \\ U_4 = d(-\omega_1{}^2 + \omega_2{}^2 - \omega_3{}^2 + \omega_4{}^2) \end{cases}$$
(2)

Where 'l' is the distance between the center of the quad-rotor and the center of a propeller. U1, U2, U3 and U4 are the movement vector components. Their relation with the propellers' speeds comes from aerodynamic calculus.
For takeoff / landing controlling only Z axis equation is taken in consideration, illustrated in Equation (1),

$$\ddot{Z} = -g + (\cos\theta \cos\phi)\frac{U_1}{m}$$

III. PROPOSED CONTROLLER DESIGN

In this Paper, a hybrid PID and MPC controller technique is proposed for proper altitude controllability of quad-rotor under the disturbance and noisy conditions. The sensors are noisy in nature and specially during windy conditions smooth takeoff and landing of quad-rotor are quite complex tasks .In this proposed control method both control strategies (PID & MPC) are utilized in such a way that it can overcomes these two majour above mentioned problems. Figure 2 shows the block diagarm of proposed technique. A decision making block switches the system modle to the appropriate controller according to the desired situation. Under the circumstances of external distrubance the error sinal of between desired altitude and original altitude (sensor output) become greater then some theroshold value, in this situation decision making block switches the quadrotor system to the PID controller which quickly stabilize it and makes takeoff and landing operation smooths. Decision making block working is defined in following two equations.

$e_z > \tau \rightarrow$ *PID controller*

$e_z < \tau \rightarrow$ MPC controller

Where τ is the therosholled value of error signal &

e_z is the error signal between deisred and actual altitude position value

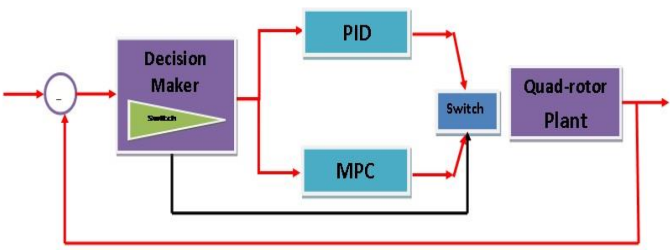


Figure 2: Proposed Hybrid Controller

Similarly under normal conditions the The system moble conntes with Model predictive controller MPC and it is selected for better controlling of the system under noisy conditions, it predicts control signal in such a way that it minimizes a define cost function which is error signal between the output and desire value over specified prediction horizon. MPC starts predicting the future control action by choosing the suitable values of control horizon *M*, prediction horizon *P* and control-weighting factor *R*. Once MPC completes its prediction process, it implements best control action and then at the next sampling interval, the control estimation is repeated again with the new available information. As a result, the performance of the system becomes increased.

IV. PID CONTROLLER DESIGN FOR DISTURBANCE REJECTION

The altitude controller used to maintain the quad-rotor distance from the ground at a desired value. For achieving stabilized quad-rotor altitude, main requirement is to work on quad-rotor z axis and remaining all other axis constant. So we select only Z axis equation form equation (8), as show in following:

$$\ddot{z} = g - (\cos\emptyset \cos\theta)\frac{U_1}{m}$$
(3)

In above equation g and $\frac{1}{m}$ are constant values, $\emptyset$ and θ must be considered as 0 for hovering condition. Using Laplace Transform and taking all constants as $g - (\cos\emptyset\cos\theta)\frac{1}{m} = K_1$, system will become:

$$Z(s) = \left(\frac{K_1}{S^2}\right)\left(\frac{0.936}{0.178S+1}\right)^2$$

(4)

So the error signals e for PID control is:

$$e_z = Z_{ref} - Z$$

Where Z_{ref} is desired value and Z is actual output.
And PID controller will become:

$$U(s) = K_p e(s) + \frac{K_i}{s} e(s) + K_d s\, e(s)$$
(5)

Kp, Ki and Kd can be selected by auto tuning method (M. Hassan Tanveer, 2013).

Let $K_B = (0.936)^2, K_c = 0.178$

$$Z(s) = \left(\frac{K_B K_1}{K_c S^4 + 1}\right) \left(K_p e(s) + \frac{K_i}{S} e(s) + S K_d e(s)\right)$$

Therefore close loop transfer function for the complete system will become:

$$\frac{Y(S)}{R(S)} = \frac{K_B K_d K_1 S^2 + K_B K_p K_1 S + K_B K_i K_1}{K_c^2 S^8 + 2 K_c S^4 + S^2 + K_B K_d K_1 S^2 + (K_p K_{1K_B} + 1) S + K_i K_{1K_B}} e(s) \quad (6)$$

Equation (6) is the controller equation, which gives desired and controlled output response.

V. NMPC Controller Design for Noise Rejection

Thus the objective function for system can be formulated as:
The dynamics of the true state is given in equation (4), can be expressed as:

$$x\ (k+1) = Ax(k) + B\,u\ (k) \quad (7)$$

Prediction Horizon length N:

$$x(k) = \begin{bmatrix} x(k+1\,|k) \\ x(k+N\,|k) \end{bmatrix}$$

$$u(k) = \begin{bmatrix} u(k\,|k) \\ u(k+N-1\,|k) \end{bmatrix}$$

and the estimation error is an additive term, applied at each time step, the error is uniformly distributed in every feedback.

$$x\ (k+i+1 \mid k) = Ax(k+i \mid k) + B\,u\ (k+i \mid k) \quad (8)$$

x(k) = M x(k) + C u(k) is the state prediction.

The optimization cost function is then become:

$$J(k) = \sum_{i=0}^{N-1} [x^T(k+i \mid k) Q x(k+i \mid k) + u^T(k+i \mid k) R u(k+i \mid k) + x^T(k+N \mid k) \bar{Q} x(k+i \mid k)] \quad (9)$$

Where Q and $\bar{Q}$ are weight coefficient. Evaluating all above equations in J.

$$J(k) = [2x^T(k) F^T u(k) + u^T(k) H u(k) + x^T(k) G x(k)] \quad (10)$$

Where,

$$H = C^T \bar{Q} C + \bar{R}$$
$$F = C^T \bar{Q} M$$
$$G = M^T \bar{Q} M + Q$$

This form can capture many common constraints such as actuation limits and error boxes. Also note that the constraint acts upon the estimated state, for that is known. Given the assumption of a bounded state uncertainty, this can be readily transformed to and from equivalent constraints on the actual state. Since the feasibility of the optimization depends only on the estimate and not the truth, the form is appropriate here.

Subject to the constraint:

$$\nabla u\,(J) = 2H\,(u) + 2Fx = 0 \quad (11)$$

And the optimal solution to the optimization problem of the plant, therefore the control law for receding horizon implementation will be:

$$u\,(k) = -[I\ 0 \ldots 0]\, H^{-1} F x(k) \quad (12)$$

VI. SIMULATION RESULTS

In this section, system simulation results are shown in terms of time history which demonstrates the strength of control method for the smooth takeoff / landing and altitude control of quad-rotor. Disturbance is added in system more or less around 100 percent of reference input, and its effect i.e: Overshoot, Response Time and Settling Time are shown in Table 3. PID parameters chosen after auto-tune are listed in Table-2. Physical Parameters of quad-rotor are chosen as mention in Table-1. Initial Value of altitude z(0) is set to origin 0 position.

Figure 3 shows the actual altitude tracking path. Figure 4 shows the noise which is added in trajectory by means of system and sensor noises. Figure 5 is the disturbance which is added in system in a particular time instance. Figure 6 compares the three different control techniques effect under disturbance and noisy conditions.

Parameter Name	**Symbol**	**Value**	**Unit**
Rotational Inertia Along Z-axis	I	0.0013	$kg.m^2$
Total Mass	M	0.65	kg
Thrust Constant	B	$3.13e^{-5}$	$N.s^2$
Drag Constant	D	$7.50e^{-7}$	$N.s^2$
Arm Length	l	0.23	m

Table 1:Quad-rotor Calibration Data

Kp	**0.0007482940731823**
Ki	**0.000007482940731823**
Kd	**0.000844205739401302**

Table 2:PID Auto-tune Parameter

Controller	Overshoot	Settling Time	Response Time
PID	10%	1 sec	4 sec
MPC	90%	2 sec	2 sec
Hybrid	10%	1 sec	2 sec

Table 3: Effect of different controllers

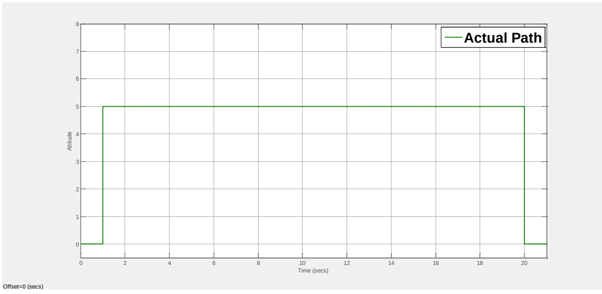


Fig. 3: Actual Altitude Path

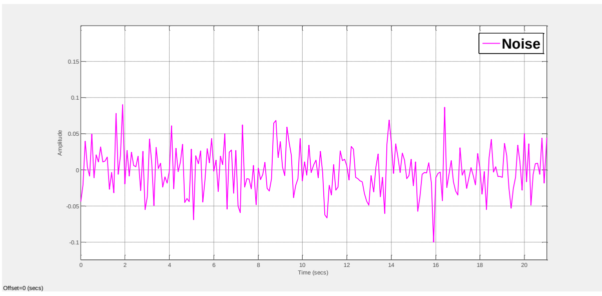


Fig. 4: Gaussian noise added in System

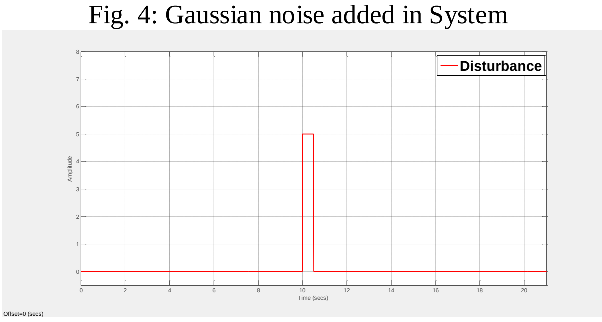


Fig. 5: Disturbance

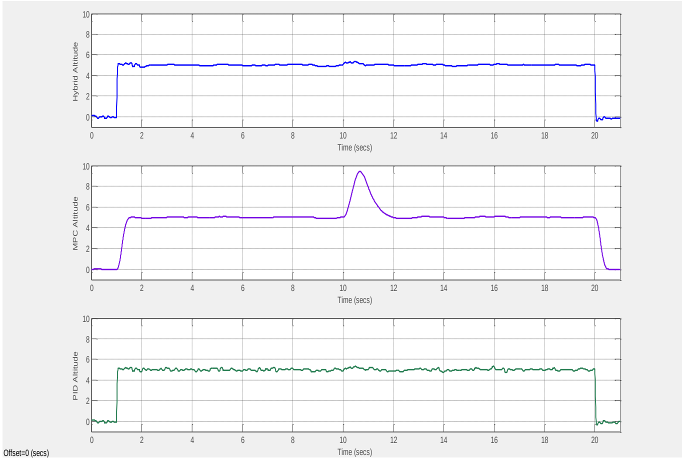


Fig. 6: Controllers Effect

VII. CONCLUSION

In this paper we present the efficient auto-tuned PID controller with MPC for Disturbance and Noise Rejection in Smooth Takeoff / Landing and Altitude Stabilization of Quad-rotor UAV. NMPC shows its robustness in overall simulations when system and sensor noises injects in system. Result shows that proposed auto-tuned PID controller works very well and quickly become stable when external disturbances are applied. From the simulation results, we can conclude that proposed switching controller works efficiently and makes the system stable under the circumstance where there is any kind of disturbance and noises imposed on it.

VIII. REFERENCES

[1] N. Guenard et al. 2006 “Control laws for the tele operation of an unmanned aerial vehicle known as an x4-flyer,” in Proc. (IEEE) International Conference on Intelligent Robots (IROS’06), Beijing, China.

[2] R. Xu and u. Ozguner, 2006. “Sliding mode control of a quad-rotor helicopter,” Proc. of the 45th IEEE conference on decision and control, pp. 4957-496.

[3] D.Hazry, S.Faiz Ahmed, 2013 “Adaptive Hybrid Control Algorithm Design for attitude stabilization Of Quad-rotor UAV”, archives Des Science Journal Vol.66, Issue.2.

[4] Tanveer, M. H., Ahmed, S. F., Hazry, D., Warsi, F. A., & Joyo, M. K. (2013). STABILIZED CONTROLLER DESIGN FOR ATTITUDE AND ALTITUDE CONTROLLING OF QUAD-ROTOR UNDER DISTURBANCE AND NOISY CONDITIONS. American Journal of Applied Sciences, 10(8), 819.

[5] Samir bouabdullah,2007. “Full Control of a Quad-Rotor,” Proc of IEEE international conference on intelligent robots and systems (iros 2007).

[6] H Bouadi, m. Bouchoucha and m. Tadjine, 2008 “Sliding mode control based on backstepping approach for an uav type-quad-rotor,” international journal of applied mathematics and computer sciences, vol. 4, no. 1, pp. 12-17.

[7] Joyo, M. K., Ahmed, S. F., Hazry, D., Tanveer, M. H., & Warsi, F. A. (2013). Position Controller Design for Quad-rotor under Perturbed Condition.

[8] Tanveer, M. H., Ahmed, S. F., Hazry, D., Joyo, M. K., & Warsi, F. A. (2013). Disturbance And Noise Rejection Controller

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Parameters of Gas-Liquid Flow Distribution Uniformity in Upward Multi-Pass Channels

Zuradzman M. Razlan[1], D. Hazry[1], R. Heng[1], A.B. Shahriman[1], Khairunizam WAN[1], S. Faiz Ahmed[1], Nazrul H. ADNAN[1], M. Hirota[2], N. Maruyama[2], A. Nishimura[2], and H. Hisyam[3].

[1]Centre of Excellence for Unmanned Aerial Systems, Universiti Malaysia Perlis, 01000 Kangar, Perlis, Malaysia.
[2]Department of Mechanical Engineering, Mie University, Tsu-city, Mie, Japan.
[3]Lean Applied Pte. Ltd., 47620 Subang Jaya, Selangor, Malaysia.
zuradzman@unimap.edu.my

Abstract: The gas-liquid flow distributions in multi-pass upward parallel channels that simulate the evaporator for an electric vehicle air-conditioner system were examined experimentally. In this paper, the attentions are to discover the most influenced parameter to the flow distribution uniformity by using design of experiment method. Experiments were conducted in an isothermal air-water flow system. In the mist-flow inlet, the water distribution was insensitive to the backpressure conditions and its uniformity was improved in comparison with that in the stratified-flow inlet. The flow distribution uniformity for gas phase is influenced mostly by superficial air velocity, and the flow distribution uniformity of liquid phase is mostly influence by 2-way interaction of parameters which are flow pattern and superficial air velocity.

Keywords: Mal-distribution, Stratified flow, Mist flow, Backpressure, Design of Experiment.

1 INTRODUCTION

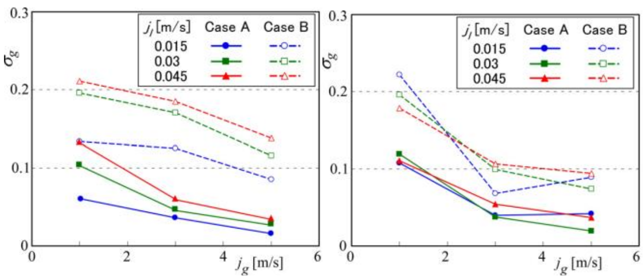
In this study, we already examined experimentally the gas-liquid flow distribution characteristics in multiple upward channels that simulate the compact evaporator used in the small air-conditioning system, with attention to the influences of the backpressure conditions at the branch outlets and of flow-inlet conditions at the entrance of the dividing header on the gas-liquid distributions as reported in Zuradzman et. al [1]. The gas-liquid flows of the refrigerant have been simulated by the air-water two-phase flows under isothermal condition that are suitable for grasping fundamental flow characteristics in the channel. The distribution ratios of air and water in the branches have been measured under the upward parallel flow condition. It is expected that the data of gas-liquid distributions obtained under these specified inlet and outlet conditions are helpful not only to understand the fundamental two-phase flow characteristics in the multi-pass channels but also as a database to examine the reliability of results obtained by numerical simulations.

2 METHODOLOGIES

This paper focus on analysis from the result of standard deviation of gas and liquid phase (flow distribution uniformity) has been done and reported recently by Zuradzman et. al. [1] as in Fig. 1 and 2. It is important for acknowledgement of which parameters and its interaction among each other that influenced the most to the uniformity of the flow distribution by using Minitab 15 statistical software program, a design of experiment (DOE) analysis has been done to study which parameter that influenced the most.

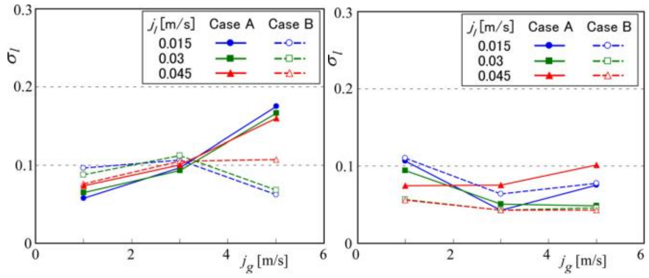
Table 1 Summary of the experimental conditions or parameters from Zuradzman et. al [1].

Fluids	Isothermal air and water
Superficial air velocity at the header entrance j_g	1.0 m/s, 3.0 m/s, 5.0 m/s
Superficial water velocity at the header entrance j_l	0.015 m/s, 0.03 m/s, 0.045 m/s
Pressure condition at the branch outlets	Case A (non-uniform) Case B (uniform)
Flow-inlet condition at the header entrance	Stratified-flow inlet Mist-flow inlet
Header attitude	Horizontal
Branch attitude	Upward



(a) Stratified-flow inlet **(b)** Mist-flow inlet

Fig. 1 Standard deviations of the air distribution ratios σ_g



(a) Stratified-flow inlet **(b)** Mist-flow inlet

Fig. 2 Standard deviations of the water distribution ratios σ_l

3 ANALYSIS RESULT

In this paper, an analysis of parameters or factors that influenced the uniformity of flow distribution is discussed. Seems the liquid superficial velocity has been observed clearly that has a small influence to the uniformity of flow distributions, it has been decided to focus the analysis with these sources; (1) Flow pattern, (2) Backpressure, (3) Superficial air velocity (4) 2-way and 3-way interactions between parameters. This analysis is important for discovery the most valuable parameter that contributed to the uniformity of the flow distribution and the best setting of the parameters level to archive the lowest σ_g and σ_l.

As in the previous session, total of 36 combinations of test configuration with 4 parameters has been done and σ_g and σ_l has been calculated and being plot as in Fig. 1 and 2. Superficial water velocity j_l has a minor influence to the uniformity of flow distribution and by considering only 3 parameters as in table 2. The σ_g and σ_l that to be analyzed is reduces to 8 combinations. By, repeating the combination 3 times which the first, second and third 8 is the test data at j_l = 0.015 m/s, 0.030 m/s and 0.045 m/s respectively. Total data to be analyze are 24 combinations. By using the "Design of Experiment" method, or in mathematics and statistics study in area of "Analysis of Variance (ANOVA)" (Brooks, 1979), the analysis has been done by using a program called Minitab 15 statistical software.

First, the result of gas phase flow distribution uniformity analysis is addressed. An ANOVA table as in Table 3 has been made by using General Linear Model in Minitab 15 software from the 24 combination results of σ_g. From the table, clearly shows only backpressure and j_g are the main parameters contribute to the uniformity of gas phase flow distribution as the *P-value* in the table are less than 0.05. *P-value,* calculated from *F* ranging from 0 to 1. It is a hypothesis test to check whether the parameter is significant to the contribution of the uniformity of the flow distribution or not. (*P-value* < 0.05≡significant, *P-value* > 0.05≡not significant)

For further understanding on how large the contribution of these parameters, by calculate the percentage of *SS* value to the *SS* total from table 3, a simplified Pareto chart can be plot to show clearer contribution of each parameter and its interaction among each other to the uniformity of gas phase flow distribution as in Fig. 3. The abscissa shows the sources which are the parameters and its interaction combinations. The ordinate shows the percentage of contribution to the uniformity of the gas phase flow

Table 2 Summary of the selected analyzed parameters and its level.

Parameter	Level
Flow pattern at header entrance	Stratified-flow, Mist-flow
Backpressure	Non-uniform, Uniform
j_g (m/s)	1.0 m/s, 5.0 m/s

Table 3 Summary of the ANOVA table for σ_g.

Source	*DF*	*SS*	*MS*	*F*	*P-value*
flow pattern	1	0.000043	0.000043	0.08	0.7860
backpressure	1	0.035378	0.035378	63.30	0.0000
j_g	1	0.04134	0.04134	73.97	0.0000
flow pattern *backpressure	1	0.000241	0.000241	0.43	0.5210
flow pattern * j_g	1	0.001049	0.001049	1.88	0.1900
backpressure * j_g	1	0.000276	0.000276	0.49	0.4920
flow pattern * backpressure * j_g	1	0.000552	0.000552	0.99	0.3350
Error	16	0.008942	0.000559		
Total	23	0.08782			

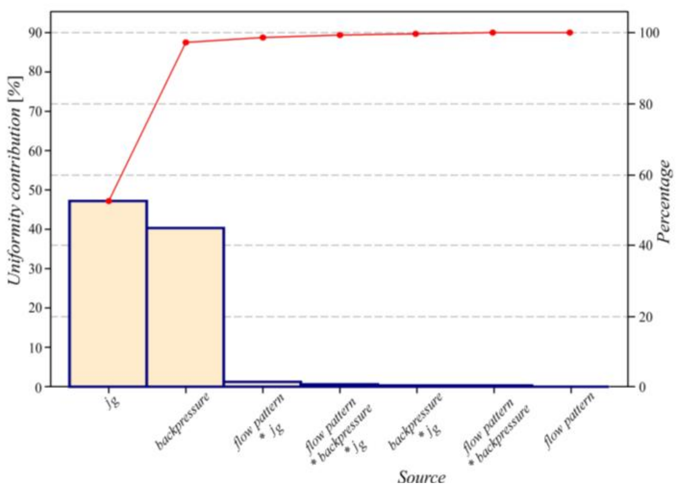


Fig. 3 Pareto chart of sources (parameters) contribute to the uniformity of gas phase flow distribution

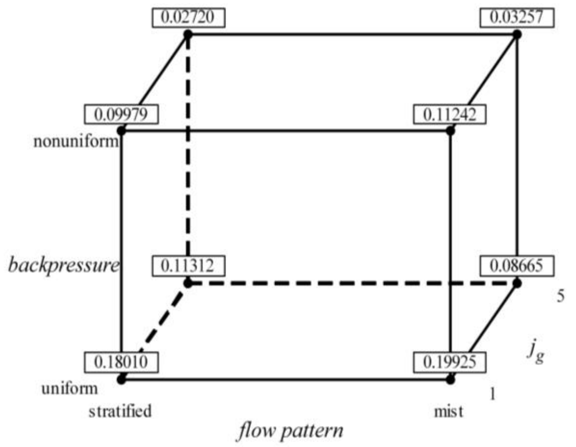


Fig. 4 Cube plot with means value of σ_g at each combination of setting level for each parameter.

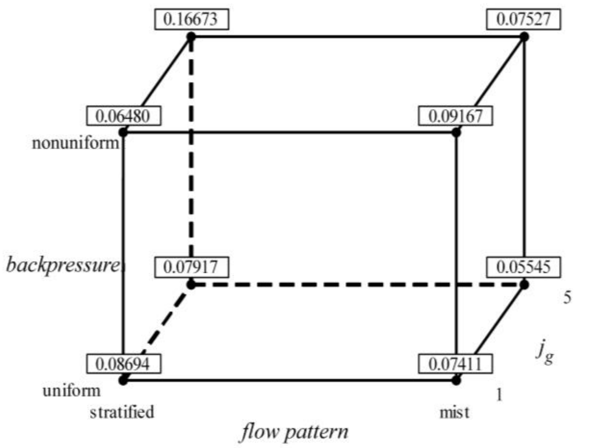


Fig. 6 Cube plot with means value of σ_g at each combination of setting level for each parameter.

Table 4 Summary of the ANOVA table for σ_l.

Source	*DF*	*SS*	*MS*	*F*	*P-value*
flow pattern	1	0.0038346	0.0038346	9.80	0.0060
backpressure	1	0.0039629	0.0039629	10.13	0.0060
j_g	1	0.0013102	0.0013102	3.35	0.0860
flow pattern * backpressure	1	0.0002948	0.0002948	0.75	0.3980
flow pattern * j_g	1	0.0062611	0.0062611	16.00	0.0010
backpressure * j_g	1	0.0047	0.0047	12.01	0.0030
flow pattern * backpressure * j_g	1	0.0043281	0.0043281	11.06	0.0040
Error	16	0.006261	0.0003913		
Total	23	0.0309525			

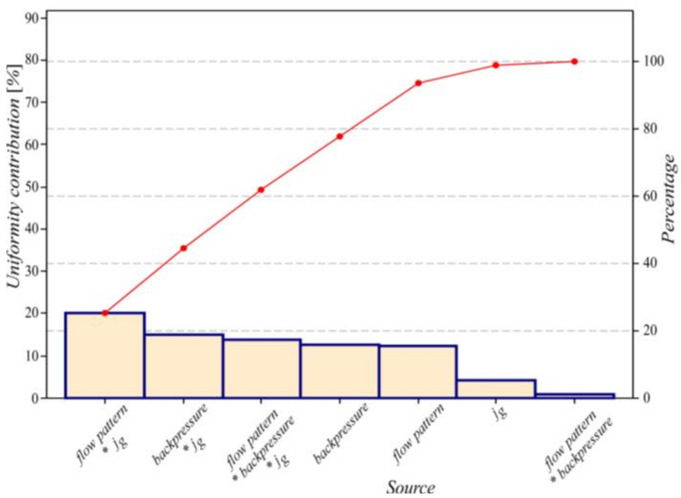


Fig. 5 Pareto chart of sources (parameters) contribute to the uniformity of liquid phase flow distribution

distribution. From the chart, the total of contribution by all the parameters to the uniformity of flow distribution of gas phase is 89.82%. The other 10.18% is the experimental error or considered unknown. From this 89.82% of contribution, the j_g with 47.07% continuing with backpressure with 40.28% contribute the most for the uniformity of gas phase. The other parameters and all the interaction shall be classified as not significant to the contribution of uniformity of flow distribution.

The next analysis is to find the best combinations of parameters and its level to create the best setting for the best uniformity of flow distribution. By using Minitab 15 software again with design of experiment cube plot tool, the result yield as in Fig. 4. This cube plot shows the mean value of σ_g at each combination of the 3 parameters and its level as in Table 2. From the cube plot, the smallest mean value of σ_g is 0.02720 which is located at left, back and upper side of the plot. Thus to ensure the σ_g at minimum level, the setting of each parameter should be as follow. Backpressure should be at non-uniform condition, j_g should be set at 5.0 m/s and flow pattern should be set with stratified flow. Noted, the flow pattern is considered a not significant parameter as explain in table 3 and Fig. 3, thus make the differential of σ_g value between mist-flow and stratified-flow in Fig. 4 is relatively small.

Parameters that contribute the most to the uniformity of liquid phase flow distribution shall be discussed next. As previously explained, an ANOVA table from the 24 combination of parameters experiment σ_l result is been made as in Table 4. Different from gas phase ANOVA table, this table, the P-value for 3-ways interaction shows lower value than 0.05, meaning that it is significant to the contribution to the uniformity of liquid phase flow distribution. Only one of the 2-ways interaction combinations is not significant to the uniformity of flow distribution, i.e. flow pattern and backpressure. Nevertheless, j_g P-value is higher than 0.05, however due to its interactions among other parameters significant to the flow distribution, makes j_g significant as a main factor as others.

Fig. 5 shows a Pareto chart of contribution percentage to the uniformity of liquid phase flow distribution by its main parameters solely and also by its parameters interaction among each other. From the chart, the total of contribution by all the parameters and its interactions to the uniformity of flow distribution of liquid phase is 79.77%. The other 20.23% is the experimental error. The chart shows that the 2-ways interaction of flow pattern and j_g is the most influenced to the uniformity of liquid phase flow distribution that contributes 20.23% continue with combination of backpressure and j_g with 15.18%. From this observation of Fig. 5, it shows that the interaction among parameters are more important than the parameter itself to the contribution of the uniformity of liquid phase flow distribution.

Since interaction contributed more than the parameter itself, from the experiment to find the best setting for each factor that can sustain minimizing the standard deviation of liquid by a design of experiment cube plot tool as explain earlier. Fig. 6 is the result of cube plot for mean value of σ_l. From the figure, to archive the smallest value of σ_l the parameters should be set as follows. Flow pattern should be mist flow, the backpressure should be uniform condition and j_g should be set at 5.0 m/s.

4 CONCLUSIONS

An analysis of variances that contribute to the uniformity of gas and liquid phase flow distribution by using a computer program called Minitab 15 statistical software has been made. The results are summarized as follows.

(1) 24 selected standard deviation's data has been analyzed by using design of experiment method through Minitab 15 statistical software. For uniformity of gas phase flow distribution, the main parameter that contributed the most is j_g. For uniformity of liquid phase flow distribution, it was the 2-way interaction between flow pattern and j_g that contribute the most.

(2) The best setting of level of each parameter for getting the best uniformity of gas phase flow distribution are:
  (a) j_g shall be at 5.0 m/s.
  (b) Backpressure condition shall be set at non-uniform.
  (c) Flow pattern at header entrance however shall be either stratified or mist-flow.

(3) Also, the best setting of level of each parameter for getting the best uniformity of liquid phase flow distribution are:
  (a) Flow pattern at header entrance shall be Mist-flow.
  (b) Backpressure condition shall be set at uniform.
  (c) j_g shall be at 5.0 m/s.

NOMENCLATURE

j	superficial velocity (m/s)
M	mass flow rate (g/min)
P	pressure at the exit of branch (kPa)
DF	degree of freedom
SS	sums of squares
MS	means of squares
F	F test, equality of variance between factor levels.
P-value	P-values range from 0 to 1
Greek Symbols	
σ	standard deviation
Subscripts	
g	gas phase or air
l	liquid phase or water
i	*i* th branch (*i*= 1…10)

REFERENCES

[1] Zuradzman Mohamad Razlan, Ryota Isobe, Yasuhiro Mizuno, Hiroaki Goshima, Masafumi Hirota, Naoki Maruyama, "Gas-liquid Distributions in Upward Multi-pass Channels of Compact Evaporator", *14th International Heat Transfer Conference (IHTC14)*, paper no. IHTC14-22588

An Empirical Examination of Inter-firm Capital Relationships in Mazda's Yokokai using the IDE Spatial Model

M. Sakamoto[4], S. Tagawa[1], T. Ito[1], S. Matsuno[1], R. Mehta[2], V. Berdonosov[3], and S. Ikeda[4]

1 Dept. of Business Administration, Ube National College of Technology, Yamaguchi, Japan
(Tel : 81-836-35-7115; Fax : 81-836-35-7115)
(tagawa@ube-k.ac.jp, ito@ube-k.ac.jp, matsuno@ube-k.ac.jp)
2 School of Management, New Jersey Institute of Technology, U.S.A
(Tel :1-973-596-6419; Fax :1-973-596-3074)
(mehta@njit.edu)
3 Faculty of Computer Science, Komsomolsk-on-Amur State University of Technology, Russia
(Tel :7-962-287-51-41; Fax :7-4217-535259)
(berd1946@gmail.com)
4 Dept. of Computer Science and Systems Engineering, University of Miyazaki, Japan
(Tel : 81-985-58-7392; Fax : 81-985-58-7392)
(sakamoto@cs.miyazaki-u.ac.jp)

Abstract: Just as workflow and transactional linkages, inter-firm capital relationships are similarly known to have an impact on corporate performance. Because of the superior performance of Japanese networks known as Keiretsu, many companies have begun to recognize the importance of capital linkages whereas other companies have placed emphasis on improving coordination of managerial resources. In this paper, we review the literature on network organizations with a lens focused on examining the capital relationships in Yokokai, the Mazda's Keiretsu. Employing regression analysis, three significant indices, influence, degree, and efficiency, were selected from nine indices, including others that include dyadic redundancy, dyadic constraint, effective size, constraint, hierarchy, and density. We propose a new approach, called IDE spatial model, to calculate the strength of the inter-firm's relationships. In order to ascertain the rational inter-firm relationships, network indices and corporate performance are analyzed. Therefore, this research suggests a new perspective to examine the rational inter-firm relationship that can be used in any network organization.

Keywords: influence, degree, efficiency, the IDE spatial model, Keiretsu

I. INTRODUCTION

Just as workflow and transactional linkages, inter-firm capital relationships are similarly known to have an impact on corporate performance. Because of the superior performance of Japanese networks known as Keiretsu, many companies have begun to recognize the importance of capital linkages whereas other companies have placed emphasis on improving coordination of managerial resources. In this paper, we propose a new approach, called IDE spatial model, to calculate the strength of the inter-firm's relationships. In order to ascertain the rational inter-firm relationships, network indices and corporate performance are analyzed. Therefore, this research suggests a new perspective to examine the rational inter-firm relationship that can be used in any network organization.

This paper is organized as follows. In Section 2, we briefly review some relevant literature of quantitative approaches of Keiretsu. Section 3 introduces the IDE spatial model. Section 4 shows the results, and discusses the implications of the results. Finally in Section 5 we conclude by a summary of this paper.

II. BACKGROUND

One of the important factors in structural analysis, a strong tie and/or weak tie is popular way to describe the different strength of a relationship.

However, quantitative approaches are required to discover the determinants of corporate performance. Relationships among firms based on distances between their locations and the frequency of face-to-face communication among engineers in the automobile industry has been examined [1]. Moreover, Fukuoka et al. reported a new finding on relationships among member firms in Nissan's using transactions and cross shareholdings [2]. Recently Ito et al. applied graph theory to network organization analysis, and identified some characteristics such as centrality, size of network as being important facets of relationships [3]. Transactional inter-firm's relationships are the key concepts in these studies. In a Keiretsu, capital inter-firm's relationship is known as cross shareholding. To the best of our knowledge, only a few studies have examined cross shareholdings from an organizational network vantage point; therefore this investigation attempts to shed light on one of the most advanced quantitative analysis using data gathered from Mazda's Yokokai Keiretsu.

III. METHOD

Many spatial paradigms, such as the IDCBI model and the DEC spatial model have been developed to discern the relationship between network indices and corporate performance. One of the key questions of this investigation is to determine if all these models are suitable for the analysis of capital network. In order to identify the factors that impact corporate performance, we employed regression analysis using three significant indices, influence, degree, and efficiency, were selected from nine indices. The others six are dyadic redundancy, dyadic constraint, effective size, constraint, hierarchy, and density. We propose a new approach, called IDE spatial model, to calculate the strength of the inter-firm's capital relationships.

1. Outline of the IDE spatial model

Generally influence means one kind of power to affect persons or events, or causing something without any direct or discernible effort. Influence reflects the power to influence or have an impact on other member firms directly and indirectly in a network. Consequently, influence will be divided into two parts: direct influence and indirect influence. Suppose that A is the matrix of the direct network, and A^n means the indirect influence from one firm to another firm by n steps. Then influence is calculated as follows.

$$T = A + R = A + A^2 + A^3 + \cdot\ \cdot\ \cdot + A^n$$
$$= A(I - A)^{-1}$$

where

T: Total influence;
A: direct influence;
R: indirect influence;
I: Identity matrix.

An integral index known as centrality has many facets that include degree, closeness, and betweenness. Degree expresses a firm's potential communication activity. In cross shareholding network, degree includes two categories: in-degree and out-degree, because cross shareholding networks are considered to be asymmetric organizations. In-degree refers to a firm accept investment from other member firms, whereas out-degree reflects a firm that only buy stocks of other firms within the network. Degree is calculated as below.

$$C_D(p_k) = \sum_{i=1}^{n} a(p_i, p_k)$$
$$i = 1, 2, \cdots, n; \quad k = 1, 2, \cdots, n$$

where

$a(p_i, p_k) = 1$; if and only if p_i and p_k are connected by a line

= 0; otherwise

In this paper, percent data of inter-firm's transactions are collected from Yokokai.

Efficiency is another important index. Effective size of the network is the number of alters that ego has, minus the average number of ties that each alter has to other alters. Suppose that A has ties to three other actors. Further, suppose that none of these three has ties to any of the others. Thus the effective size of ego's network is three. Alternatively, suppose that A has ties to three others, and that all of the others are tied to one another. A's network size is three, but the ties are "redundant" because A can reach all three neighbors by reaching any one of them. The average degree of the others in this case is 2 (each alter is tied to two other alters). Therefore, the effective size of the network is its actual size, i.e., 3, reduced by its redundancy, i.e., 2, to yield an efficient size of 1. Efficiency expresses the effective size of ego's network by its actual size. That is, what proportion of ego's ties to its neighborhood is "non-redundant"? Thus, efficiency tells us how much impact ego is getting for each unit invested in using ties. A firm can be effective without being efficient; and an actor can be efficient without being effective.

2. Data collection and visualization

In order to measure influence, degree, and efficiency of each firm in Yokokai, cross shareholding data in the Yokokai Keiretsu in 1997 and 2004 have been collected from our interviews and the publications of the Japan Auto Parts Industries Association and Automotive Parts Publishing Company [4].

In 1997, 53 financial institutions, 190 parts suppliers, and 11 carmakers are included in Yokokai. The detailed information of the numbers of financial institutions, parts-supplies, and carmakers in Mazda' cross shareholdings network in 1997 and 2004 are shown in Table 1.

Table 1 Numbers of financial institutions, parts suppliers and carmakers in Yokokai in 1997 and 2004

	1997	2004
Financial Institutions	53	36
Parts-suppliers	190	177
Carmakers	11	11
Total	254	224

In this paper, a singleton, which can be found in any network, is defined as a firm that exists independently, without any relationship with other firms. Therefore, a singleton generally is an isolated firm in a given network. In 1997, 132 singletons including 26 financial institutions, 105 parts-suppliers, and 1 carmaker have been found. In 2004, 1 financial institution, 120 parts-suppliers, and 2 carmakers are singletons. All of the singletons have been removed in this paper because the number of singletons has no any impact on the calculation results.

The cross shareholdings relationships among the companies were identified through three dimensional modeling. A bar shows the percentage of the transaction between each pairs of firms. The three dimensional capital inter-firm relationships in Yokokai are depicted in Figure 1 and 2 respectively.

The cross shareholdings relationships among the companies were identified through three dimensional modeling. A bar shows the percentage of the transaction between each pairs of firms. The three dimensional inter-firm capital relationships in Yokokai are shown in Figure 1 and 2 respectively.

Compared with Figure 1, we find that the interactivity between financial institutions and parts-suppliers is lower, whereas the interactivity between parts suppliers is higher. The reason is that the interactive relationship among financial institutions, suppliers and carmakers changed rapidly with the varying of economic environment. For suppliers, the control from carmakers is becoming looser than before. Now suppliers can more easily forge a new capital relationship with other firms.

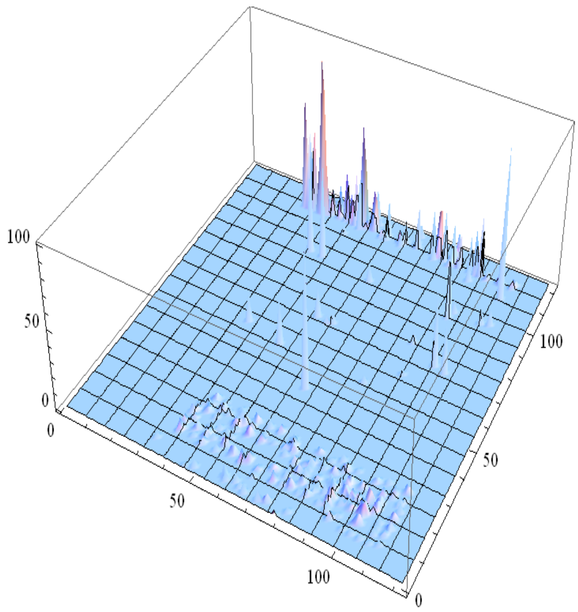


Figure 1 Inter-firm's capital network in 1997

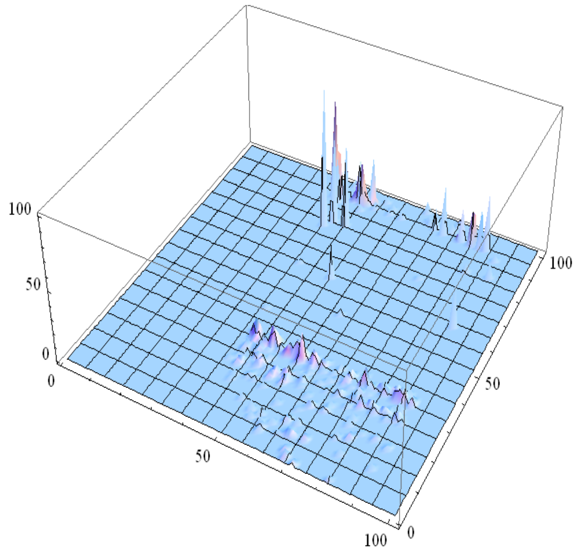


Figure 2 Inter-firm's capital network in 2004

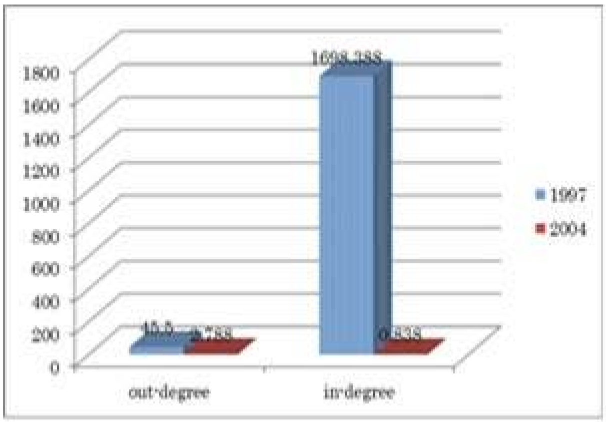


Figure 3 Out-degree and in-degree in 1997 and 2004

IV. RESULTS AND DISCUSSIONS

1. Centrality changes

In an asymmetric network, degree can be divided into two parts: out-degree and in-degree. The result of the out-degree and in-degree of the whole cross shareholdings is shown in Figure 3.

The out-degree and in-degree in 2004 are lower than that in 1997. Lower centrality means that much more companies invest their money to buy parts suppliers they interested. In other word, parts supplier can accept much more different companies' investment without permission from the top firm of their Keiretsu. Especially the dramatically declining of in-degree inflects the structure changes happened in Yokokai.

2. Determinants of profits

We calculated partial correlation coefficient of the nine indices in multiple regression model, and found density, effective size and other 4 indices are not significant. Thus we removed those 6 variables and re-estimated the regression model. The results in 1997 are shown in Table 1.

TABLE 1 RESULTS OF THE VARIANCE ANALYSIS

	Sum of Squares	Degree of Freedom	Mean Square	F Test
Total	4.239E+11	61		36.721
Regression	2.777E+11	3	9.257E+10	Prob.
Residual	1.462E+11	58	2.521E+09	0.000

Coefficient of determination: 0.65510
Multiple correlation coefficient: 0.80938
Adjusted R-square: 0.79829
AIC: 1,521.98
DW ratio: 2.5382

As shown in Table 1, the coefficient of determination is 0.65510, and DW ratio is 2.5382, and the probability is 0.000. Therefore, the regression model is significant. The estimation of regression coefficients is shown in Table 2.

TABLE 2 ESTIMATION OF REGRESSION COEFFICIENTS

	Influence	Degree	Efficiency	Intercept
Partial regression coefficient	4.17 E+7	-7.300E+4	5.575E+5	-5.509E+5
Standard coefficient	82.896	-82.398	0.289	0.000
T value	8.536	-8.487	3.465	-3.455
Degree of Freedom	58	58	58	58
Prob.	0.000	0.000	0.001	0.001
Correlation coefficient	0.4756	0.470	-0.018	
Partial correlation coefficient	0.746	-0.744	0.414	

The estimation of regression coefficients in 2004 has been calculated also. Based upon Table 2 and the calculation of 2004, influence, degree, and efficiency are significant. Degree is negative, while influence and efficiency are positive. Degree is an index of number of stockholdings, and influence is an index of structural positions, then the hypothesis of "better structural position, higher performance" is upheld.

The number of influence is decreasing. And the number of degree and efficiency are increasing. It means that the importance of structural position is still effective, but the capital relationship with other firms and the managing cost of investment is becoming more and more important because of the membership keiretsu is loosening.

V. CONCLUSION AND FUTURE WORKS

In this paper, the new approach of IDE analysis is proposed, and its validity is proved. We ascertained that (1) the centrality of Keiretsu is becoming low, (2) not all network indices, only influence, degree, and efficiency has an impact on profits. (3) The importance of influence is becoming lower, and the degree and efficiency is becoming higher due to the trend in the loosening of Keiretsu. However, additional research advocated developing our knowledge on rational capital relationships with other parts suppliers.

Acknowledgment: This research was partially supported by JSPS KAKENHI Grant Number 24510217.

REFERENCES

[1]Dyer H. J. (1996) "Specialized Supplier Networks as a Source of Competitive Advantage: Evidence from the Auto Industry", *Strategic Management Journal*, Vol. 17, 271-291

[2]Fukuoka S., Ito T., Passerini K. and Sakamoto M. (2006) An Analysis between Transaction and Cross Shareholdings in the Keiretsu of Nissan, *Managing Information in the Digital Economy Issues & Solutions*, 163-169, IBIMA International Conference, Bonn Germany

[3]Ito T., Tagawa S., Matsuno S., Uchida Y., Sakamoto M., Ikeda S., Mehta R. (2013) An Analysis of Network Structure in Mazda's Yokokai using the DEC Spatial Model, Proceedings of the 2013 International Conference on Business Administration, Marketing, and Economics (MAME 2013), Recent Advances in Business Administration, Marketing and Economics, pp.77-81, September 28-30, 2013, Venice, Italy

[4]JAPIA&APPC (1997, 2004), *Japanese Automotive Parts Industry*, Automotive Parts Publishing Company, (Japanese Edition)

A Comparative Study of Inter-firm Transactional Relationship between Kyohokai and Yokokai

M. Sakamoto[4], M. Hasama[1], T. Ito[1], Y. Uchida[1], R. Mehta[2], Y. Park[3], and S. Ikeda[4]

1 Dept. of Business Administration, Ube National College of Technology, Yamaguchi, Japan
(Tel : 81-836-35-7115; Fax : 81-836-35-7115)
(hasama@ube-k.ac.jp, ito@ube-k.ac.jp, uchida@ube-k.ac.jp)
2 School of Management, New Jersey Institute of Technology, U.S.A
(Tel :1-973-596-6419; Fax :1-973-596-3074)
(mehta@njit.edu)
3 Dept. of Business Administration, Prefectural University of Hiroshima, Japan
(Tel : 81-82-251-9826; Fax :-)
(ecventure@pu-hiroshima.ac.jp)
4 Dept. of Computer Science and Systems Engineering, University of Miyazaki, Japan
(Tel : 81-985-58-7392; Fax : 81-985-58-7392)
(sakamoto@cs.miyazaki-u.ac.jp)

Abstract: Although two major Japanese manufacturers, Toyota and Mazda, compete in the automotive market, their corporate performance, respectively, is different because the internal resources and the way they coordinate inter-relationships among their network partners is known to be divergent. Both these two companies have their own collaborating organizations, Kyohokai for Toyota and Yokokai for Mazda. Kyohokai, in contrast to Yokokai, is a central-oriented organization with strong interactive cooperation among network partners. Consistent with the stream of research on the coordination of internal resources and external inter-firm relationships, the purpose of this research is to identify and contrast the best practices of Toyota to Mazda using a comparative approach. This paper reviews relevant l literature on network organizations to focus on the differences between Kyohokai and Yokokai. Specifically, it makes a contribution by proposing a new perspective to identify the determinants of corporate performance and clarify the difference among the external inter-firm's relationships between Kyohokai and Yokokai, thus ascertaining the rational structure of network organizations.

Keywords: out-influence, in-degree, sales, the OIS model, Kyohokai, Yokokai, Keiretsu

I. INTRODUCTION

Although two major Japanese manufacturers, Toyota and Mazda, compete in the automotive market, their corporate performance, respectively, is different because the internal resources and the way they coordinate inter-relationships among their network partners is known to be divergent. Both of these two companies have their own collaborating organizations, Kyohokai for Toyota and Yokokai for Mazda. Kyohokai, in contrast to Yokokai, is a central-oriented organization with strong interactive cooperation among network partners. Consistent with the stream of research on the coordination of internal resources and external inter-firm relationships, the purpose of this research is to identify and contrast the best practices of Toyota to Mazda using a comparative approach. This research makes a contribution to the literature by proposing a new perspective to identify the determinants of corporate performance and clarify the difference among the external inter-firm's relationships between Kyohokai and Yokokai, thus ascertaining the rational structure of network organizations.

This paper is organized as follows. In Section 2, we briefly review some relevant researches of quantitative approaches of Keiretsu. Section 3 introduces the OIS spatial model. Section 4 shows the results, and discusses the implications of the results. Finally, Section 5 provides conclusions and summarizes this study.

II. BACKGROUND

Relationship, one of the important factors in structural analysis, has basically been analyzed using quantitative and qualitative approaches. Nowadays, quantitative approach is widely used due to the development of computer technology and mathematical theories. Quantitative approaches are required to ascertain the structure of any given network. Inter-organizational relationships in the Keiretsu have been analyzed with quantitative analysis tools such as CONCOR and other statistical methods [1].

Furthermore, Fukuoka et al. reported a new finding in relationships among member firms in Nissan's Keiretsu using transactions and cross shareholdings [2]. Recently Ito et al. begin to apply graph theory to network organization analysis, and clarifies some characteristics such as centrality, size of network [3-9]. A comparative study is one of the useful tools to uncover the common characteristics and differences in successful cases. After reviewing the relevant literature of network organizations, we found that only a few comparative studies have been published. Therefore, we focus on the common successful factors and differences between Kyohokai and Yokokai in this paper.

III. METHOD

In order to identify the determinants of supplier revenues, we calculated the network indices including influence, degree, effective size, efficiency, constrain, hierarchy, and density. We found that only influence and degree are significant. Thus, we analyzed the relationship between two significant explanatory variables: out-influence and in-degree, and explained variable: sales, which is one measure of corporate performance. Therefore, we propose a new approach, known as the OIS spatial model, to identify the common factors and also to illustrate the differences between Kyohokai and Yokokai.

1. Outline of the OIS model

Generally, a graph consists of a set of nodes and a set of arcs. Two nodes are connected if a path between these two nodes. Path is one of the important concepts in graph theory. According to graph theory, a path is a sequence of nodes such that the nodes and the arc are adjacent. A walk is a sequence of nodes and arcs such that the nodes and arcs are adjacent. The difference between path and walk is that a path is a walk that does not include any node twice, except that it its first node might be the same as its last. For digraphs, walks can travel arcs only in the direction of the arrows.

In general, influence means one kind of powers to affect persons or events, or causing something without any direct or apparent effort. Influence reflects the power to influence or have an impact on other member firms directly and indirectly in a network. Consequently, influence will be divided into two parts: direct influence and indirect influence in a network. Suppose that A is the matrix of the direct network, and A^n means the indirect influence from one node to another node by n steps. Then, influence is calculated as follows.

$$T = A + R = A + A^2 + A^3 + \cdot \ \cdot \ \cdot + A^n$$
$$= A(I - A)^{-1}$$

where

T: Total influence;
A: direct influence;
R: indirect influence;
I: Identity matrix.

Because transaction network is asymmetric, influence will be divided into two parts: out-influence and in-influence.

An integral index known as centrality has many facets that include degree, closeness, and betweenness. Degree expresses a firm's potential communication activity. In cross shareholding network, degree includes two categories: in-degree and out-degree, because cross shareholding networks are considered to be asymmetric organizations. In-degree refers to a firm accept investment from other member firms, whereas out-degree reflects a firm that only buy stocks of other firms within the network. Degree is calculated as below.

$$C_D(p_k) = \sum_{i=1}^{n} a(p_i, p_k)$$
$$i = 1, 2, \cdots, n; \quad k = 1, 2, \cdots, n$$

where

$a(p_i, p_k) = 1$; if and only if p_i and p_k are connected by a line

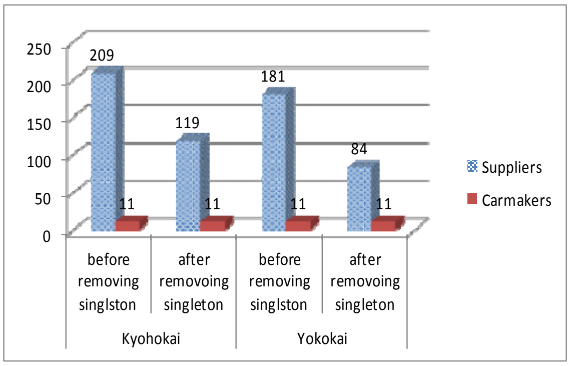
= 0; otherwise

Like influence, degree also will be divided into out-degree and in-degree. In this paper, percent data of inter-firm's transactions are collected from Yokokai.

2. Data collection

In order to ascertain the relationship between network indices and corporate performance, transaction data in the keiretsu of Kyohokai and Yokokai have been drawn from personal interviews as well as the publications of the Japan Auto Parts Industries Association and Automotive Parts Publishing Company [10].

In 2002, 209 parts suppliers, and 11 carmakers are included in Kyohokai. A singleton can be found in any network. Moreover, a singleton is defined as the node exists independently, without any relationship with other nodes in this paper. Therefore, a singleton is an isolated node in a given network generally. In Kyohokai, 90 singletons included. The detailed information of the numbers of parts-supplies, and carmakers in Kyohokai



and Yokokai in 2002 are shown in Fig. 1.

Fig. 1 numbers of financial institution, parts

The transactional relationships among the companies were identified through graph modeling. A tie shows the percentage of the transaction between each pair of firms. We collected directed and weighted data to measure out-influence and in-degree of each firm. The transactional network of inter-firm relationships in Kyohokai and Yokokai could be shown as Fig. 2 and 3 respectively.

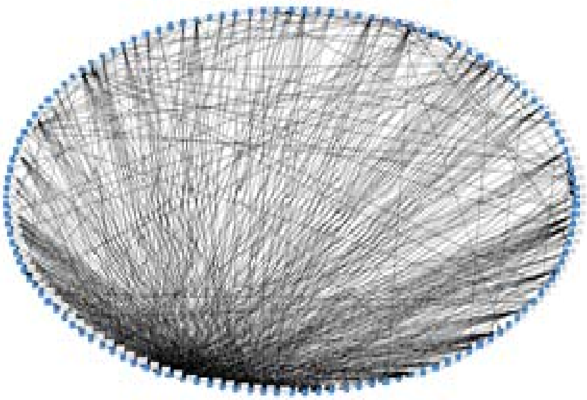


Fig. 2 Transactional network in 2002 in Kyohokai

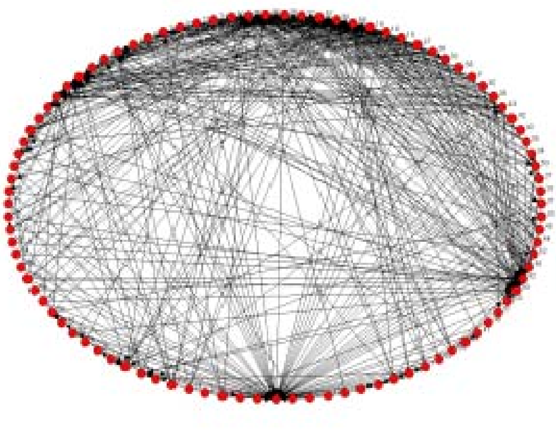


Fig. 3 Transactional network in 2002 in Yokokai

IV. RESULTS AND DISCUSSIONS

1. Centrality analysis

The overall centrality of Kyohokai and Yokokai is calculated as follows.

Table 1 Out-degree and in-degree of Kyohokai and Yokokai

	Out-degree	In-degree
Kyohokai (Toyota)	0.341%	23.326%
Yokokai (Mazda)	49.915%	1005.184%

Form Table1, it is evident that out-degree and in-degree of Yokokai are larger than of Kyohokai. High centrality means the suppliers purchase their parts from a couple of specific companies. It also means that compared with Yokokai, the companies in Kyohokai have more chances to purchase their parts. This could possibly be the main reason why Kyohokai's firm is more efficient. Consequently, the hypothesis that "the lower the centrality, the higher the performance" holds.

2. Determinants of sales

We calculated partial correlation coefficient of the seven indices in multiple regression models, and found effective size, density and other 3 indices are not significant. Thus, we removed those 5 variables and reanalyzed the regression model. The results are displayed in Table 2.

TABLE 2 RESULTS OF THE VARIANCE ANALYSIS

	Sum of Squares	Degree of Freedom	Mean Square	F Test
Total	3.16E+13	81		12.722
Regression	7.70E+12	2	3.85E+12	Prob.
Residual	2.39E+13	79	3.02E+11	0.000

Coefficient of determination: 0.24361
Multiple correlation coefficient: 0.493561
Adjusted R-square: 0.47378
AIC: 2,415.89
DW ratio: 2.1997

From Table 2, the coefficient of determination is 0.243610, and DW ratio is 2.1997. The probability is 0.000. Therefore, the regression model is significant. The estimated regression coefficients are shown in Table 3.

Based upon Tables 3, out-influence, and in-degree are significant. Out-influence is negative, and in-degree is positive. Out-influence is an index of the structural position of selling its parts, and in-degree an index of purchasing parts from other firms, then the hypothesis of "more purchasing power is associated with higher performance" is upheld.

TABLE 3 ESTIMATION OF REGRESSION COEFFICIENTS

	Out-Influence	In-Degree	Intercept
Partial regression coefficient	-3.550E+5	6.489E+3	4.45187E+5
Standard coefficient	-0.1697	0.4672	0.000
T value	-1.7337	4.7736	2.557
Degree of Freedom	79	79	79
Prob.	0.0869	0.000	0.0125
Correlation coefficient	-0.1595	*0.4635*	
Partial correlation coefficient	-0.0971	0.4732	

Basically, high purchasing power means a specific firm is highly capable of producing and selling component parts.

V. CONCLUSION AND FUTURE WORKS

In this paper, the new approach of OIS model was proposed. This investigation revealed that (1) "lower centrality, higher performance" linkage holds. (2) Out-influence has a negative impact on sales while in-degree has a positive impact on sales. However, additional research is suggested to develop the theory and knowledge base on rational relationship with other parts suppliers. For instance, it is necessary to clarify the individual firm's position in transactional network to improve its efficiency. The linkage between structural position of each individual firm and their performance is also advocated.

Acknowledgment: This research was partially supported by the Ministry of Education, Culture, Sports, Science, and Technology, Grant-in-Aid for Exploratory Research, 21510171, 2009.

REFERENCES

[1]Lincoln R. L. and Gerlach M.L. (2004) *Japan's Network Economy Structure, Persistence, and Change,* Cambridge University Press

[2]Fukuoka S., Ito T., Passerini K. and Sakamoto M. (2006) An Analysis between Transaction and Cross Shareholdings in the Keiretsu of Nissan, *Managing Information in the Digital Economy Issues & Solutions*, 163-169, IBIMA International Conference, Bonn Germany

[3] T. Ito, K. Passerini, M. Sakamoto (2008) Structure Analysis of Keiretsu of Toyota, Encyclopedia of Networked and Virtual Organizations, pp.1542-1548, Idea Group Publishing

[4]T. Ito, C. Medlin, K. Passerini, M. Sakamoto (2009) Influence Trust and Trade in the Keiretsu of Toyota: A Centrality *Analysis*, *Trust*, *Globalisation d Market Expansion*, Chapter 8, pp.101-118, Nova Science

[5]Ito T., Matsuno S., Xia Z., Sakamoto M., and Rajiv Mehta (2010) An Analysis of Interactive Influence in Mazda's Yokokai Keiretsu, Artificial Life and Robotics, Volume 15, Number 3, 249-252, Springer Japan

[6]Ito T., S. Matsuno, R. Mehta, M. Sakamoto, S. Ikeda (2013), An analysis of inter-firms distances and corporate performance using DEC spatial model, Recent Advances in Energy & Environment Integrated Systems, Proceedings of the 2nd International Conference on Integrated Systems and Management for Energy, Development, Environment and Health (ISMAEDEH '13), pp.80-85, April 23-25, 2013, Morioka City, Iwate, Japan

[7]Ito T., S. Tagawa, R. Mehta, M. Sakamoto, S. Ikeda (2013), A research on inter-firms relationship using network-based DEA –A case study of Mazda Yokokai–, Recent Advances in Energy & Environment Integrated Systems, Proceedings of the 2nd International Conference on Integrated Systems and Management for Energy, Development, Environment and Health (ISMAEDEH '13), pp.92-97, April 23-25, 2013, Morioka City, Iwate, Japan

[8]Ito T., M. Hasama, R. Mehta, M. Sakamoto, S. Ikeda (2013), Measuring efficiency of firms in Yokokai using the IDCBI spatial model, Recent Advances in Energy & Environment Integrated Systems, Proceedings of the 2nd International Conference on Integrated Systems and Management for Energy, Development, Environment and Health (ISMAEDEH '13), pp.116-122, April 23-25, 2013, Morioka City, Iwate, Japan

[9]Ito T., Tagawa S., Matsuno S., Uchida Y., Sakamoto M., Ikeda S., Mehta R. (2013), An Analysis of Network Structure in Mazda's Yokokai using the DEC Spatial Model, Proceedings of the 2013 International Conference on Business Administration, Marketing, and Economics (MAME 2013), Recent Advances in Business Administration, Marketing and Economics, pp.77-81, September 28-30, 2013, Venice, Italy

[10]JAPIA&APPC (2005) *Japanese Automotive Parts Industry*, Automotive Parts Publishing Company, (Japanese Edition)

A Path Analytic Model and Measurement of the Relationships between Green Supply Chain Management Implementation and Corporate Performance

S. Matsuno*, M. Hasama, Y. Uchida, and T. Ito

Ube National College of Technology, 2-14-1 Tokiwadai, Ube, Yamaguchi, Japan
**matsuno@ube-k.ac.jp*

Abstract: This paper proposes and empirically tests a model of the relationships between green supply chain management (GSCM) activities and corporate performance. From the literature review, we identified five constructs, namely, environmental commitment, supplier collaboration, supplier assessment, information sharing among suppliers, and business process improvement. These explanatory variables are used to form a structural model explaining the environmental and economic performance. The model was analyzed using the data from a survey of sample of manufacturing firms in Japan. The results suggest that the degree of supplier collaboration has an influence on the environmental performance directly. While, the impact of supplier assessment on the environmental performance is mediated by the information sharing and/or business process improvement. And, the environmental performance has a positive relationships on the economic performance.

Keywords: Green supply chain management, Corporate performance, Empirical study, Path modeling

I. INTRODUCTION

Green supply chain management (GSCM) is an inter-firm collaborative initiative concerning environmental load reduction and environmental preservation throughout the entire life cycle of a product. Against a backdrop of an international trend of tightening regulations on specific chemical substances, as typified by the Restriction of Hazardous Substances (RoHS) Directive and Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) in Europe, as well as social demands for environmental load reduction and preservation, a heightening of environmental awareness among consumers, and other factors, many firms have engaged in GSCM activities in recent years [1]-[8].

However, along with the pursuit of economic efficiency throughout the entire supply chain, which is the main focus of conventional supply chain management, GSCM also seeks to achieve environmental load reduction, and many issues related to GSCM require investigation, for example, examination of cost effectiveness. In addition, the development of inter-firm information systems and information transfer format standardization are urgently required; REACH in Europe, for instance, requires the communication and sharing of chemical substances information between agents/partners in supply chains.

This paper is organized as follows. In Section 2, we briefly review some relevant researches about the effectiveness of GSCM implementation. We analyze the status of GSCM implementation at firms in Japan on the basis of our original questionnaire survey and also consider the relationship between GSCM activities and performance, focusing on the role of inter-firm information sharing in Section 3. Section 4 presents a structural equation model concerning factors that affect GSCM performance and use path analysis to test the validity of the model. Finally in Section 5 we conclude by a summary of this paper.

II. LITERATURE REVIEW

A previous research concerning effects obtained from the implementation of GSCM is the empirical study that Zhu and Sarkis (2004) conducted targeting 281 manufacturing firms in China [7]. They used statistical analysis methods to test effects that GSCM has on environmental and economic performance and whether quality control and JIT activities have a moderator effect on performance. Although the findings showed that GSCM implementation has a positive effect on a firm's environmental performance, they did not provide sufficient conclusion to measure whether the effect on economic performance is positive or negative.

Large and Thomsen (2011) tested the effect that GSCM has on environmental and purchasing performance [2]. After identifying factors that affect GSCM performance on the basis of literature review, they used survey data on purchasing firms in Germany to perform structural equation modeling and tested the validity of their model. Their findings showed that green supplier assessment and active collaboration have a positive effect on environmental performance, and that they are promoted according to the strength of the purchasing department's strategic role and commitment to environmental management. The findings also indicated that improvement in environmental performance has a positive effect on purchasing performance.

In light of the abovementioned previous research, we engage in further explorative discussion concerning GSCM activities and performance at firms in Japan, focusing on the role that inter-firm information sharing plays in GSCM implementation.

III. DATA AND ANALYSIS

1. Outline of the survey

We conducted a mail questionnaire survey from November 2010 to February 2011 with the research objective of investigating "the current state and issues in supply chain greening and inter-firm collaboration". Since ordinarily the introduction and implementation of GSCM is likely to be affected by firm size, we targeted Japanese manufacturers with 500 or more employees in this survey. The number of valid responses obtained from randomly sampled firms was 126, but detailed results are omitted here [9][10].

2. GSCM implementation and internal management

First, we confirmed the status of implementation of concrete inter-firm cooperation activities in GSCM (see Table 1). We were able to ascertain that many firms provide information and exchange opinions concerning environmental regulations and cooperate concerning the prohibition or reduction of use of regulated chemicals contained in raw materials and components. In addition, we found that some progressive firms implemented the detailed GSCM activities cooperate concerning eco-design and engage in joint research or joint development of eco-friendly products or materials with business partners under inter-firm cooperation and collaboration.

Then, we confirmed the state of development of internal management systems for overall environmental management. On the whole, high scores were obtained for all questions (based on a 6-point scale), and it was found that the introduction of EMS and development of organizations in charge of environmental management are particularly advanced. It is likely that this state of internal management systems development shows the strength of overall commitment to environmental management, including the implementation of GSCM.

Table 1. Status of GSCM implementation (n = 124)

	Mean	*S.D.*
We provide information and exchange opinions concerning environmental regulations.	4.02	1.32
We cooperate concerning the prohibition or reduction of use of regulated chemicals contained in raw materials and components.	4.47	1.46
We cooperate in the development of resource-saving and resource-recycling manufacturing processes.	3.50	1.21
We cooperate in improving the efficiency of distribution and transport systems (joint distribution or modal shifts).	3.65	1.42
We cooperate in the review of packaging and the reuse and recycling of containers and packaging.	3.87	1.18
We cooperate concerning eco-design (eco-friendly design).	3.49	1.24
We engage in joint research or joint development of eco-friendly products or materials with business partners.	3.31	1.42

Note: 1 = We do not engage in it at all. 6 = We engage in it very actively.

3. Supplier assessment

We analyzed what sort of requirements manufacturing firms have for suppliers concerning implementation of GSCM and, on the other hand, what sort of assistance they provide.

We confirmed that many firms have a variety of requests for suppliers. For example, request that suppliers understand each firm’s environmental philosophy and requiring them to observe a code of conduct or guidelines. Also, although the scores are comparatively low, there are firms that conduct on-site environmental inspections of suppliers and engage in continuous auditing and monitoring, and support the development of suppliers' EMS.

4. GSCM implementation and information sharing among firms

We examined the status of information communication/exchange activities through information networks among firms. Since ordinarily inter-firm information communication/exchange through information networks plays a major role in the effective operation and implementation of conventional supply chain management, it can be expected to play a similar role in GSCM.

Information communication/exchange between firms via information networks occurs mainly with respect to information about purchase order placement/receipt or suppliers/customers, and information about billing, payment, and settlement. At the same time, we found that detailed communication/exchange between firms also occurs with respect to information about components and measured values of regulated chemicals.

The background to this is likely progress in preparations for and compliance with REACH, which went into effect in June 2007 in Europe. REACH requires communication and sharing of chemical substances information between agents in supply chains. Confirmation for matters such as the presence or absence and the content of chemical substances subject to regulation in molded articles is required, as is such information from manufacturers as applications and amount used by application for each chemical substance.

5. GSCM implementation and performance

We confirmed awareness among firms of the status of matters such as business process improvement accompanying GSCM implementation. And we confirmed awareness among firms of improvements in environmental and economic performance resulting from GSCM implementation.

Here, we consider quantitative reduction of various types of waste, regulated chemicals, and the like to be environmental performance, and reduction of various costs, such as purchasing and distribution costs, to be economic performance. Although it is of course necessary to also examine improvement in economic performance on the basis of financial indicators such as sales and profit margins, it is generally difficult to eliminate from these financial indicators the impact of factors other than GSCM implementation. Accordingly, in this paper we perform analysis in which the concept of economic performance is limited to the abovementioned cost reductions.

On the whole, above-average scores are given for both environmental and economic performance improvement. However, firms do not enjoy the reduction in purchasing and distribution costs.

IV. PATH ANALYSIS CONCERNING GSCM ACTIVITIES AND PERFORMANCE

In light of the above analysis, we present a structural equation model concerning factors that affect GSCM performance [11]. The hypothetical path model is shown in Fig. 1. The model uses the seven composite variables that is to say, Environmental commitment (5 items), Supplier collaboration (7 items), Supplier assessment (6 items), Information sharing among suppliers (3 items), Business process improvement (3 items), and Environmental performance (5 items) as explanatory variables and Economic performance (4 items) as the dependent variable, and assumes a causal chain at multiple stages. We assume a positive path in all relationships between factors.

Next, we provide an overview of the path analysis. Here, we omit the factor analysis process for examining the validity of the assumed latent variables and use the value obtained by dividing the total of the values of the items pertaining to each latent variable by the number of items. In other words, we use composite scales to represent the latent variables. Table 2 shows the descriptive statistics of each variable. For each scale, Cronbach's α is greater than 0.8, and reliability (internal consistency) is high.

We empirically test the path model using these variables, performing sequential multiple linear regression analysis using the standard partial regression coefficient as the path coefficient.

When we took Environmental commitment, Supplier collaboration, and Supplier assessment as

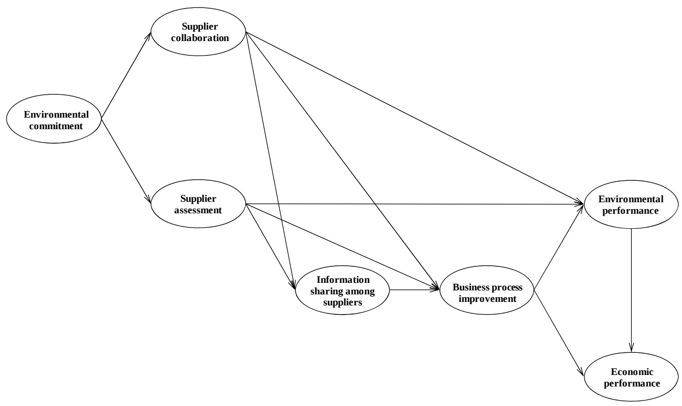


Fig.1. Hypothetical path model

Table 2. Descriptive statistics of composite variables

	# of items	*α*	*Mean*	*S.D.*
Environmental commitment	5	0.93	4.75	0.99
Supplier collaboration	7	0.89	3.77	1.03
Supplier assessment	6	0.88	3.64	1.19
Information sharing among suppliers	3	0.81	2.70	1.31
Business process improvement	3	0.87	3.34	1.05
Environmental performance	5	0.89	3.48	1.16
Economic performance	4	0.86	3.07	1.13

Note: each score on a 6-point scale

explanatory variables and Information sharing among suppliers as the explained variable, no significant positive path from Supplier collaboration was found. That is to say, some non-significant paths were found in the causal chain between factors that affect GSCM performance. However, all other paths were significant at the 1% or 5% level. Also, all coefficients of determination were significant, and the relationships between factors assumed in this paper were almost supported.

V. CONCLUSION

This paper proposes and empirically tests a model of the relationships between GSCM activities and corporate performance. We identified five constructs, namely, environmental commitment, supplier collaboration, supplier assessment, information sharing among suppliers, and business process improvement.

The results suggest that the degree of supplier collaboration has an influence on the environmental performance directly. While, the impact of supplier assessment on the environmental performance is mediated by the information sharing and/or business process improvement. And, the environmental performance has a positive relationships on the economic performance.

ACKNOWLEDGMENTS

This research was partly supported by the Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (C) 24530445 and authors would like to thank the organization.

REFERENCES

[1] Hoejmose S, Brammer S, and Millington A (2012), Green supply chain management: The role of trust and top management in B2B and B2C markets. Industrial Marketing Management 41(4):609-620

[2] Large R and Thomsen C (2011), Drivers of green supply management performance: Evidence from Germany. Journal of Purchasing and Supply Management 17(3):176-184

[3] Lopez-Gamero M, Molina-Azorin J, and Claver-Cortes E (2009), The whole relationship between environmental variables and firm performance: Competitive advantage and firm resources as mediator variables. Journal of Environmental Management 90(10):3110-3121

[4] Matsuno S, Hasama M, and Ito T (2011), Greening of the supply chain and its impact on information sharing among firms based on the transaction cost economics (in Japanese). Transactions of the Japan Society for Production Management 17(3):11-16

[5] Testa F and Iraldo F (2010), Shadows and lights of GSCM: determinants and effects of these practices based on a multi-national study. Journal of Cleaner Production 18(10-11):953-962

[6] Yang S, Lu C, Haider J, and Marlow P (2013), The effect of green supply chain management on green performance and firm competitiveness in the context of container shipping in Taiwan. Transportation Research Part E: Logistics and Transportation Review 55:55-73

[7] Zhu Q and Sarkis J (2004), Relationships between operational practices and performance among early adopters of green supply chain management practices in Chinese manufacturing enterprises. Journal of Operations Management 22(3):265-289

[8] Zhu Q, Sarkis J, and Lai K (2012), Green supply chain management innovation diffusion and its relationship to organizational improvement: An ecological modernization perspective. Journal of Engineering and Technology Management 29(1):168-185

[9] Matsuno S, Mikami Y, and Ito T (2012), An analysis of green supply chain management practices and its performance focused on the role of information sharing among firms (in Japanese). Transactions of the Japan Society for Production Management 18(2):11-16

[10] Matsuno S and Tokinaga S (2012), Greening of the supply chain and its impact on corporate networks: Analysis based upon a questionnaire survey (in Japanese). Journal of Information and Management 32(4):23-30

[11] Matsuno S, Hasama M, Uchida Y, and Ito T (2013), Green supply chain management activities and corporate performance: Evidence from Japan. Proc. of the 1st International Conference of Production Management, in Ho Chi Minh, Vietnam

A Study of Open Source Cloud System for Small and Medium Enterprise

Y. Uchida[1], S. Matsuno[1], T. Ito[1], M. Hasama[1], and M. Sakamoto[2]

[1] *Dept. of Business Administration, Ube National College of Technology, Ube 755-8555, Japan*
Tel : 81-836-35-7567; Fax : 81-836-35-7567
uchida@ube-k.ac.jp, matsuno@ube-k.ac.jp, ito@ube-k.ac.jp, hasama@ube-k.ac.jp
[2] *Dept. of Computer Science and Systems Engineering, University of Miyazaki, Miyazaki 889-2192, Japan*
Tel : 81-985-58-7392; Fax : 81-985-58-7392
sakamoto@cs.miyazaki-u.ac.jp

Abstract: The use of cloud services among SMEs remains low, and R&D and technical support for SMEs is an urgent problem. Even so, there have been almost no academic studies on the relationship between cloud computing and SME information infrastructure, or on the future direction of this. With this background, we considered and designed a cloud system framework suitable for SMEs, built a prototype model and investigated the ease of implementation. Services that deliver functionality equivalent to Amazon EC2 which is representative cloud service but delivered through other technologies are called EC2 clones. OpenStack is not EC2 clone, and is a typical open source project for providing Infrastructure as a Service (IaaS) type cloud services. OpenStack was used in this study to implement a simple cloud service model.

Keywords: Cloud, Open Source, Small and Medium sized Enterprise, OpenStack

I. INTRODUCTION

Today's information systems must be able to continue stable operation even during crises such as the recent Great East Japan Earthquake, or can be quickly restored to operation. This is true even for the information infrastructure of small and medium sized enterprise (SME). This means that an ICT department business continuity plan (BCP) [1], designed to continue operations under disaster and other emergency situations, is essential to ensure preventative measures to minimize impact and to quickly restore the most important and most essential operations.

The "Survey on the State of Cloud Usage by Small and Medium Sized Enterprise" (2011) [2], showed that about 10% of SMEs "use cloud computing," highlighting the urgent problem for R&D and technological support for SMEs. Even so, there have been almost no academic studies on the relationship between cloud computing and SME information infrastructure, or on the future direction of this.

Previously, we worked on a research project called "Open Source Information System Research and Development for Supporting the IT Strategy of Small and Medium Sized Enterprise" [Grant-in-Aid for Scientific Research (C)] (2008-2010), and proposed a model open source information system to support the IT strategy of SMEs, and together with developing a prototype, considered usage-related problems. Based on the our findings from that research project, in this study, we consider and design a cloud system framework appropriate for small to medium sized enterprise, and investigate and consider the issues related to operation and administration.

II. CLOUD SYSTEMS

1. Cloud computing

The United States National Institute of Standards and Technology (NIST) defines "cloud" as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources, that can be rapidly provisioned and released with minimal management effort or service provider interaction [3].

Cloud computing includes the following five characteristics:

- On Demand and Self Services
- Broad Network Access
- Resource Pooling
- Rapid Elasticity
- Measured Services

Services can be categorized into the following three service types:

- SaaS (Software as a Service: Functionality is delivered over the network)
- PaaS (Platform as a Service : Application development environment and customization features are delivered over the network)

- IaaS (Infrastructure as a Service: Even virtual machines and operating systems are delivered over the network)

2. Cloud research trends

"Above the Clouds: A Berkeley View on Cloud Computing," [4] a report of the Reliable Adaptive Distributed Systems Laboratory (RAD Lab) of the United States, highlighted ten obstacles facing cloud computing, including availability of service, data lock-in, and data confidentiality and auditability.

In a Gartner report called "Survey of 2,014 global CIOs," [5] when asked to predict when more than half of the information processing of the computers of their companies will be moved to the cloud, 53% of global respondents said "by 2015," but for Japan, only 25% gave the same response. Nineteen percent of global respondents said "after 2021" or "not possible to migrate," while the response was 43% for Japan. This shows that Japan is very late to the game in areas such as cloud research and cloud implementation.

On the subject of research trends in Japan, other than a cloud and public infrastructure study by the Information-technology Promotion Agency [6], most are studies related to security [7].

3. Cloud computing

Cloud systems can be categorized as shown in Fig.1. In particular, depending on the cloud system service architecture, they are configured according to the following three subsystems: SaaS (Software as a Service: functions are provided over the network), PaaS (Platform as a Service: application development environment, customization functions are provided over the network), and IaaS (Infrastructure as a Service: even virtual machines and operating systems are provided over the network). For open source cloud system implementations, OpenStack Compute (Nova) is an example of a resource toolset, while OpenStack Object Storage (Swift) is an example of a cloud storage toolset. Even SMEs can easily use these open source projects to develop tools that provide functionality similar to (functional compatibility) Google Apps or Amazon EC.

4. Amazon EC2

Amazon EC2 is a good example of IaaS. Amazon EC2 [8] stands for "Amazon Elastic Compute Cloud." It

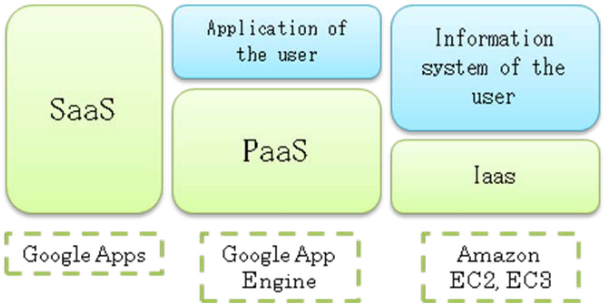


Fig.1. Cloud systems

refers to a web service that presents computing capabilities in the cloud and that are easily scalable. It can be used according to the following processes.

- Immediately launch and execute by selecting a preconfigured template image. Or, create an Amazon Machine Image (AMI) that includes applications, libraries, data, and related settings.
- Configure security and network access on Amazon EC2.
- Select the desired instance type and operating system, and using web service APIs or the various administration tools that are provided, start, stop and monitor the AMI instance as much or as little as needed.
- For each instance, decide whether to run from multiple locations, use a fixed IP end point, or add robust block storage.
- Pricing is based on actual resources consumed, such as total time the instance was run or amount of data transferred.

Of these processes, the AMI selection process is shown in Fig.2.

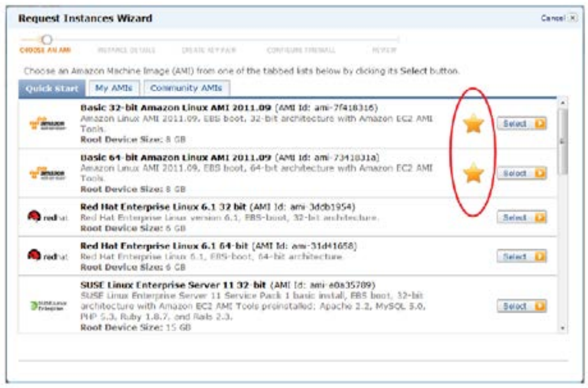


Fig.2. AMI selection

5. OpenStack

OpenStack [9] is a representative open source to offer IaaS type cloud service. OpenStack is IaaS cloud computing project begun in 2010 by Rackspace Inc. [10] and NASA (National Aeronautics and Space Administration) [11]. In 2012, nonprofit organization "OpenStack Foundation" [12] was established. A component to operate it in GUI includes OpenStack Dashboard (horizon) to show it in Fig.3, and OpenStack manages compute, storage, and networking resources. The most recent version at the time of 2013-11-21 of OpenStack is “OpenStack 2013.2 (Havana)”.

Fig.3. OpenStack Dashboard

III. RDO

RDO [13] is a community of people using and deploying OpenStack on Red Hat and Red Hat-based platforms. Community site openstack.redhat.com of RDO was established like Fig.4.

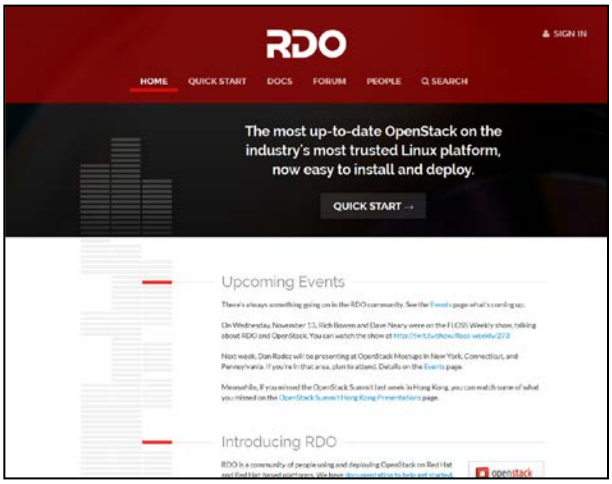


Fig.4. openstack.redhat.com of RDO

IV. Simple cloud model by OpenStack

1. Prototype system

Table 1 shows the specifications of the prototype system tested in this study.

Table 1. Prototype system specifications

Form factor	Space saving desktop
CPU	Intel Core i3 processor (with virtualization support)
Memory	4GB
Hard disk	200GB
OS	CentOS 6.4（x86_64）
Linux kernel	2.6.32
Cloud platform software	OpenStack 2013.2 (Havana)

2. Installation of RDO

By installing RDO, the cloud environment can be used as follows.

・ Install of Software repository
sudo yum install -y http://rdo.fedorapeople.org/openstack/openstack-grizzly/rdo-release-grizzly-3.noarch.rpm

・ Install of the Packstack Installer
sudo yum install -y openstack-packstack

・ Install of OpenStack using Packstack
packstack –allinone

Now its single node OpenStack instance is up and running.

・ Visit the Dashboard
Log in to the Openstack dashboard at http://localhost/dashboard - the username is "admin".

・ Enabe SSH
Adding Port "22" on the default security group

・ Key pair
Create or impor a key pair

・ Add an image
"Fedora19" in the "http://cloud.fedoraproject.org/fedora-19.x86_64.qcow2" [14]

・ Launch the instance
In the main portion of the screen, click the "Launch" button for the "F19" image.

・ Associate Floating IP
In the main portion of the screen, followed by the "Associate Floating IP" link for the instance you ju

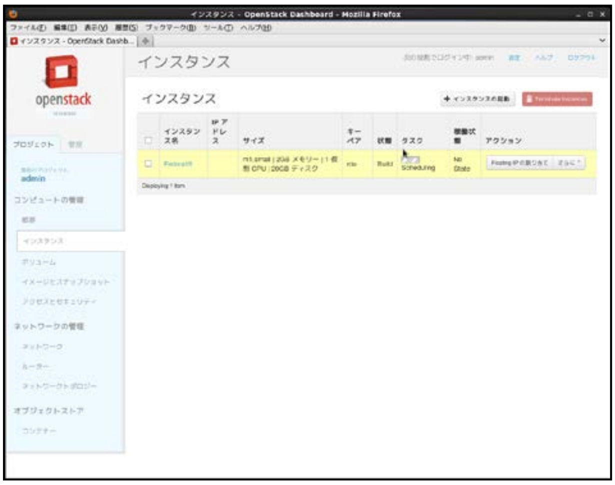
st launched.

・ SSH to the Instance

```
$ ssh -l root -i my_key_pair.pem floating_ip_address
```

The above steps enable an instance to be presented (see Fig.5).

Fig.5. A running instance



V. CONCLUSION

We considered and designed a cloud system framework suitable for SMEs, performed implementation testing and tested operation. In the future, it will be necessary to run the prototype system and investigate and consider issues related to operation and administration.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 24530444.

REFERENCES

[1] Asian Disaster Reduction Center (2012), BCP status of the Private Sector in the APEC Region 2012. Report:1-16, http://www.adrc.asia/publications/bcp/survey_2012.pdf

[2] Information-technology Promotion Agency (2011), Factual Investigation on Cloud Usage by Small to Medium Enterprise . Information-technology Promotion Agency (in Japanese)

[3] National Institute of Standards and Technology, The NIST Definition of Cloud Computing . http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

[4] Electrical Engineering and Computer Sciences (University of California at Berkeley) , Above the Clouds: A Berkeley View of Cloud Computing . Technical Report No.UCB/EECS-2009-28, http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.html

[5] Gartner (2011), Worldwide Survey of 2,014 CIOs (In Japanese). press release, http://www.gartner.co.jp/press/html/pr20110303-01.html

[6] Information-technology Promotion Agency (2013), Report – “Research on Cloud Computing and Public Infrastructure (in Japanese) . Information-technology Promotion Agency, Japan

[7] Hori et. al (2009), Cloud Computing Security Research Trends (in Japanese). Information Processing Society of Japan Report CSEC “Computer Security” 2009-CSEC-47(4):1-6

[8] Amazon, Amazon Elastic Compute Cloud (Amazon EC2). http://aws.amazon.com/jp/ec2/

[9] The OpenStack Foundation, OpenStack Open Source Cloud Computing Software. http://www.openstack.org/

[10] Rackspace, US Inc., Rackspace: The Leader in Hybrid Cloud. http://www.rackspace.com/

[11] National Aeronautics and Space Administration, NASA. http://www.nasa.gov/

[12] The OpenStack Foundation, Foundation » OpenStack Open Source Cloud Computing Software. http://www.openstack.org/foundation/

[13] Red Hat, Inc., RDO. http://openstack.redhat.com/

[14] The Fedora Project, Fedora Project Homepage. http://fedoraproject.org/en/

A Visualization of Patent Strategies in Japanese ICT Companies Based on Text-Mining

T. Tokumitsu[1], T. Okada[1], I. Nakaoka[2] and Y. Park[4]

1 Advanced Course of Management Information Engineering,
Ube National College of Technology, Yamaguchi, Japan
(tokumitsu0310@gmail.com, takuya.takuya.7391@gmail.com)
2 Dept. of Business Administration, Ube National College of Technology, Yamaguchi, Japan
(Tel : 81-836-35-5130)
(nakaoka@ube-k.ac.jp)
3 Dept. of Business Administration, Prefectural University of Hiroshima, Japan
(Tel :81-82-251-9826)
(ecventure@pu-hiroshima.ac.jp)

Abstract: Japanese ICT companies had kept strong competitive advantage until the collapse of the bubble economy which occurred in the early 1990s. Since then, most of the companies still have been struggling with poor performance. The companies have implemented various management reforms for the last twenty years in order to recover competitive advantages. Although effects by those attempts appeared for the short-term, recovery of the competitive advantage did not continue for the long-term. We propose following two big reasons for continued slump and repeated failures in their turnaround. First, the Japanese companies lagged behind changes of consumer needs. Second, they had implemented wrong technology innovation strategies. Therefore, we recognize the necessity to understand issues of technology innovation strategies of Japanese ICT companies based on patent documents recognized as an ample source of technical and commercial knowledge. In this paper, we disclose and analyze patent strategy of Japanese large ICT companies, Sony, Sharp and Panasonic in smart phone market by correspondence analysis and self-organizing maps.

Keywords: patent strategies, visualization, F-term, correspondence analysis, self-organizing maps

I. INTRODUCTION

Japanese ICT companies had sustained their competitive advantages until the early 1990s, but most of them have lapsed into poor business performances caused by the collapse of the bubble economy since 1991. In order to recover their competitive advantages, they have restructured their operations such as corporate downsizings, business reconstructions and re-examinations of long-term transactional relationships with their cooperative firms for last twenty years. However, the restructurings did not lead to long-term effects on their competitive advantages. The following three reasons are discussed. First, Japanese ICT companies have the innovator's dilemma by pursuing excessive qualities [1]. Second, the structural model of most digital consumer products has change from the integral architecture to the module architecture [2]. Third, they failed to capture values from their new products developed with their hard efforts because of the rapid price reduction in digital consumer product market [3]. We consider that Japanese ICT companies have implemented wrong technology innovation strategies. Impairments of their ability of research and development are unlikely because 28 of Japanese ICT companies are included in the world 100 major innovative companies in 2012, moreover Japanese major ICT companies such as Sony, Panasonic, Sharp, NTT and Nidec Corporation belong in the higher rank. Instead we hypothesize two reasons for their long-term slumps. Firstly, their technology innovation strategies have not led to their turnaround. Secondly, their major products did not changes from products whose needs were decreasing such as feature phone while leading products in ICT industry were changing to products in new field such as smart phone. In order to test the hypotheses, we need to compare Japanese ICT companies and some other foreign ones in technology innovation strategies.

As our exploratory approach, we focus on Japanese three ICT companies, Sony, Sharp and Panasonic, and clarify their patent strategies in smart phone market.

Since Apple launched iPhone in 2007, the market structure has dramatically changed from feature phone to smart phone. Therefore, we visualize characteristics and transitions of their patent strategies associated with smart phone since 2000, and discuss the results.

This paper is organized as follows. In Section 2, we briefly review some relevant researches of visualization of patent analyses. Section 3 shows our exploratory analyses by correspondence analyses and self-organizing maps and the results. Finally in Section 4 we describe conclusion of this paper.

II. BACKGROUND

Patent documents are an ample source of technical and commercial knowledge. The patent is one of the indicators of capacity for technological development. There are some researches aimed at visualizing and analyzing patents, or proposing efficient text-mining approaches for creating patent maps.

YG. Kim *et al* understand advances of emerging technologies and forecast its trend in the future in the following way. First, Keywords from patent documents related to a target technology field are extracted. Second, patent documents with the keywords are clustered by k-Means algorithm. Third, a semantic network of keywords is formed. Final, a Patent Map is created [4].

B. Yoon *et al*. note that citation analysis is subject to some crucial drawbacks and propose a network-based analysis, an alternative method for citation analysis. Such new indexes as technology centrality index, technology cycle index, and technology keyword clusters are suggested for in-depth quantitative analysis [5].

Y.H. Tseng *et al*. describes a series of text mining techniques that conforms to the analytical process used by patent analysts. These techniques include text segmentation, summary extraction, feature selection, term association, cluster generation, topic identification, and information mapping [6].

However, there are few researches focusing on management strategies based on text-mining analysis of patents. Therefore, this paper describes technology innovation strategies of Japanese ICT companies based on the analysis.

III. ANALYSIS

In this section, patents strategies of Japanese ICT companies are analyzed. In Japan, patent documents are archived in IPDL (Industrial Patent Digital Library), and are browsed and retrieved ubiquitously. All patents are classified according to each three classification codes, IPC(International Patent Classification), FI(File Index) and F-term(File Forming Term). Both of FI and F-term are classification codes for segmentalizing IPC. Especially, F-term is given to every patent from F-term list based on the technical items indicated in patent documents. Thus, almost every patent has multi-F-term code. This paper refers to patents with F-term related to smart phone described in "The technical development actual condition analysis investigation report about the smart phone according to patent document analysis (patent map)" [7]. Top 20 most frequently used F-terms in smart phone industry are shown in Table 1.

Table 1 Top 20 most frequently used F-terms in smart phone industry

F-term	Contents
5K127	Telephone function
5K027	Telephone circuits
5K067	Mobile radio communication systems
5K201	Telephonic communication services
5K023	Telephone set structure
5K101	Telephonic communication linked with other devices
5C122	Studio devices
5B084	Information transfer between computers
5E501	Digital calculator user interface
5C164	Two-way televisions, distribution of moving picture or the like
5K036	Display and transmission of telephone numbers
5B087	Position input by displaying
5B020	Input from keyboards or the like
5C022	Studio equipment
5K048	Selective calling equipment (for remote control and measurement)
5C025	Television transmitter and receiver circuitry
5C082	Control or circuits of indicating devices
5C087	Alarm systems
5B058	Card readers/writers or multifunction peripheral apparatus
5K024	Special services for exchanges 2

1. An approach based on the number of patent publications

As first our approach, the numbers of patent publications associated with smart phone in each of the three companies are shown in Figure1.

The numbers of patent publication in Sharp and Panasonic tend to increase slightly in Figure1. As contrasted to the two companies, the number of Sony increases steadily until 2009, since then, substantially decreases. The decreasing in the number of their patent publications is caused by a decreasing in the number of their patent applications. We consider that investment into research and development is decreased in Sony.

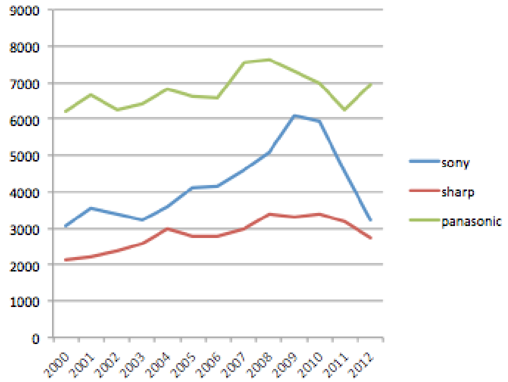


Figure 1 The number of patent publications

2. An approach by the correspondence analysis

We disclose their patent strategies based on typical words associated with smart phone by the correspondence analysis. Results of the analyses of Sony, Sharp, Panasonic and the total of the three companies are shown in figure from 2 to 5. The reference data in the analysis is the numbers of each their patent document in each year, including nouns shown in Table 1. These figures are based on dates of patent publication, and patents are applied to products in the companies.

Focusing on each of their strategies, we reveal their characteristics and differences. In figure2, every word collects vertically around 2004.

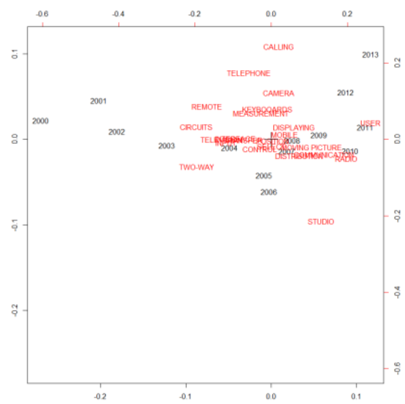


Figure 2 Correspondence analysis result of Sony

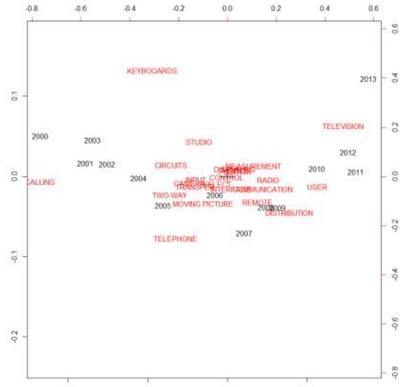


Figure 3 Correspondence analysis result of Sharp

Figure 4 Correspondence analysis result of Panasonic

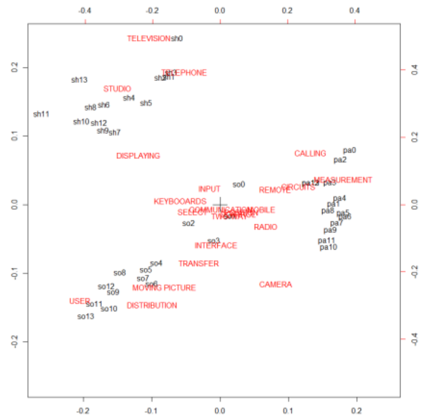


Figure 5 Correspondence analysis result of each company

Moreover, many of the words are close to each other in 2007 and 2008. We deduced that it is because they responded to changes of smart phone industry in 2007 when iPhone was launched. In figure 3 and 4, every word is plotted horizontally. That is, Sharp and Panasonic are not likely to intensively respond to the changes. We deduce that the two companies invested most of their management resources in other businesses such as TV. Compared with Sharp and Panasonic, Sony has close connections with more of the words in figure5.

Therefore, we consider that Sony addressed various technology developments for adapting to changes of the market. In addition, we need to analyze camera from the other perspective because its word is far from the others in each of the companies and the years.

3. An approach by self-organizing maps

As presented above, annual connections between keywords used in contents of the only 20 F-terms and each of the 3 companies, Sony, Sharp and Panasonic were established by correspondence analysis. However, above results are not comprehensive. Therefore, we conduct analyses covering 160 words from top 100 F-terms as one of exploratory approaches. Self-organizing maps（SOM）is used for more multidimensional data analysis[8]. A result of the three companies is shown in figure 6. In addition, colors of nodes depend on the number of patent publications including “CAMERA”. That is, the deeper used pink is, the more the number is, and the deeper used blue is, the fewer the number is.

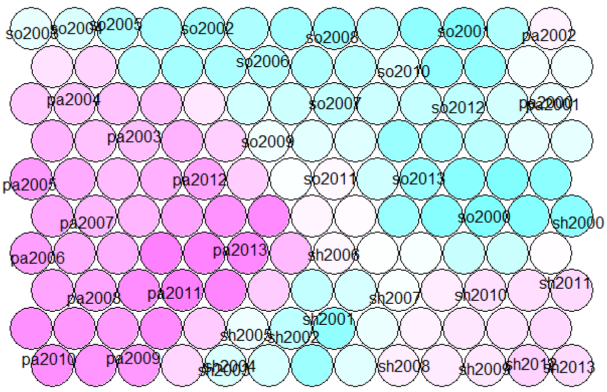


Figure 6 SOM Map of each companies

In figure 6, each company of the years gathers into small area as to Sharp and Panasonic. However, each of the years is plotted widely as to Sony. Thus, Sony has variously changed their technology innovation strategies in order to adapt to environment changes. Moreover, Panasonic and Sharp since 2008 have increased their patent publications associated with “CAMERA”. We consider that Sharp and Panasonic conduct incremental innovations. That is, they have continued own previous strategies.

V. CONCLUSION AND FUTURE WORKS

In this paper, we visualized patent strategies of large ICT companies, Sony, Sharp and Panasonic, in Japanese smart phone market by the correspondence analysis and SOM. From the results, we revealed the differences among the three companies in the market. SONY is likely to be an exploration company implementing various attempts for adapting to change of the market. Sharp and Panasonic are likely to be exploitation companies utilizing their own past strategies rather than adapting to the market changes.

Our three future works are as follows. We analyze foreign ICT companies such as Samsung and Apple in the same way, and compare them with the three companies. Furthermore, the results in this paper need to be compared with results based on existing methods for business analysis and their attempts conducted in the real world.

REFERENCES

[1] Christensen C. M. “The Innovator's Dilemma: The Revolutionary Book That Will Change the Way You Do Business”, HarperCollins (2003)
[2] Oshika T. and Fujimoto T. “Empirical analysis of international trade theory and Product Architecture Theory”, RIETI Discussion Paper Series 06-J-015 (2006) (in Japanese)
[3] Nobeoka K., Ito M. and Morita K. “Case of digital consumer electronics: the failure of value capture by the commoditization”, RIETI Discussion Paper Series 06-J-017 (2006) (in Japanese)
[4] Kim Y.G., Suh J.H. and Park S.C. “Visualization of patent analysis for emerging technology”, Expert Systems with Applications 34 (2008) 1804–1812
[5] Yoon B.G. and Park Y.T. ”A text-mining-based patent network: Analytical tool for high-technology trend”, Journal of High Technology Management Research 15 (2004) 37–50
[6] Tseng Y.H., Linb C.J. and Linc Y.I. “Text mining techniques for patent analysis”, Information Processing and Management 43 (2007) 1216–1247
[7] Patent Tec “The technical development actual condition analysis investigation report about the smart phone according to patent document analysis (patent map) CD-ROM version” (2012) (in Japanese)
[8] T.Kohonen ”Self-Organizing Maps”, Springer (2000)

Dissipative Particle Dynamics Simulation for Vesicle Shape Change

Yoshiyuki Oofuji[1], Naohito Urakami[1], Masayuki Imai[2], Takashi Yamamoto[1]

[1]*Department of Physics and Information Sciences, Yamaguchi University, Yamaguchi 753-8512, Japan*
[2]*Department of Physics, Tohoku University, Sendai 980-8578*
Tel: 81-83-933-5690; Fax: 81-83-933-5690
urakami@yamaguchi-u.ac.jp

Abstract: Spherical vesicles can change to various shapes such as oblate, prolate, stomatocyte-like, and starfish-like, according to the osmotic pressure difference between the inner and outer vesicles. The shape changes of vesicles are very important for understanding the activities of living cells. In this study, we investigated the process of vesicle shape change by carrying out dissipative particle dynamics simulations. We prepared spherical vesicles in which the difference between the numbers of lipids forming the inner and outer leaflets (ΔN) varied. If ΔN was small, with a decrease in the number of water beads inside a vesicle (N_w), a transformation from sphere to oblate, then oblate to stomatocyte-like was observed. If ΔN was large, a transformation from sphere to prolate, then prolate to tube-like occurred. The vesicles shape changes were in good agreement with the experiments. To investigate the mechanism of vesicle transformation in detail, we performed simulations by moving lipids between the inner and outer vesicle leaflets to vary ΔN. As a result, a transformation between the prolate and oblate vesicles was observed. These results indicated that the shapes of vesicles were determined by ΔN and N_w.

Keywords: Vesicle Shape Change, Biomembrane, Dissipative Particle Simulation, Area Difference Model

I. INTRODUCTION

Lipid molecules are the main constituent of biomembranes, and they can self-assemble into various types of structures such as micelles, cylinders, and bilayers (lamellae or vesicles) in an aqueous solution. Vesicles have a closed surface with a bilayer structure, and they have been used as a model system of biomembranes in many experiments designed to investigate the various phenomena observed in living cells, such as the formation of lipid domains [1-4] on the membrane and the diffusion of lipids [5] in the membrane. In addition, the vesicle is a strong candidate for the transport of drug components across the cell membrane, and it plays an important role in cosmetic and food materials.

One of the fascinating properties of lipid membranes is that they form various shapes; they may be spherical, prolate, oblate, stomatocyte-like, starfish-like, etc. [6-8]. Moreover, the vesicles change their shape according to external stimuli. For example, depending on the osmotic pressure difference between the inner and outer vesicles, the water molecules move from the inside to the outside, and a spherical vesicle thus becomes oblate or prolate. With further reduction in the internal volume of the vesicle, an oblate vesicle becomes shaped like a stomatocyte or starfish. In contrast, a prolate vesicle changes shape to a tube or pear [6, 7]. In this way, the vesicles transform to various shapes.

The area difference elasticity (ADE) model [9] has been suggested as a means of understanding the various shapes of vesicles. The ADE model introduces two important parameters: an excess area and an intrinsic area difference. The excess area is the area-to-volume ratio, and the intrinsic area difference is determined by the number of lipids in the outer and inner leaflets. The vesicle shapes are determined by minimizing the sum of the bending energy of the membrane and the energy resulting from the intrinsic area difference, and various vesicle shapes were interpreted on the basis of the ADE model. However, it is difficult to compare the theoretical results with the experimental results because of the difficulty in counting the number of lipids in the two leaflets.

Such difficulty motivated computational studies of vesicles. In the past decade, many computational studies on vesicles have been reported. For instance, vesicle formation from random state [10], formation of lipid domains and budding process [11, 12], and fusion and fission of vesicles [13] have been reported using dissipative particle dynamics (DPD) simulations [14, 15]. In addition, the vesicle shapes consisting of triblock copolymers [16] and the transformation of vesicle shapes in small systems [17] were investigated by DPD simulations. Therefore, DPD simulation is an efficient method of investigating the shape changes of vesicles.

In this work, we carried out DPD simulations to reproduce the various shapes of vesicles observed in the experiment. The transformation of vesicles from a

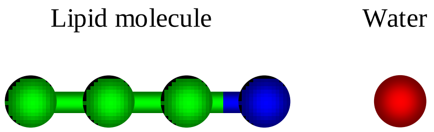


Fig. 1 Illustrations of coarse-grained models used in our simulation. The lipid molecule is represented by one hydrophilic head bead (H) and three hydrophobic tail beads (T). The water molecule is represented by a single hydrophilic bead (W).

spherical shape that accompanied a decrease in the internal volume of the vesicle was investigated. Moreover, the transformation of vesicles induced by the flip-flop of lipids between the outer and inner leaflets was also explored.

II. SIMULATION MODELS AND METHODS

We used coarse-grained models for lipids and water molecules. As shown in Fig. 1, the lipids were represented by one hydrophilic head bead (H) and three hydrophobic tail beads (T), and the water molecules were represented by a single hydrophilic bead (W). All beads have the same mass m. The numbers of lipid and water molecules were 10,000 and 400,000, respectively.

In the DPD simulation, the position and velocity of each bead i are denoted by $(\boldsymbol{r}_i, \boldsymbol{v}_i)$. The time evolution is governed by Newton's equation of motion. The forces acting between the i-th and j-th beads are a conservative force $\boldsymbol{F}_{ij}^{(\mathrm{C})}$, a dissipative force $\boldsymbol{F}_{ij}^{(\mathrm{D})}$, and a pairwise random force $\boldsymbol{F}_{ij}^{(\mathrm{R})}$. These three forces are given as follows:

$$\begin{aligned} \boldsymbol{F}_{ij}^{(\mathrm{C})} &= a_{ij}\,\omega(r_{ij})\hat{\boldsymbol{r}}_{ij}, \\ \boldsymbol{F}_{ij}^{(\mathrm{D})} &= -\gamma\omega^2(r_{ij})(\hat{\boldsymbol{r}}_{ij}\cdot\boldsymbol{v}_{ij})\hat{\boldsymbol{r}}_{ij}, \\ \boldsymbol{F}_{ij}^{(\mathrm{R})} &= \sigma\omega(r_{ij})\theta_{ij}(t)\hat{\boldsymbol{r}}_{ij}, \end{aligned} \tag{1}$$

where a_{ij} is the maximum repulsion force between i-th and j-th beads, $\hat{\boldsymbol{r}}_{ij} = \boldsymbol{r}_{ij}/r_{ij}$, $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$, $r_{ij} = |\boldsymbol{r}_{ij}|$, and $\boldsymbol{v}_{ij} = \boldsymbol{v}_i - \boldsymbol{v}_j$. The parameters γ and σ are related to each other by the fluctuation-dissipative theorem:

$$\sigma^2 = 2\gamma k_{\mathrm{B}} T, \tag{2}$$

where k_{B} and T are the Boltzmann constant and the thermostat temperature, respectively. The weight function $\omega(r)$ is chosen as follows:

$$\omega(r) = \begin{cases} 1 - r/r_0 & r \le r_0 \\ 0 & r > r_0 \end{cases} \tag{3}$$

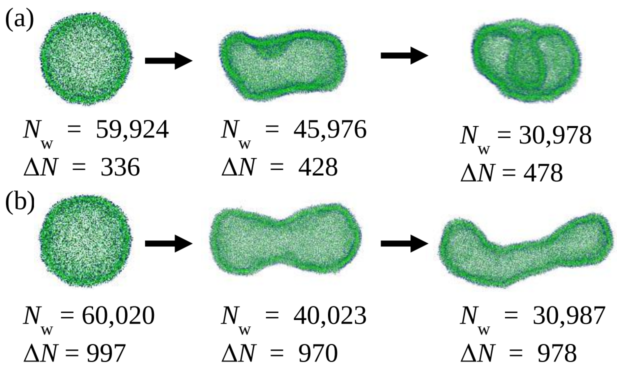


Fig. 2 Shape changes from spherical vesicles, with a decrease in the number of water beads in vesicles. In the initial spherical vesicles, (a) (N_{w}; ΔN) = (59,924; 336) and (b) (60,020; 997). Water beads are not displayed for clarity.

where r_0 is the cutoff radius, and $\theta_{ij}(t)$ is a random variable satisfying the following:

$$\begin{aligned} &\langle \theta_{ij}(t) \rangle = 0 \\ &\langle \theta_{ij}(t)\theta_{kl}(t') \rangle = (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\delta(t - t'). \end{aligned} \tag{4}$$

The spring force for lipids is given by the equation:

$$\boldsymbol{F}_{ij}^{(\mathrm{S})} = -C(1 - r_{i,i+1}/b)\hat{\boldsymbol{r}}_{i,i+1} \tag{5}$$

where $r_{i,i+1}$ denotes the distance between connected beads, b is the equilibrium bond distance, and C is the force constant.

The parameters a_{ij} was set to 25ε for between hydrophilic beads (H and W) and for between hydrophobic beads (T), where ε is the energy unit. For between hydrophilic and hydrophobic beads, a_{ij} was set to 200ε. The noise parameter σ was set to 3.0 and $k_{\mathrm{B}}T = \varepsilon$. In the spring force, $C = 100\varepsilon$ and $b = 0.45r_0$. The number density was $3.0r_0^{-3}$. For the other simulation parameters, the same parameters as in the reference [12] were used. Our simulations were carried out using the COGNAC within the OCTA program [18].

In our simulations, in order to investigate the shape changes of vesicles, we varied two parameters: the number of water beads inside a vesicle, N_{w}, and the difference between the number of lipids, $\Delta N = N_{\mathrm{out}} - N_{\mathrm{in}}$, where N_{out} and N_{in} denoted the number of lipids in the outer and inner leaflet of a vesicle, respectively. The parameters N_{w} and ΔN correspond to the excess area and the intrinsic area difference in the ADE model. N_{w} was reduced by moving the water beads inside a vesicle to the outside, and ΔN was varied by the flip-flop of the lipids between the inner and outer leaflets.

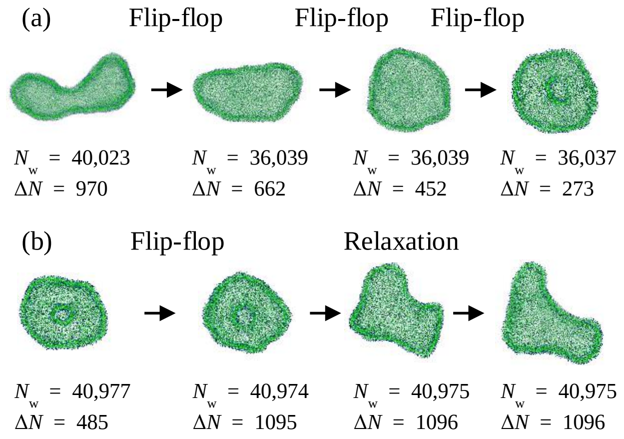


Fig. 3 Transformation of vesicle shapes by changing ΔN. (a) Tube-like vesicle transforms to stomatocyte by moving the lipids from the outer to the inner leaflets in a step-by-step manner. (b) Stomatocyte vesicle transforms to a triangular oblate shape by moving the lipids from the inner to the outer leaflets at the first step.

III. SIMULATION RESULTS

1. Transformation of vesicle shape by changing N_w

We examined the transformation of vesicle shape from spherical to the other shapes. In order to reproduce the volume change of a vesicle in our simulations, 1,000 water beads inside a vesicle were selected randomly, and transferred outside the vesicle in every 5,000 steps. In this way, the number of water beads inside the vesicle (N_w) decreased during simulation runs. As shown in Fig. 2, we prepared two types of spherical vesicles, N_w = 59,924, ΔN = 336 and N_w = 60,020, ΔN = 997. For a small ΔN, with a decrease in N_w, the spherical vesicle became oblate at N_w = 45,976, and stomatocyte-shaped at N_w = 30,978 (Fig. 2 (a)). In contrast, for a large ΔN, with a decrease in N_w, the vesicle shape changed from spherical to prolate at N_w = 40,023, and to a tube at N_w = 30,987 (Fig. 2 (b)). The transformation sequence obtained in the simulation is in good agreement with the experiment [6, 7]. During the simulations, ΔN changed because the spontaneous flip-flop of lipids occurred. Moreover, the variation of N_w occurred because of the spontaneous transfer of water beads between the inside and outside of a vesicle. However, these variations in ΔN and N_w were small and could be neglected. Therefore, the results show that the difference in the initial structures of vesicles is important in determining which branch of transformation the sphere is subject to (oblate or prolate).

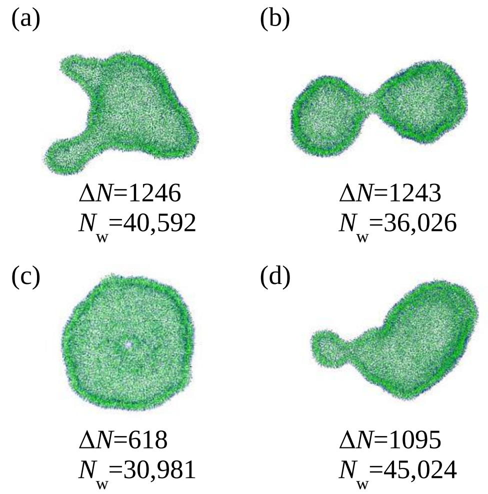


Fig. 4 Various shapes of vesicles obtained by our simulations by changing N_w and ΔN. (a) Starfish-like shape, (b) dumbbell, (c) torus, and (d) racket.

2. Transformation of vesicle shapes by changing ΔN

In order to confirm that the difference of ΔN determines vesicle shape, we investigated the shape changes caused by the flip-flop of lipids between the inner and outer leaflets. During the simulation run, the water beads were not moved from the inside to the outside vesicle but there was a spontaneous transfer of water beads, so N_w was kept almost constant. As shown in Fig. 3(a), the lipids in the tube-like vesicle were moved from the outer to the inner leaflets, that is, ΔN changed from 970 to 273 in a step-by-step manner. The tube-like vesicle became oblate and stomatocyte-shaped. In the stomatocyte vesicle in Fig. 3(b), the lipids moved from the inner to the outer leaflets, and ΔN changed from 485 to 1096 at the first step. In fact, the stomatocyte-shaped vesicle changed to a triangular oblate shape. From the results, we concluded that the shapes of vesicles were determined by ΔN. The transformation sequences obtained in the simulations were also observed in the experiment [8], which was carried out on the condition that the volume of vesicles remained almost unchanged.

3. Various shape changes of vesicles

In the simulations, we reproduced various shape changes of vesicles resulting from changes in N_w and ΔN, as shown in Fig. 4. The starfish-like vesicle was derived from the stomatocyte vesicle by increasing ΔN. The dumbbell-like vesicle was obtained from the prolate vesicle by decreasing N_w and increasing ΔN. The oblate vesicle became torus-shaped following a decrease in N_w. The racket was obtained from a spherical vesicle by increasing ΔN. These vesicle shapes were observed in the experiments [6-8]. It was possible to reproduce the various vesicle shapes using our model for the simulations.

IV. CONCLUSION

We carried out DPD simulations to investigate the transformation of vesicle shapes from spherical to the other shapes with a decrease in N_w. For small ΔN, the vesicle changed from spherical to oblate. In contrast, for large ΔN, the vesicle changed from spherical to prolate. With further reduction in N_w, the oblate vesicle became stomatocyte-like, and the prolate vesicle changed to a tube. The results were in good agreement with the experimental results [6, 7]. To investigate the mechanism of the vesicle shape changes in detail, we carried out the simulations by moving lipids between the inner and outer leaflets of the vesicle. We observed the vesicle shape change from a tube to prolate, oblate, and finally, stomatocyte-like when lipids were moved from the outer to the inner leaflet, corresponding to a decrease in ΔN. In contrast, there was a transformation from stomatocyte-like to triangular oblate when lipids were moved from the inner to the outer leaflet, corresponding to an increase in ΔN. These results indicate that the shapes of vesicles are determined by N_w and ΔN. The vesicle shapes obtained in the simulations were comparable with those predicted by the ADE model. In addition, we also obtained the starfish, dumbbell, torus, and racket shapes by varying N_w and ΔN. In the future, we must analyze these vesicle shapes in detail, and we will clarify the mechanism of various vesicle shape changes. Information about the shape changes of vesicles would help us to understand the motion of living cells when reacting to external stimuli.

ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Number 24540437, 25247070.

REFERENCES

[1] Veatch SL, Keller SL (2002) Organization in lipid membranes containing cholesterol. Phys. Rev. Lett., 89:268101(1-4).
[2] Veatch SL, Keller SL (2005) Miscibility phase diagrams of giant vesicles containing sphingomyelin. Phys. Rev. Lett., 94:148101(1-4).
[3] Masui T, Urakami N, Imai M (2008) Nano-meter–sized domain formation in lipid membranes observed by small angle neutron scattering. Eur. Phys. J. E, 27:379-389.
[4] Cicuta P, Keller SL, Veatch SL (2007) Diffusion of liquid domains in lipid bilayer membranes. J. Phys. Chem. B, 111:3328-3331.
[5] Scherfeld D, Kahya N, Schwille P (2003) Lipid dynamics and domain formation in model membranes composed of ternary mixtures of unsaturated and saturated phosphatidylcholines and cholesterol. Biophys. J., 85:3758-3768.
[6] Hotani H (1984) Transformation pathways of liposomes. J. Mol. Biol. 178(1):113-120.
[7] Yanagisawa M, Imai M, Taniguchi T (2008) Shape deformation of ternary vesicles coupled with phase separation. Phys. Rev. Lett. 100:148102(1-4).
[8] Sakashita A, Urakami N, Ziherl P, Imai M (2012) Three-dimensional analysis of lipid vesicle transformations. Soft Matter, 8:8569-8581.
[9] Seifert U (1997) Configurations of fluid membranes and vesicles. Adv. Phys., 46:13-137.
[10] Yamamoto S, Maruyama Y, Hyodo S (2002) Dissipative particle dynamics study of spontaneous vesicle formation of amphiphilic molecules. J. Chem. Phys., 116:5842-5849.
[11] Yamamoto S, Hyodo S (2003) Budding and fission dynamics of two-component vesicles. J. Chem. Phys., 118:7937-7943.
[12] Laradji M, Kumar PBS (2004) Dynamics of Domain Growth in self-assembled fluid vesicles. Phys. Rev. Lett., 93:198105(1-4).
[13] Granfmüller A, Shillcock J, Lipowsky R (2007) Pathway of Membrane fusion with two tension-dependent energy barriers. Phys. Rev. Lett., 98:218101(1-4).
[14] Groot RD, Warren PB (1997) Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. J. Chem. Phys., 107:4423-4435.
[15] Groot RD, Madden TJ (1998) Dynamic simulation of diblock copolymer microphase separation. J. Chem. Phys., 108:8713-8724.
[16] Li X, Pivkin IV, Liang H, Karniadakis GE (2009) Shape Transformations of membrane vesicle from amphiphilic triblock copolymer. A dissipative particle dynamics simulation study, Macromolecules, 42:3195-3200.
[17] Noguchi I, Urakami N, Imai M, Yamamoto T (2010) Simulation of shape transformations of lipid bilayer vesicles (in Japanese). Kobunshi Ronbunshu, 67:605-610.
[18] Aoyagi T, Sawa F, Shoji T, Fukunaga H, Takimoto J, Doi M, (2002) A general-purpose coarse-grained molecular dynamics program. Comput. Phys. Commun., 145:267-279.

Intelligent mechanisms in *E. coli* in processing carbon sources

Zhongyuan Tian and Hiroshi Matsuno*

Graduate School of Science and Engineering, Yamaguchi University,
1677-1 Yoshida, 753-8512 Yamaguchi-shi, Yamaguchi, Japan
Tel/Fax : 81-83-933-5697
hmatsuno@yamaguchi-u.ac.jp

Abstract: *E. coli* is “wise” enough to take suitable responding time, and suitable responding behaviors, when facing different kinds and intensities of stimulations. According to the time cost of a respond to a signal, we divide the intracellular processes into 2 levels: *central dogma level* and *post translation level*. *The central dogma level* includes the whole process of a gene transcribed to an mRNA then translated to a protein. *The post translation level* includes other intracellular process of non direct relation with genetic DNA or mRNA, for example protein-protein interaction, protein-metabolite interaction, protein localization, protein motion, etc. The time consuming for an action in *the central dogma level* is longer than that in *the post translation level*.

In this study, firstly, we constructed a systematical network of glucose, PTS, glycogen and chemotaxis system, which was mediated by EI, HPr and ATP. This network plays a function as a switch, which controls *E. coli*’s motion within the first few seconds. By using this switch, we unveiled a fact that the shorter time stimuli result in *the post translation level* reactions for quicker response. But further longer time stimuli will activate more time consuming *the central dogma level* reactions, that is this type of signals will be handled in the process of transcription or translation.

Secondly, different intensities of signals result in different kinds of actions, was illustrated by ppGpp example. A low concentration ppGpp will cause RNAP express rRNA & table genes only; meanwhile, a high concentration ppGpp will cause RNAP express mRNA genes only. When *E. coli* is not in stringent occasion, without the binding of ppGpp, RNAP will express mRNA and other rRNA & stable genes in a proportion of 2:8.

In all, a bacterium, like an *E. coli*, not only can identify different kinds of signals; but also can distinguish a signal’s time-length and intensity, and then adopt different respond actions.

Keywords: chemotaxis, PTS, ATP, ppGpp, the central dogma, post translation

I. INTRODUCTION

A human’s body will initiatively or positively take reactions via reflex arc system or deep tendon reflexes system, facing to different kinds and intensities of stimulations. *E. coli*, a bacterium, is it “wise” enough to take suitable reactions towards the variety signals? Here the word “suitable” includes suitable responding time, and suitable responding behaviors. Our answer is “Yes”. In *E. coli*, different kinds, time length or intensities of signals can drive suitable functioning units (e.g. proteins) on suitable levels to make suitable reactions.

According to the time cost of a respond to a signal, we divide the intracellular processes into 2 levels: *central dogma level* and *post translation level* (**Fig 1**). *The central dogma level* includes the whole process of a gene transcribed to an mRNA then translated to a protein. *The post translation level* includes other intracellular process of non direct relation with genetic DNA or mRNA, for example protein-protein interaction, protein-metabolite interaction, protein localization, protein motion, etc. The time consuming for an action in *the central dogma level* is longer than that in *the post translation level*.

For *E. coli*, the average length of coding sequences is 1068 bp [1], and the maximal transcription speed is about 40-80 bp/sec [2, 3], the maximal translation speed is about 20 aa/sec [3, 4], then the average time cost for a gene expression in an optimal condition is about 40 second. But in reality, a gene’s expression may costs

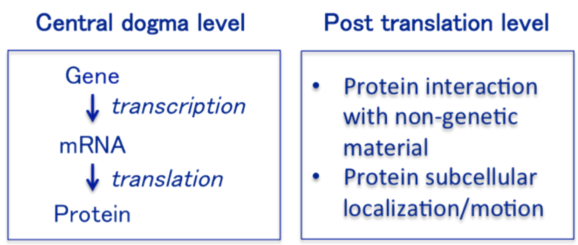


Fig. 1 According to the time cost of intracellular action, we divide the intracellular processes into two levels: *the central dogma level* and *the post translation level*

from several minutes to several days. And the time consuming in post translation level is commonly much shorter then that of central dogma level. For example, the for a chemotaxis reaction in *E. coli*, only need few seconds, or even in a sub-second time scale [5], which includes at least 6 steps actions of *the post translation level* from a signal to a chemoreceptor, …, at last to a flagella.

In this context, examples (a) uses a systematical model of PTS-Chemotaxis-glycogen to describe different levels' responds resulted from a same kind of stimuli coming with different time lengths.

According to intensity of a signal, the respond should be classified into *low*, *mid* and *high* intensity cases, which is illustrated by the example (b): Using a ppGpp-RNAP model to explain different kinds of actions are caused by different intensities of the stimuli.

II. Example (a): A same kind of stimuli coming with different time lengths result in different levels' responds

In studies on *E. coli*, PTS, glycogen, and flagellum are common objects, but always studied separately. PTS is the transport system for an *E. coli* intaking cultural glucose [6]. Glycogen is a polymer functioning as a carbohydrate intracellular storage [7]. Flagellum together with FilM and CheY forms the *E. coli* motion driver, which gives in time response to the five kinds of chemoreceptors [8].

Intra these 3 systems, EI, HPr, $EIIA^{Glc}$ and ATP are key mediators. Un-phosphorylated EI (EI) inhibits the auto-phosphorylation of CheA (P~CheA), which in turn stops the transferring of phosphate group from P~CheA to the flagella motor CheY; this will causes flagella rotating in clock-wise (CW) direction and *E. coli* will tumbling here [6]. The binding of un-phosphorylated HPr (HPr) with GlgP (HPr::GlgP) catalyzes a quicker glycogen decomposition, $EIIA^{Glc}$ regulates expression levels of *glgBXCAP*, *ptsHIcrr* and *ptsG*, indirectly via cAMP/CRP complex [9]. ATP is a necessary member of CheA auto-phosphorylation (**Equ. 1**) and ADPG synthesis (**Equ. 2**) [10]. We can find, despite $EIIA^{Glc}$ works up to *the central dogma level*, all the other 3 mediators are mainly functioning in *the post translation level*.

$$CheA + ATP \leftrightharpoons P{\sim}CheA + ADP \qquad \textbf{(1)}$$

$$G1P + ATP \overset{GlgC}{\Longrightarrow} ADPG + PPi \qquad \textbf{(2)}$$

Shorter time stimulation causes responses on the post translation level. The analysis in this section is based on an assumption: proteins levels are sufficient to perform their functions. In **Fig. 2**, glucose and ATP are 2 stimuli of outside (in the culture) and inside (intracellular) respectively. When an *E. coli* meets glucose at the first time, it quickly activates the poles located PTS (un-phosphorylated) by passing phosphate group to glucose. Meanwhile glucose as a chemotaxis signal has been captured by poles located chemoreceptors. Signals form both these 2 pathways repress the CheA auto-phosphorylation, which at last result in *E. coli* run-forward by rotating flagella in

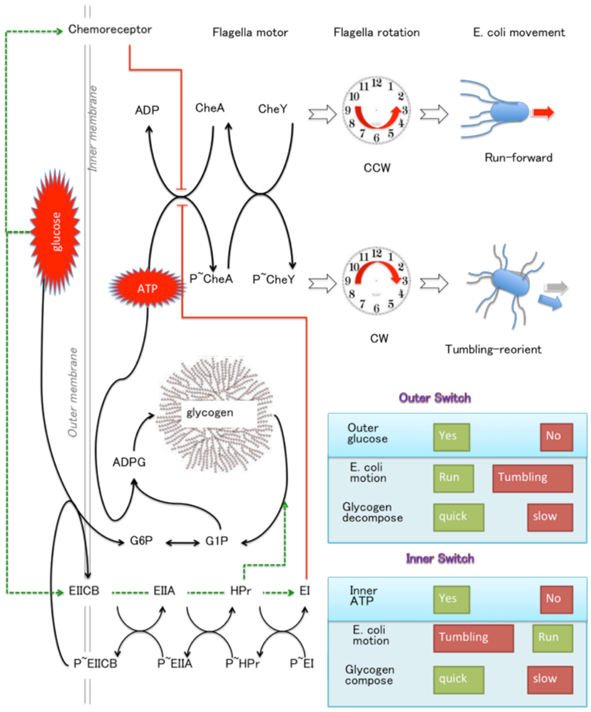


Fig. 2 Shorter time stimulation causes responses on *the post translation level*, which is realized by the network of PTS, glycogen and chemotaxis system.

counter clock-wise (CCW) direction. Simultaneously, poles located HPr::GlgP complex makes pole-located-glycogen a quick decomposition to supply more phosphate group as soon as possible, since glycogenolysis is a quicker process and needless energy driven.

If then outside glucose disappears, chemoreceptors will lose its inhibition on CheA auto-phosphorylation, so as the phosphorylated EI (P~EI). Now, *E. coli* will stop run-forward, instead, turn direction (tumbling), until catching another glucose.

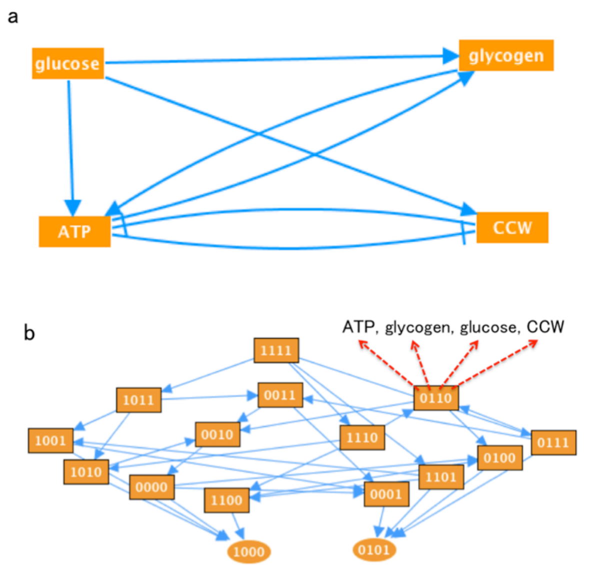


Fig. 3 Boolean network and states transition analysis result of the switch of ATP, glucose, glycogen and chemotaxis system.

If following the first step, the outside glucose is there still; the un-phosphorylated CheY will continue to drive *E. coli* run-forward. But accompany with quick glycogenolysis driven by EI, much more ATP is produced in the pathway of glycolysis and TCA cycle. Accompany with the sudden accumulation of intracellular ATP, the glycogenesis pathway is turned on, which is caused by glycogen content varies inexpertly; and the CheA auto-phosphorylation process will be returned on by the higher pressure of ATP and lower pressure of the inhibitions, which results in *E. coli* tumbling-reorient.

After the aforementioned preparation period, if there is still a glucose signal, then *E. coli* will "realize": there really has a "glucose banquet". At this time, more phosphate grout is transferred from PEP to PTS for glucose uptake, ATP is in a quite flat level. *E. coli* runs into the glucose now.

In all, for shorter time stimulation, when an *E. coli* meets glucose (outside stimulus), it quickly takes suitable types of motion and uptake simultaneously. But if the time length of this decision left for *E. coli* is limited in only a few seconds, a longer time scale would become re-oriented by Brownian motion [5]. When intracellular phosphate group (ATP) lacks (inside stimuli), the corresponding flagella motion and intracellular behavior will change.

In order to clarify the relationship with glucose, ATP, glycogen and flagella, we construct a Boolean network with Ginsim [11] (**Fig. 3a**). And from the result of its analysis (**Fig. 3b**), we can find, the system of double signals (glucose and ATP) and double responses (Flagella and glycogen) reaches 2 stable states, of which: 1000 represents there is only ATP, but no glucose and glycogen, *E. coli* will tumble there to reorient; 0101 means if there is no ATP and glucose, but only glycogen remains, the *E. coli* will go straight. This mechanism can be understood as a switch of E. coli facing shorter time stimulation. In one word, E. coli now either tumbles there with ATP inside, or run forward with glycogen inside. We should not forget, afore-talked mechanism only occurs at the first few seconds.

Longer time stimulation causes responses on the central dogma level. If an *E. coli* is emerged in a glucose culture, then the stimulating time is long enough to activate the reaction of *the central dogma level*. This process includes a series of complex networks, such as PEIIAGlc&cAMP and FDP&Cra subpathways, (As show in **Fig. 4**), which were explained in the paper [9].

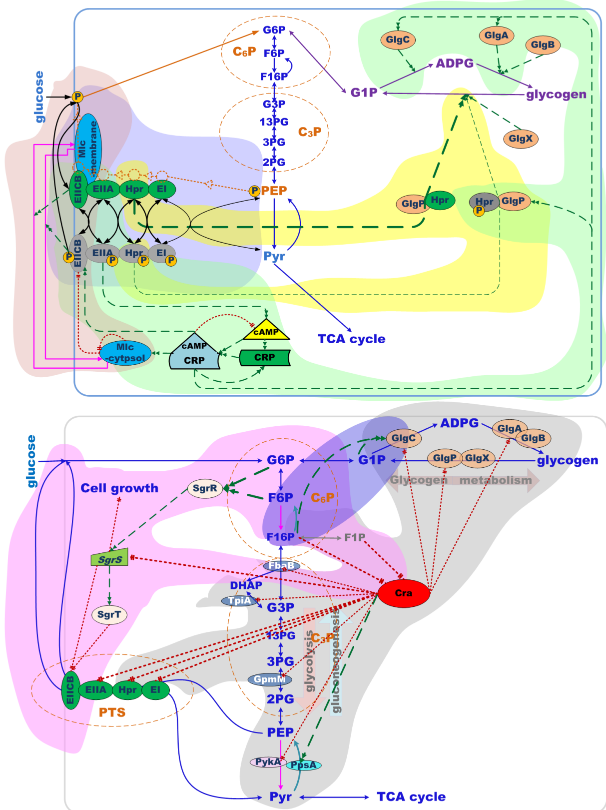


Fig. 4 Regulation mechanisms inter PTS and glycogen metabolism: PEIIAGlc&cAMP pathway, FDP&Cra pathway and HPr subcellular localization

III. Example (b): Using a ppGpp-RNAP model to explain that different kinds of actions are caused by different intensities of the stimuli

ppGpp is a stringent alarmone in *E. coli*, whose mechanism was clearly studied by Traxler et al. in their series papers [12–14]. **Fig. 5** shows that the 3 "bands" of ppGpp intensities result in 3 types of RNAP functions. If there is no starvation signal in an *E. coli* (ppGpp is null), RNAP will employ RpoD to transcribe mRNA and rRNA & stable genes in a proportion of 20% and 80%. If *E. coli* facing only amino acid starvation, a low concentration of ppGpp is produced, as a consequence, the low concentration ppGpp will bind to and drive RNAP to only express rRNA & stable genes. As all nutrition starvation occurs, High concentration ppGpp will be produced, which together with DksA binds to and drive RNAP employing RpoS to transcribe mRNA only.

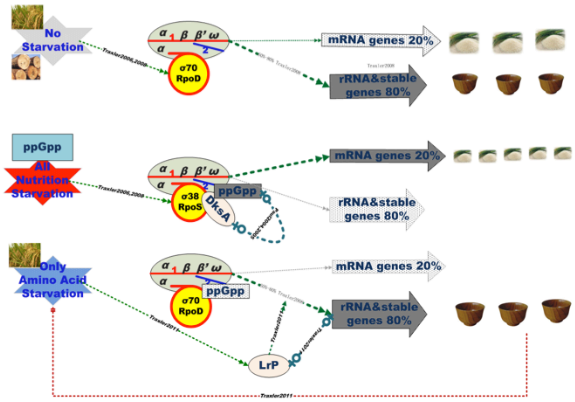


Fig. 5 Different intensities of ppGpp result in different kinds of RNAP actions.

IV. CONCLUSION

E. coli, a bacterium, is an efficient system, when facing different circumstances. In this study, firstly, by systematically constructing a network of glucose, PTS, glycogen and chemotaxis system, we unveiled a fact that the shorter time stimuli results in *the post translation level* reactions. And further longer time stimuli will activate more time consuming *the central dogma level* reaction. Secondly, different intensities of signals result in different kinds of actions, was illustrated by ppGpp example.

The last but not the least, we would like to say, biology can not be classified to higher or lower levels, they only evolve towards different directions.

REFERENCES

[1] Moriyama EN (1998), Gene length and codon usage bias in Drosophila melanogaster, Saccharomyces cerevisiae and Escherichia coli, Nucl. Acids Res 26(13):3188-3193

[2] Dennis PP (1974), Macromolecular Composition During Steady-State Growth of Escherichia coli B/r. J Bacteriol 119(1):270-281

[3] Neidhart FCE (1996), Escherichia coli and Salmonella (No. Ed. 2). Blackwell Science Ltd

[4] Young R (1976) Polypeptide-chain-elongation rate in Escherichia coli B/r as a function of growth rate. Biochem J 160(2):185-194

[5] Sourjik V (2010), Spatial organization in bacterial chemotaxis. EMBO J 29(16):2724-2733

[6] Deutscher J (2006), How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria, Microbiol. Mol Biol Rev 70(4):939-1031

[7] Wilson WA (2010), Regulation of glycogen metabolism in yeast and bacteria. FEMS Microbiology Reviews

[8] Falke JJ (1997), The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. Annu Rev Cell Dev Biol 13:457-512

[9] Tian Z (In press), Identification of Key Regulators in Glycogen Utilization in E. coli Based on the Simulations from a Hybrid Functional Petri Net Model. BMC Systems Bio

[10] Rahimpour M (2013), GlgS, described previously as a glycogen synthesis control protein, negatively regulates motility and biofilm formation in *Escherichia coli*. Biochemical Journal 452(3):559-573

[11] Naldi A (2009), Logical modelling of regulatory networks with GINsim 2.3. Biosystems 97(2):134-139

[12] Traxler MF (2006), Guanosine 3′,5′-bispyrophosphate coordinates global gene expression during glucose-lactose diauxie in Escherichia coli. PNAS 103(7):2374-2379

[13] Traxler MF (2008),The global, ppGpp-mediated stringent response to amino acid starvation in Escherichia coli. Molecular Microbiology 68(5):1128-1148

[14] Traxler MF (2011), Discretely calibrated regulatory loops controlled by ppGpp partition gene induction across the 'feast to famine' gradient in Escherichia coli. Molecular Microbiology 79(4):830-845

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

An autonomous propagation of ciliary metachronal wave on elastic surface of *Paramecium* cells

N. Narematsu and Y. Iwadate

Faculty of Science, Yamaguchi Universitys, Yamaguchi 753-8512
(Tel : 81-83-933-5760; Fax : 81-83-933-5760)
(iwadate@yamaguchi-u.ac.jp)

Abstract: Ciliary movements in protozoa show metachronal coordination so as to maintain a constant phase difference between adjacent cilia. This coordination is called as “metachronal wave”. It is now generally thought that metachronal waves arise from hydrodynamic coupling between adjacent cilia at extracellular fluid. However, under the breakdown the hydrodynamic coupling of ciliary movements at a restricted portion of a *Paramecium* cell, metachronal waves pass over the portion. We will discuss the mediator of propagation of metachronal waves.

Keywords: cilia, ciliary movement, metachronal wave, *Parmaecium*

I. INTRODUCTION

One of the most interesting features of organisms is their generation of spatiotemporally complex regularity, such as repeated structure of somite and stable rhythm of heartbeat, with no external signal. This autonomous generation of regularity is not limited to multi-cellular organisms or tissues but observed even in uni-cellar ones.

Surface of ciliated protozoa, such as *Paramecium* cells (Fig. 1), is covered with a dense array of cilia. They swim in viscous fluid by beating cilia periodically with a succession of power and recovery strokes [1], [2]. Ciliary movements are coordinated so as to maintain a constant phase difference between adjacent cilia. The coordinated beat cycles of a multitude of them seems as waves traveling on cell surface. The formation of traveling, so-called metachronal, waves needs a high degree of synchronization between the beating cilia [1], [3], [4].

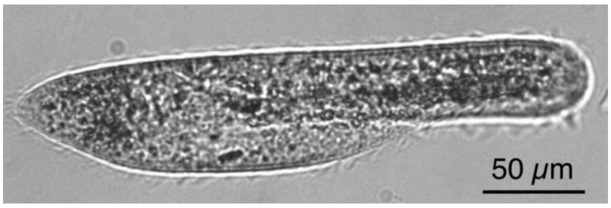


Fig. 1 *Paramecium* cell

It is still not well understood how the individual cilia coordinate with one another to produce such sequential action. From several detailed observations [1], [3]–[7] and theoretical studies [8]–[11], it has been suggested that hydrodynamic coupling may play a role, in that an individual cilium’s deflection can be transmitted by the surrounding viscous fluid to a neighboring cilium which is then induced to deflect similarly. However, is the hydrodynamic interaction the only mediator of metachronal waves?

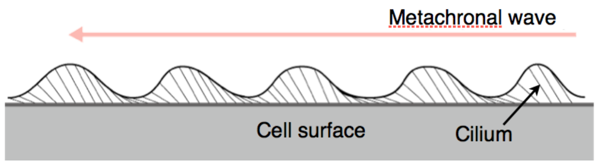


Fig. 2 Metachronal wave of cilia

In this paper, we show experimental and simulation results to suggest that, in addition to hydrodynamic coupling, cortex coupling can also play a role in the propagation of metachronal waves.

To show this, we held one end of the *Paramecium* fixed with a micropipette and then held the other end of the *Paramecium* with a micropipette that is then pulled in an oscillatory manner to induce a periodic stretching of the *Paramecium* body. We then study the metachronal wave propagation in both the unstretched and the stretched cells.

II. MATERIALS AND METHODS

Paramecium culture

Paramecium caudatum cells (strain G3, trichocysts non-discharge mutant) were cultured in an infusion of rice straw at room temperature (20-25 ^oC). Prior to each experiment, *Paramecium* cells were transferred into standard saline medium containing 2 mM KCl, 0.5 mM $CaCl_2$, 0.5 mM $MgSO_4$, 10 mM 3-Morpholinopropanesulfonic acid (MOPS)-Tris (pH 7.2).

Cell capturing

The capturing of a swimming cell was performed according to the methods described previously [12] with slight modifications. To slow their swimming, *Paramecium* was dispersed in a highly viscous medium containing 0.5% (w/v) methylcellulose, 1 mM KC1, 1 mM $CaC1_2$, and 10 mM Pipes-KOH (pH 7.0). A cell of each species was caught using a suction pipette. The diameter of the suction pipette was adjusted to about 35 µm. The pipette was connected to a glass tube and a 5 ml syringe through a silicone tube. The pipette, the glass tube and the syringe were filled with standard saline medium. The syringe was used for adjusting the height of the water surface in the glass tube to control the hydrostatic pressure.

Photolysis of caged calcium

UV application to the restricted area was performed according to the methods described previously [13]–[15], with slight modifications. An inverted microscope (TE300; Nikon, Tokyo, Japan) was used throughout the experiment. A mercury short arc lamp (USH102D; Ushio, Tokyo, Japan) was attached to the side light path of the microscope and used as the UV light source. A pinhole 400 ⌠m in diameter (43-5305, Coherent, Tokyo, Japan) was arranged at the field plane conjugated to the plane of the specimen. The UV light through the pinhole was filtered through a band-pass filter of 270–400 nm (U340; Hoya, Tokyo, Japan) and a cutoff filter of 300nm (UV30; Hoya, Tokyo, Japan) to pass 300–400 nm UV light, as described by Funatsu et al. [16]. The light was collected by a UV-transmitting objective lens (CFI S Fluor 40xH, NA 1.30; Nikon, Tokyo, Japan) to form a small image of the pinhole on the plane of the specimen. Thus, the 300–400 nm UV light was applied to the restricted portion of the *Paramecium* cell, about 2min after microinjection of NP-EGTA. The application time was controlled with an electromagnetic shutter (No. 0; Copal, Tokyo, Japan).

Cyclic stretching of a cell

Anterior and posterior of a *Paramecium* cell were sucked by two pipettes. Both pipettes were connected to piezo actuators (MC-140L, MESS-TEC, Saitama, Japan). In order to induce cyclic stretching of the cell, sequential voltage square pulses were applied to one actuator, which was connected to the pipette sucking the anterior of the cell. Frequency and amplitude of cell stretching were controlled by adjustments of those of voltage pulses. In the case when only shaking without stretching was applied to the cell, sequential voltage square pulses were applied to both actuators.

III. RESULTS

Metachronal waves pass through the portion of the cell surface where the direction of ciliary movements is artificially reversed

The direction of ciliary beat in *Paramecium* cells is reversed by a rise in intraciliary Ca^{2+} concentration [13], [17], [18]. We applied UV light to a small patch of cilia on a *Paramecium* cell that had been previously injected with NP-EGTA medium (white circle in Fig. 2). UV irradiation caused reversal of the beat direction of the cilia (Fig. 3).

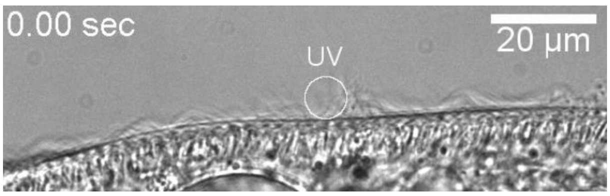


Fig. 3 UV irradiation to a *Paramecium* cell that had been previously injected with NP-EGTA

We investigated whether or not metachronal coordination was eliminated by the breakdown of hydrodynamic coupling of ciliary movements at the reversed cilia. Figure 4 shows a typical kymograph from thin image strips that recorded ciliary movement, in which the beat direction of a local patch of cilia was reversed by photolysis of NP-EGTA. Metachronal waves are shown as oblique black traces from top right to bottom left. The area of reversed cilia is indicated with a rectangle. A dotted line was drawn on the traces of a wave anterior to the reversed cilia. The line was then extended over the portion of reversed cilia to the posterior.

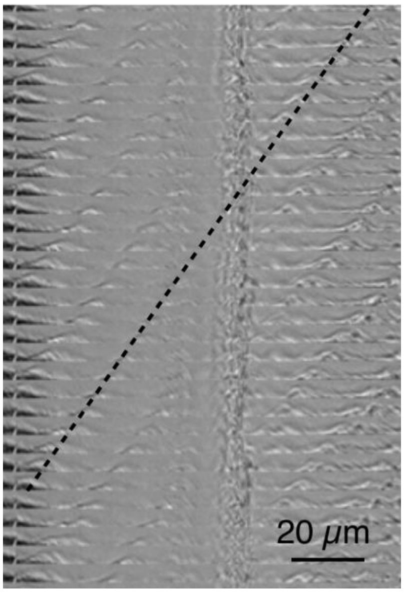


Fig. 4 Kymograph from thin image strips that recorded ciliary movement, in which the beat direction of a local patch of cilia was reversed.

Artificial cyclic stretching of cell body with low frequency decrease that of metachronal wave

Next, we sucked both ends of a live *Paramecium* cell by two micropipettes and applied a vibration with a frequency of 12.5 Hz to the one micropipette, which sucked the anterior of the cell, to induced cyclic stretching of cell body. The applied frequency of 12.5Hz is lower than that of metachronal wave of each cell. The stretching ratio of the distance between the tips of two pipettes after the stretching was adjusted to 5.7%. If the elasticity of cell cortex was a mediator of metachronal wave, frequency of metachronal wave should decrease under the cyclic stretching of a low frequency of 12.5 Hz.

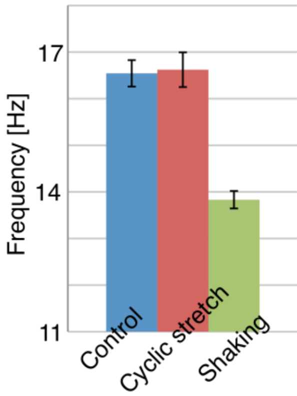


Fig. 5 Wave frequencies under cyclic stretching and shaking.

A typical result is shown (Fig. 5). Fig. 5 is wave frequencies under cyclic stretching and shaking. The frequency decreased significantly under cyclic stretching, although that did not change under shaking. Whereas, wave speeds each condition were almost same (data not shown).

Discussion

It is now generally thought that metachronal wave is mediated by the viscosity of extracellular fluid. No intracellular organelle has been proposed as a candidate for mediating the wave.

In this study, metachronal waves pass over the portion where the direction of ciliary movements was artificially reversed (Fig. 4). Direction of effective and recovery stroke of ciliary beat is not completely parallel to body axis of *Paramecium* cells [1]. Whereas, we observed the wave propagation in the direction parallel to body axis. Thus, UV application from the direction perpendicular to the body axis should not completely prevent the effect of the oblique water flow. There is a possibility in which the oblique water flow may induce the jump of the metachronal waves over the portion where the direction of ciliary movements was artificially reversed, although it strongly suggest that metachronal wave is mediated not only by the viscosity of extracellular fluid but also by something other. *Paramecium* has a contractile cytoskeletal network called the infraciliary lattice (ICL) that lies just beneath the cortical alveoli [19]. ICL may be a possible candidate as an intracellular mediator of metachronal wave.

REFERENCES

[1] Machemer H (1972), Ciliary activity and the origin of metachrony in *Paramecium*: effects of increased viscosity. J Exp Biol 57: 239–259

[2] Sugino K and Naitoh Y (1982), Simulated cross-bridge patterns corresponding to ciliary beating in *Paramecium*. Nature 295: 609–611

[3] Satir P (1963), Studies on cilia. The fixiation of the metachronal wave. J Cell Biol 18: 345–365

[4] Aiello E and Sleigh ME (1972), The metachronal wave of lateral cilia of *Mytilus edulis*. J Cell Biol 54: 493–506

[5] Rompolas P, Patel-King RS and King SM (2010), An outer arm Dynein conformational switch is required for metachronal synchrony of motile cilia in planaria. Mol Biol Cell 21: 3669–3679

[6] Okamoto K and Nakaoka Y (1994), Reconstitution of metachronal waves in ciliated cortical sheets of *Paramecium*-asymmetry of the ciliary movements. J Exp Biol 192: 73–81

[7] Okamoto K and Nakaoka Y (1994), Reconstitution of metachronal waves in ciliated cortical sheets of *Paramecium*-wave stabilities. J Exp Biol 192: 61–72

[8] Gueron S, Levit-Gurevich K, Liron N and Blum JJ, (1997), Cilia internal mechanism and metachronal coordination as the result of hydrodynamical coupling. *Proc Natl Acad Sci USA* 94: 6001–6006

[9] Guirao B and Joanny J-F (2007), Spontaneous creation of macroscopic flow and metachronal waves in an array of cilia. Biophys J 92: 1900–1917

[10] Dillon RH, Fauci LJ and Yang X (2006), Sperm motility and multi ciliary beating: An integrative mechanical Model. *Comput Math Appl* 52: 749–758

[11] Elgeti J and Gompper G (2013), Emergence of metachronal waves in cilia arrays. Proc Natl Acad Sci USA 110: 4470–4475

[12] Iwadate Y, Katoh K, Asai H and Kikuyama M (1997), Simultaneous recording of cytosolic Ca^{2+} levels in *Didinium* and *Paramecium* during a *Didinium* attack on *Paramecium*. Protoplasma 200: 117–127

[13] Iwadate Y (2003), Photolysis of caged calcium in cilia induces ciliary reversal in *Paramecium caudatum*. J Exp Biol 206: 1163–1170

[14] Iwadate Y, Kikuyama M and Asai H (1999), Photolysis of caged Ca^{2+} induces trichocyst discharge in *Paramecium caudatum*, Protoplasma 206: 11–19

[15] Iwadate Y and Nakaoka Y (2008), Calcium regulates independently ciliary beat and cell contraction in *Paramecium* cells. Cell Calcium 44: 169–179

[16] Funatsu T, Kono E and Tsukita S (1993), Time-resolved electron microscopic analysis of the behavior of myosin heads on actin filaments after photolysis of caged ATP. *J. Cell Biol* 121: 1053–1064

[17] Naitoh Y and Kaneko H (1972), Reactivated triton-extracted models of paramecium: modification of ciliary movement by calcium ions. Science 176: 523–524

[18] Naitoh Y and Eckert R (1969), Ionic mechanisms controlling behavioral responses of paramecium to mechanical stimulation. Science 164: 963–965

[19] Garreau de Loubresse N, Keryer G, Viguès B and Beisson J (1988), A contractile cytoskeletal network of *Paramecium*: the infraciliary lattice. J Cell Sci 90: 351–364

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

An analysis of the synergy in a ball throwing task

Hiroshi KIMURA, and Jun NISHII
Graduate School of Science and Engineering, Yamaguchi University.
1677-1 Yoshida, Yamaguchi, 753-8512, Japan
E-mail: {kimura, nishii}@bcl.sci.yamaguchi-u.ac.jp

Abstract: Throwing a ball is a very important task in baseball. To throw a ball into a given target, we must determine the initial state of the ball, the position, speed, and throwing direction at ball release. There are two kinds of strategies in the choice of these variables; the first is to learn a desired value set for these state variables through leaning and try to adjust the variables to the desired values. The second is to allow the variance of each variable in every trial but compensate for the values each other in a cooperative manner so as to hit the target. The purpose of this study is to elucidate the knack of the ball throwing task by observing how good thrower selects the state variables in throwing. We measured ball trajectories thrown by subjects sitting on a chair into a given target and analyzed if cooperative change among the state variables is observed. The result shows that in good throwers the variance of the release point is suppressed but cooperative change between the initial ball speed and the throwing direction was observed.

Keywords: ball throwing, synergy, UCM analysis, motor control, variability.

I Introduction

Improving the ball control in throwing is a very important task in baseball. If we neglect the effect of air friction, the trajectory of a ball is determined by the initial state variables, the release point, speed, and throwing direction. How do good throwers choose and control these state variables? There are two possibilities. The first solution is to throw a ball by adjusting each value of the state variables accurately to a desired value learned through learning (Fig. 1). In fact, "fix a release point" is a typical advice given by baseball trainers to improve the ball control of pitchers. Furthermore, Nagami reported that the variance of the ball throwing direction largely affects the ball control in throwing [4]. This result seems to suggest that the throwing angle should be fixed to a certain value in ball throwing. The second solution is to allow the variance of each initial state variable but adjust them so as to compensate for the values each other in a cooperative manner in order to throw the ball into a target (Fig. 1). For instance, when the release point is higher than a previous successful throwing, we can hit the target by adjusting the speed and throwing direction into smaller values. In fact, good pitchers and fielders show accurate throwing from various postures, which indicates that they have acquired the technique to adjust the initial state variables by mutual compensation among the variables. Bernstein called such kinds of mutual compensation among abundant state variables synergy and emphasized the importance of the synergy as a source of dexterity [1]. Scholz and Schöner proposed the UCM (Uncontrolled Manifold) analysis to evaluate the degree of synergy [6] and many recent papers have reported that synergy is utilized in many movement tasks [2][3][5].

The purpose of this study is to analyze the synergy in ball throwing task and elucidate a knack that good throwers have acquired through learning.

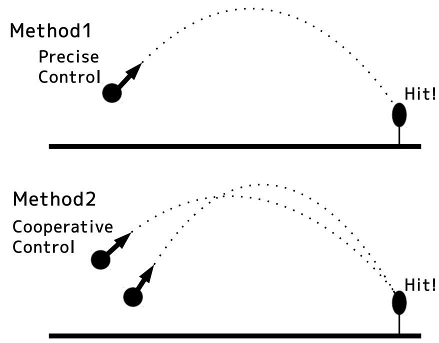


Figure 1: Two kinds of strategies in the choice of the state variables of the ball at ball release. The first is to throw a ball by adjusting state variables accurately to a desired value acquired through learning. The second is to allow the variance of each initial state variables but adjust them so as to compensate for the values each other in a cooperative manner so as to hit the target.

II Methods

We measured the ball trajectory in throwing and we analyzed the synergy between the state variables by the UCM method. The method is described below.

1 Experiment methods

The subjects were three men who had played baseball more than ten years and three men who had shorter or no experience of playing baseball. The profiles of the subjects are shown in table 1. The target was a 0.2×0.2 m square and placed at the height of 0.9 m and the distance of 3.5 m from the throwing position (Fig. 2). Subjects were fixed their chest to the back of the chair by a belt and directed to throw a ball by overhand throwing into the target by their prefer ball speed. They threw a ball 40 times in total. The interval for every throwing was 1 minute, and they were allowed to take a break when they felt fatigue. We measured the trajectories of the ball by a motion capture system (Himawari GV 200, Library co.), by 300 fps, and determined the ball position at ball release. The initial ball speed and the throwing angle were determined by the ball positions in the first 5 frames from the ball release.

Table 1: The subject's profile

Subjects	Age	Baseball career (year)	Height (cm)
A	22	10	180
B	21	11	167
C	20	10	178
D	22	0	173
E	22	0	174
F	22	0	178

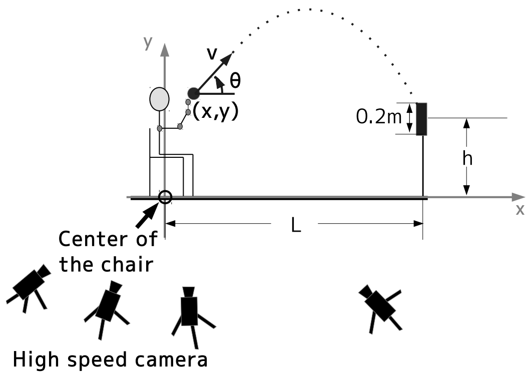


Figure 2: Overview of the measurement experiment. The target was 0.2×0.2 m square and placed at the height of $h = 0.9$ m with the distance of $L = 3.5$ m from the throwing position.

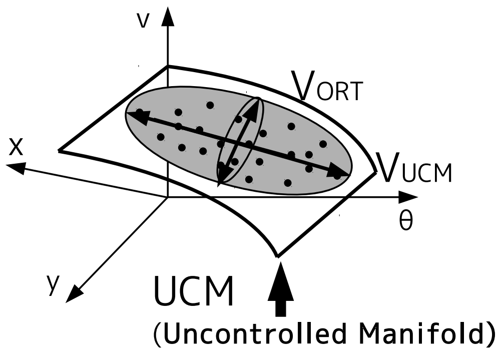


Figure 3: Schematic view of the UCM analysis. In this study, the UCM refers to the manifold that expresses the combination of the state variables, the position (x, y), speed v, and throwing direction θ at ball release, that brings the ball to the center of the target.

2 The UCM method

In this section, the UCM method used in the analysis of the synergy in the ball throwing is described. There are infinite number of choices in the initial state variables of the ball at ball release, the position (x, y), the speed v, and the throwing direction θ, to hit the target (Fig. 2). In this study, the UCM is defined as the manifold that expresses the combination of the solutions in the space of the state variables (Fig. 3).

When the state variables at ball release is $\boldsymbol{D} = (x, y, v, \theta)$, the ball height h at the target position is given by

$$h(\boldsymbol{D}) = (L-x)\tan\theta - \frac{1}{2}g\left(\frac{L-x}{v}\right)^2 (1+\tan^2\theta) + y, \quad (1)$$

where L is the distance between the target and throwing position and g is the amplitude of the gravitational acceleration. A small change in the state variables $\boldsymbol{\varepsilon}$ does not affect the height when $\boldsymbol{\varepsilon}$ satisfies

$$\nabla_{\boldsymbol{D}} h(\overline{\boldsymbol{D}}) \cdot \boldsymbol{\varepsilon} = 0. \quad (2)$$

Let the state variables at the ball release in the i-th throwing as $\boldsymbol{D}^i = (x^i, y^i, v^i, \theta^i)$ and the average as $\overline{\boldsymbol{D}} = (\overline{x}, \overline{y}, \overline{v}, \overline{\theta})$. The parallel and orthogonal components of the deviation $\boldsymbol{\sigma}^i = \boldsymbol{D}^i - \overline{\boldsymbol{D}}$ to the UCM, $\boldsymbol{\sigma}^{i\parallel}$ and $\boldsymbol{\sigma}^{i\perp}$, respectively, are given by

$$\boldsymbol{\sigma}^{i\perp} = (\boldsymbol{\sigma}^i \cdot \hat{\boldsymbol{\varepsilon}}^\perp)\hat{\boldsymbol{\varepsilon}}^\perp \quad (3)$$

$$\boldsymbol{\sigma}^{i\parallel} = \boldsymbol{\sigma}^i - \boldsymbol{\sigma}^{i\perp} \quad (4)$$

where $\boldsymbol{\varepsilon}^\perp = \nabla_{\boldsymbol{D}} h(\overline{\boldsymbol{D}}) / |\nabla_{\boldsymbol{D}} h(\overline{\boldsymbol{D}})|$ is the normal vector of the UCM plane. The averages, $\sigma^\parallel$ and $\sigma^\perp$, are given by

$$\sigma^\parallel = \frac{1}{n}\sum_{i=1}^{n} \left|\boldsymbol{\sigma}^{i\parallel}\right| \quad (5)$$

$$\sigma^\perp = \frac{1}{n}\sum_{i=1}^{n} \left|\boldsymbol{\sigma}^{i\perp}\right|. \quad (6)$$

When $\sigma^\parallel$ is larger than $\sigma^\perp$, such distribution of the state variables suggests the existence of the synergy that suppresses the variance of the height h. To judge the existence of synergy among the state variables synergy, we used the degree of synergy S_h defined by

$$S_h = \frac{\sigma^\parallel - \sigma^\perp}{\sigma^\parallel + \sigma^\perp} \quad (7)$$

By the definition, the degree of synergy takes the value between -1 and 1. $S_h > 0$ indicates the existence of synergy and larger value of S_h means larger degree of synergy.

3 Correlation coefficient

We also calculated the correlation coefficient to examine whether positive correlation between state variables and ball throwing performance exists. The correlation coefficient C is given by

$$C = \frac{\sum_{i=1}^{n}(p_i - \overline{p})(s_i - \overline{s})}{\sqrt{\sum_{i=1}^{n}(p_i - \overline{p})^2}\sqrt{\sum_{i=1}^{n}(s_i - \overline{s})^2}}, \quad (8)$$

where p_i is hit probability, s_i is a state variable in the i-th throwing, and $\overline{p}$ and $\overline{s}$ are the averages.

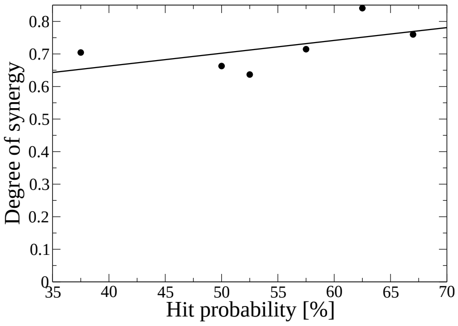


Figure 4: The relation between synergy and task performance. The horizontal axis shows the hit probability in 40 ball throwings by each subject. The vertical axis shows the degree of the synergy among state variables at ball release. The solid line is the regression line of the data, and the correlation coefficient is 0.56.

III Results and Discussion

In this section, we show the results of the UCM analysis and discuss the relation between the synergy among the state variables and the ball throwing performance.

Fig. 4 shows the relation between synergy among state variables and the hit probability. The degrees of the synergy of all subjects were over 0.6, which indicates that the variance of the state variables are compensated each other to adjust the ball height at the target position. The correlation between the hit probability and the degree of the synergy was 0.56, which suggests that good throwers tend to show large synergy.

Fig. 5 shows the relationship between the variance of each state variable at ball release and hit probability. As the hit probability increases, the variances of the horizontal and vertical ball release position decreases (Fig. 5(a)(b)), and the correlation coefficients were -0.89 and -0.60, respectively. This result indicates that the variance of the ball release position tends to be smaller in better throwers.

The correlation coefficient between the variance of the state variables, the initial ball speed and the throwing direction, and hit probability were 0.14 and -0.20, respectively (Fig. 5(c)(d)). Therefore, the correlation between them were weak.

Our results described above have shown that in better throwers the degree of the synergy among the state variables of the ball at ball release was higher, on the other hand, the variance of the release position was smaller than beginners. The variance of the initial speed and the throwing direction have only a weak relation with the hit probability. These results seem to suggest that the release point seldom participates in the synergy among state variables, and the synergy is exploited between the initial speed and the throwing direction (Fig. 6).

To ascertain this hypothesis we computed the degree of the synergy between two state variables, the initial ball speed and the throwing direction, that suppresses the variance of the ball height at the target position. Fig. 7 shows the result and suggests that its defference with the synergy shown in Fig. 4 was small and 5 % at the maximum (in the subject whose hit probability was 57.5 %). This result

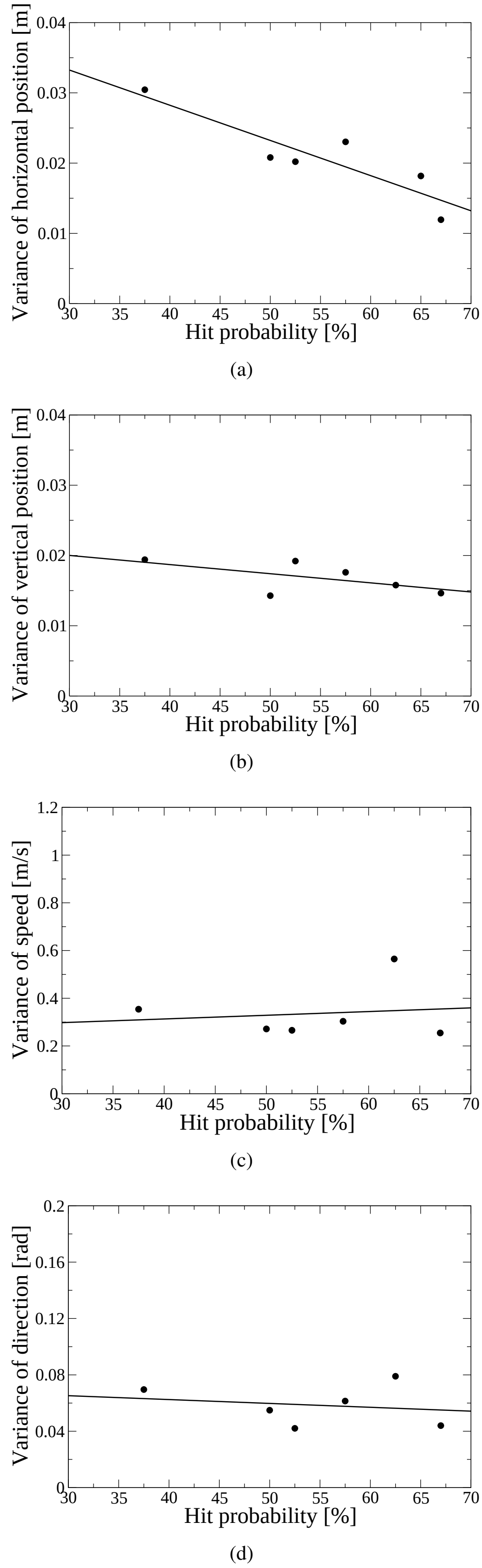


Figure 5: The relation between the variances of the state variables and the task performance. The horizontal axis is the hit probability. The vertical axis in (a), (b), (c), and (d) are the variance of the horizontal and vertical ball release position, initial ball speed, throwing direction, respectively. The solid line is the regression line for each data, and the correlation coefficient in (a), (b), (c), and (d) were -0.89, -0.60, 0.14, and -0.2, respectively.

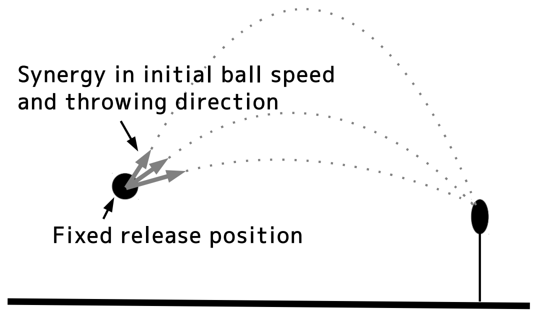


Figure 6: Schematic view of the control of the initial state of the ball at ball release by good throwers. The variance of the release point is suppressed, and synergetic control is found between the initial ball speed and the throwing direction.

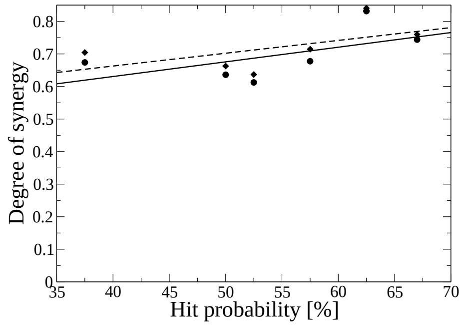


Figure 7: The degree of the synergy between the initial speed and the throwing direction. The circles are the data points and the solid line is the regressive line. The diamonds and the dotted line are the same as the circles and the line in Fig. 4.

shows that the contribution of the ball release position into the synergy to adjust the ball height at target position is small and the synergy is realized by adjusting two parameters, the initial ball speed and the direction.

IV Conclusion

In this study, we analyzed the synergy in ball throwing task to elucidate how good throwers determine the initial state variables of a ball at ball release. The results have shown that the synergy works in the selection of the initial speed and the direction of the ball in all subjects, and the degree of the synergy tends to increase in good throwers. On the other hand, the contribution of the ball release position to the synergy is low and the variance of the position is smaller in good throwers. Bernstein proposed the hypothesis that the degrees of freedom (DOF) of our body is frozen in the early phase of motor learning and the DOF would be freed as learning proceeds to obtain motor dexterity, but our results show that freezing of the DOF seems to be also utilized in a skilled task.

This result shows that the instruction mentioned in the section 1, ”fix a release point”, is an appropriate advice to improve the ball control. However, our result also shows that improvement of the coordination between the initial ball speed and throwing direction would be also important technique in ball throwing, therefore, such coordination should be taken into the practice. For example, throwing a ball by changing the speed and angle might be effective practice on this purpose.

In this study, we analyzed the state variables of a ball at ball release, however, it is also possible to analyze the synergy that coordinates arm joint trajectories, which might give us a cue to know how the variance of the release position is suppressed by good throwers. Furthermore, the subjects in this study threw balls by sitting on a chair, however, ball throwing in baseball is performed by using the whole body. In this case, the degrees of freedom of the body involved in the throwing task increases considerably, hence, the analyses for such movements would be a challenging task but attractive to know the knack for skilled throwing: which variables should be precisely controlled and how synergy of multiple degrees of freedom should be utilized.

References

[1] N. A. Bernstein, *The coordination and regulation of movement*. Pergamon Press, London, 1967.

[2] J. Feng, J. P. Scholz, and M. L. Latash, The role of kinematic redundancy in adaptation of reaching. *Experimental Brain Research,* vol. 176, pp. 54-69, 2007.

[3] M. L. Latash, J. P. Scholz, and G. Schöner, *Motor Control Stratigies Revealed in the Structure of Motor Variability*. Motor Control and Motor Variability, vol. 31, no. 1, pp. 26-31, 2002.

[4] T. Nagami, J. Morohoshi, K. Kanosue, *Relationship between ball control and movements of hand and fingers in baseball pitchers*. スポーツ・アンド・ヒューマン・ダイナミクス講演論文集, pp. 51-55, 2009

[5] J. Nishii, Y. Hashizume, S. Kaichida, H. Suenaga, Yoshiko Tanaka, *Constraint and exploitation of redundant degrees of freedom during walking*. Robotics and Autonomous Systems, Vol. 60, pp. 679-684, 2012

[6] J. P. Scholz, and G, Schöner, The uncontrolled manifold concept: Identifying control veriables for a functional task. *Experimental Brain Research,* vol. 176, pp. 54-69, 2007.

A low dimensional feedback control model that exploits abundant degrees of freedom

J. Nishii and T. Hamamura

Graduate School of Science and Engineering, Yamaguchi University
1677-1 Yoshida, 753-8512 Yamaguchi-shi, Yamaguchi, Japan
Tel: +81-83-933-5691
nishii@yamaguchi-u.ac.jp

Abstract: *When animals perform skilled motor task, they often adaptively choose a solution depending on the circumstances by utilizing abundant degrees of freedom (DOFs) of their body. For instance, a spinal frog can wipe an irritating stimuli off by its foot even if a leg joint are fixed.In this example, multiple leg joints cooperatively work so as to control the foot position by utilizing available joints. Such compensatory control among multiple joints is called joint synergy. We developed a simple neural network model that realizes synergetic control that exploits abundant degrees of freedom. The model takes an architecture of a sand-glass type neural network, which extracts essential variables in the middle layer to perform a task from sensory signals. The neural activity of the middle layer is modified by a feedback control so as to bring the activity to the desired value. The activity of the middle layer affects the activity of the output layer that send control signals to actuators. The model architecture is very simple, however, the results of the computer simulation of a 3-link arm control reproduced the synergetic control observed in the above example of the spinal frog.*

Keywords: motor control, sand-glass type neural network, abundant degrees of freedom, Bernstein problem

1 Introduction

When humans and animals perform skilled motor task, they adaptively choose a solution depending on the circumstances by utilizing abundant degrees of freedom (DOFs) of their body. For instance, Latash et al. reported that a spinal frog can wipe an irritating stimuli off by its foot even if a leg joint are fixed [1]. In this example, multiple leg joints cooperatively work so as to control the foot position by a compensatory manner by utilizing available joints. Such compensatory control among multiple joints is called joint synergy. Bernstein, a Russian physiologist who contributed the development of biomechanics in the first half of the 20th century, insisted that the ability to solve a given task by multiple ways, in other words, by utilizing the synergy among the abundant DOFs, is a typical characteristics of the intelligence that the living bodies show [2]. The purpose of this paper is to propose a neural network model that realizes such synergetic control.

2 A low dimensional feedback control model

In this section, we propose a low dimensional feedback control model that extracts low dimensional performance variables from sensory signals, and exploits an abundant degrees of freedom (Fig. 1).

The proposed model is a simple neural network model that takes an architecture of a sand-glass type neural network (SNN). The SNN is a neural network model with five layers and capable of nonlinear principal component analysis when fewer number of cells are prepared in the middle layer than those in the input layer and the SNN learned the identity transformation [3, 4].

In our model the SNN learns the relation between the sensory signals from the input layer and the control signals from the output layer. The number of the cells in the middle layer is fewer than that of the input layer, hence, it is expected that the middle layer compresses the input signals into low-dimensional variables. After learning each neural cell in the middle layer consists the basis to control the actuators. Furthermore, we add a learning for some cells in the middle layer so as that the firing patterns express essential variables to perform a task, so called performance variables [5], from sensory signals, e.g., the foot position in the above example of the frog from muscle lengths. When the neural activity of the middle layer is modified by a feedback control so as to bring the performance variables to the desired values, e.g., the target foot position, the activity of the middle layer brings the control signals from the out-

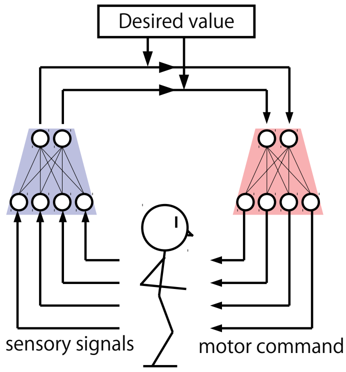


Figure 1: A framework of the movement control by our nervous system proposed by Ting (2007). Ting proposed that essential framework in movement control is the extraction of some basic variables to achieve a task and the expansion of the basic variables into muscle synergies. Our model gives a concrete neural network model that realizes the Ting's idea. (modified from [6]).

put layer to desired one which accomplish the target task. In other words, the low-dimensional feedback control for the middle layer will realizes the motor control that utilizes motor synergy.

3 Simulation experiment

In order to examine the validity of our model, we applied our control model to an arm reaching task.

3.1 Arm model

The controlled object is a 3-link arm with three pairs of muscles for each joint in a horizontal plane. The length and mass of each link is 0.3 m and 1.5 kg, respectively (fig. 2).

The muscle model used in our experiment is based on Kumamoto et al. [7] and Kambara et. al. [8], and the muscle tension T_i exerted by the motor signal u_i for the i-th muscle M_i $(i = 1, 2, \cdots, 6)$ takes the form:

$$T_i = c_i u_i - (k_{0i} + k_{1i} u_i) l_i - (b_{0i} + b_{1i} u_i) \dot{l_i}, \quad (1)$$

where k_{0i} and k_{1i} are damping coefficients, b_{0i}, b_{1i} are viscous coefficients, and c_i is a constant. l_i is muscle length that depends on the joint angle θ_j ($j = i - 1 \bmod 2$):

$$l_i = l_{0i} - a_i \theta_j, \quad (2)$$

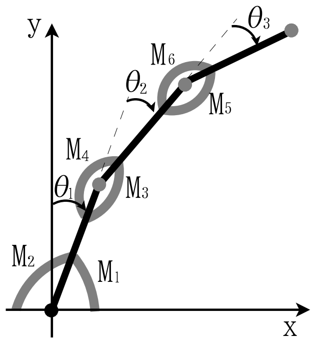


Figure 2: 3-link arm model with 6 muscles.

Table 1: Muscle parameters

	M_1	M_2	M_3	M_4	M_5	M_6
c [N]	42	42	28	28	14	14
k_0 [N/m]	240	240	160	160	80	80
k_1 [N/m]	105	105	70	70	35	35
b_0 [Ns/m]	350	350	350	350	350	350
b_1 [Ns/m]	350	350	350	350	350	350
l_0 [m]	0.12	0.12	0.12	0.12	0.12	0.12
a [m]	0.034	-0.034	0.034	-0.034	0.034	-0.034

where a_i is the moment arm of i-th joint and l_{0i} is the neutral length of the muscle M_i when $\theta_j = 0$. In other words, the first term of the right hand of eq. (1) represents the joint torque exerted by the i-th muscle, the second term the elastic force that is proportional to the muscle length, the third term the viscous force. The torque of j-th joint is given by the difference of the muscle tension around the joint:

$$\tau_j = a_{2j-1} T_{2j-1} + a_{2j} T_{2j} \quad (3)$$

The muscle parameters are shown in the Table 1. In the computer simulation, Webots (Cyberbotics co.) were used for the dynamical computation of the arm movement.

3.2 Controller

The neural network used as a controller has five layers, and the number of cells of each layer is 6,8,3,8, and 6 from the input layer to output layer, respectively. The network receives each muscle length l_i $(i = 1, 2, \cdots, 6)$ and sends motor signals u_i to each muscle. Back-propagation method was used to train the network and the teacher data were given as follows.

(1) The learning was instructed so as to send the control signal to keep the postures represented by input signals. We recorded various

pairs of motor commands $\boldsymbol{u}$ and the muscle length $\boldsymbol{l}$ when the arm came to rest by the given command, and used the data as the teacher data set to train the neural network. The motor commands to the flexor and extensor muscle were chosen so as to satisfy the relation $u_{2j-1} + u_{2j} = 1.0$ $(j = 1, 2, 3)$.

(2) One neuron in the third (middle) layer was trained to output the y-axis value of the hand position.

The above learning (1) and (2) were done alternatively. The input-output relation of each neuron was given by a sigmoidal function $f(x) = \dfrac{1}{1+\exp(-x)}$ and the output range $[0.1, 0.9]$ of the third and the fifth layer was scaled into the range $[-0.3, 0.9]$ [m] that expresses the y-axis value of the hand position and $[0, 1]$ of the motor command u_i, respectively.

3.3 Simulation

The task is to bring the y-axis value of the hand position to the desired value y^d. The motor command was computed by the controller in every 0.04 second based on the input signal that expresses the muscle lengths.

The activity of the neural cell in the middle layer that expresses y-axis value of the hand position was modulated by the following feedback:

$$y \leftarrow y + \epsilon(y^d - y), \tag{4}$$

where ϵ is the feedback gain and $\epsilon = 0.5$ in this simulation. The reaching task is judged as the end when $\dot{\theta}_i$ $(i = 1, 2, 3)$ becomes less than $\dfrac{\pi}{36}$ [rad/s] for one second.

3.4 Experimental results

Fig. 3(a) shows the hand trajectory and change of arm posture during reaching from the initial posture $\boldsymbol{\theta} = (\frac{\pi}{4}, -\frac{\pi}{4}, -\frac{\pi}{4})$ [rad] to the desired one $y^d = 0.4$ [m]. Fig. 3(b) is the arm trajectory with the same condition in Fig. 3(a) except the point that perturbation is applied during reaching. The perturbation was the anti-clockwise joint torque $\boldsymbol{\tau} = (-6.0, -4.0, -2.0)$ [N·m] for one second during arm movement. The hand overshot the target position, but reached the target afterward by the different posture from the final posture in Fig. 3(a), i.e., the proposed control model dynamically generates reaching trajectory.

Fig. 3(c) is the results when the shoulder joint is fixed at the initial angle from the same initial posture as in Fig. 3(a). The trajectory is much different from Fig. 3(a), however, the hand approached the target by using remaining DOF, knee and wrist joints. This result shows that the proposed model realizes the compensatory control by utilizing abundant degrees of freedom even if some DOFs are restricted like the example of the spinal frog mentioned in the introduction.

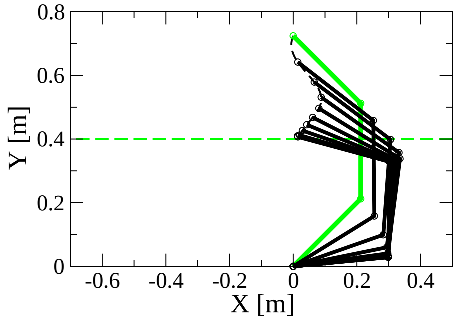
4 Conclusion

In this study we proposed a control model that utilizes abundant degrees of freedom. The model architecture is a simple five-layer neural network with a feedback control for the middle layer. The results of the computer simulation of a 3-link arm, however, have shown that the model can perform a synergetic control that exploits abundant DOFs of the body and reproduced the example of the spinal frog, i.e., the model brought the hand to a given target by using three joints if all joints are available, but even if one joint is fixed, the model computed the solution to perform the given task by utilizing the other joints.

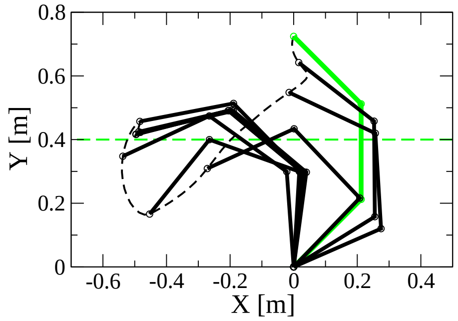
The important feature of this model is that the trajectory of each DOF is not controlled as in traditional control systems in the field of engineering, instead only the essential variables, such as the hand position, are controlled, and the control signals for each DOF is autonomously generated in a compensatory manner so as to realize the desired value of the essential variables. The combination of the desired value of essential variables forms a manifold in the space of the variables and the manifold is called uncontrolled manifold (UCM) [9], and many recent studies have reported that our nervous system stabilizes the UCM in many movement tasks [5, 9, 10]. Our model would be a most primitive and simplest model that realizes the synergetic control that stabilizes the UCM.

Acknowledgement

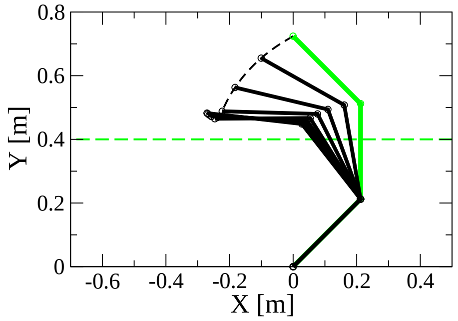
This work was partially supported by KAKENHI (25282183).



(a)



(b)



(c)

Figure 3: Change of arm posture during reaching. The initial posture (gray line) is given by $\boldsymbol{\theta} = (\frac{\pi}{4}, -\frac{\pi}{4}, -\frac{\pi}{4})$ [rad], and the target hand position is $y^d = 0.4$. (a) is normal reaching (b) is the case with perturbation, and (b) is the result when shoulder angle is fixed. The start and end of the perturbation is indicated by arrows.

References

[1] I. M. Gelfand and M. L. Latash. On the problem of adequate language in motor control. *Motor Control*, 2(4):306–313, 1998.

[2] N. A. Bernstein. *The co-ordination and regulation of movements.* Pergamon Press, 1967.

[3] B. Irie and M. Kawato. Acquisition of internal representation by multilayered perceptrons. *Electronics and Communications in Japan, Part III: Fundamental Electronic Science*, 74(11):112–118, 1991.

[4] M. Scholz. Validation of nonlinear pca. *Neural Processing Letters*, 36(1):21–30, 8 2012.

[5] M. L. Latash, J. P. Scholz, and G. Schöner. Toward a new theory of motor synergies. *Motor Control*, 11:276–308, 2007.

[6] L. Ting. Dimensional reduction in sensorimotor systems:a framework for understanding muscle coordination of posture. *Progress in Brain Res*, 165:299–321, 2007.

[7] M. Kumamoto, T. Oshima, T. Yamamoto. Control properties induced by the existence of antagonistic pairs of bi-articular musclesmechanical engineering model analyses. *Human Movement Science*, 13(5):611–634, 1994.

[8] H. Kambara, K. Kim, D. Shin, M. Sato, and Y. Koike. Learning and generation of goal-directed arm reaching from scratch. *Neural Network*, 22(4):348–61, 2009.

[9] J. P. Scholz and G. Schöner. The uncontrolled manifold concept: Identifying control variables for a functional task. *Experimental Brain Research*, 126:289–306, 1999.

[10] J. Nishii, Y. Hashizume, S. Kaichida, H. Suenaga, and Y. Tanaka. Constraint and exploitation of redundant degrees of freedom during walking. *Robotis and Autonoumos Systems*, 60(5):679–684, 2012.

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Extension of Genetic Toggle Switch Based on the Effective Search of State Transitions

M. Sugii[1], A. Fauré[2] and H. Matsuno[2]

[1]*Media and Information Technology Center, Yamaguchi University, 2-16-1 Tokiwadai, Ube-shi, Yamaguchi 755-8611, Japan*

[2]*Graduate School of Science and Engineering, Yamaguchi University, 1677-1 Yoshida, Yamaguchi-shi, Yamaguchi 753-8512, Japan*

Tel : 81-836-85-9916; Fax : 81-836-85-9910

manabu@yamaguchi-u.ac.jp

Abstract: We propose a new procedure to effectively design a mathematical model in two steps. The first step is the creation of possible network architectures under a logical formalism. The second step is the creation of dynamic models, both using a system of differential equations for the analysis of dynamics in the model and hybrid functional Petri net as a common platform for knowledge sharing between biologists and computer scientists.

We extended a genetic toggle switch from 2-state to 3-state by our proposed procedure. Differential equations for 3-state genetic toggle switch were produced by extending the ones for 2-state genetic toggle switch. Nullclines were drawn as a three-dimensional diagram, which shows three stable states of the genetic toggle switch as three cross points of three nullclines.

Keywords: Synthetic biology, Artificial genetic circuit, Logical modeling, Dynamic modeling.

I. INTRODUCTION

In the context of synthetic biology, artificial genetic circuits are designed in the following way: after setting a biological target phenomenon to be investigated, reaction parameter estimations among related molecules are conducted based on the dynamic analyses with mathematical models. Finally a system of biological reactions is developed with these molecules in vivo or in vitro (Fig.1). Gradner et al. produced an artificial genetic toggle switch [1]. They set a simple biological phenomenon and estimated the parameters from biological information, constructing a system of differential equations enabling bistable genetic circuit. Many possible models can usually be considered for the target phenomenon, which motivates us todevelop an effective method to select the best candidates. We propose a new procedure to effectively design a mathematical model in two steps and extended a genetic toggle switch from 2-state to 3-state by our proposed procedure.

II. METHODS

1. 2-state genetic toggle switch

The key of our proposed method is to separate the design process of a mathematical model into two steps. Our proposed method enables us to realize a suitable

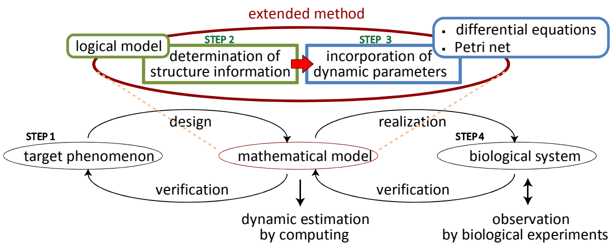


Fig.1. Framework of synthetic biology

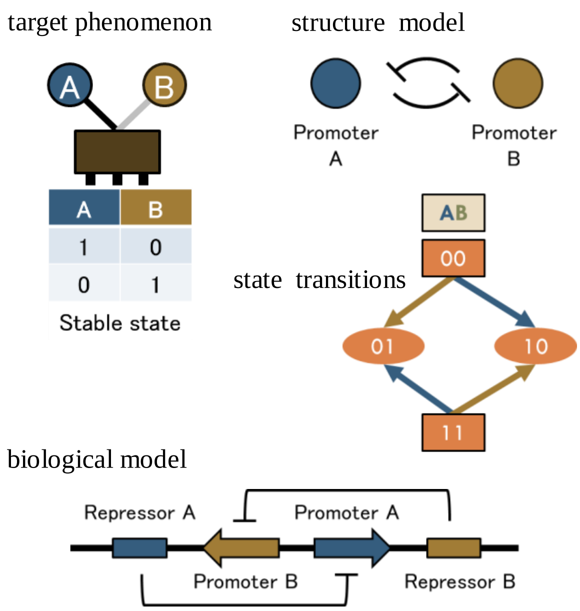


Fig.2a. 2-state genetic toggle switch [1]

artificial genetic circuit in more reduced time and cost than existing methods.

Fig.2 shows the 2-state genetic toggle switch [1]. This biological model consists of two genes A and B being led by prompter regions for each. Repressors A and B are produced from these two genes A and B, which are initiated by two promoter regions B and A, respectively. Repressors A and B inhibit the expressions of promoters A and B, respectively (biological model). This 2-state gene toggle switch is finally transformed into either of two stable states, where either of two genes (target phenomenon) is expressed depending on an initial amount of repressors A and B. A kinetic model and nullclines are given as a system of differential equations as shown in Fig.2b.

2. Extension to 3-state toggle switch

3-state toggle switch has three stable states as shown in Fig.3a (target phenomenon). Among many possible logical structure models, three models are presented in Fig.3a (Creating structure model, Model 1, 2, 3). State transitions, which are produced from GINSim [2], are given as well (State transitions, Model 1, 2, 3). Model 3 can be eliminated from the candidates because of the complicated state transitions, in which more states are used than Models 1 and 2 and inappropriate cycle behaviors are found. Possible structures of artificial genetic circuit to be selected in this way will be sent to the next step, where dynamic parameters are incorporated in the selected structures. Resulting dynamical models shall be a system of differential equations and a hybrid functional Petri net model, which can be executed on the tools such as MATLAB and Cell Illustrator [3], respectively.

3. Simulation of dynamic models

Two different types of simulations are conducted (Fig.3b, Simulation of dynamic models). One is for the analysis of dynamic behavior of the obtained artificial genetic circuit model on a system of differential equations, and the other is for acquiring biologically important behavior for the designing genetic circuit in vivo and/or in vitro. Experts in microbiology need to be participated in the process of parameter determination of these models in order to choose appropriate values from biological aspects.

4. Realization by biological experiments

E.coli will be used for the realization of the artificial genetic circuits. Selection of condition for expression

differential equations

$$\frac{du}{dt} = \frac{a}{1+v^n} - u \qquad \frac{dv}{dt} = \frac{b}{1+u^m} - v$$

nullclines

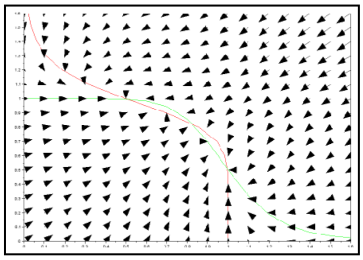


Fig.2b. Differential equations of 2-state genetic toggle switch [1]

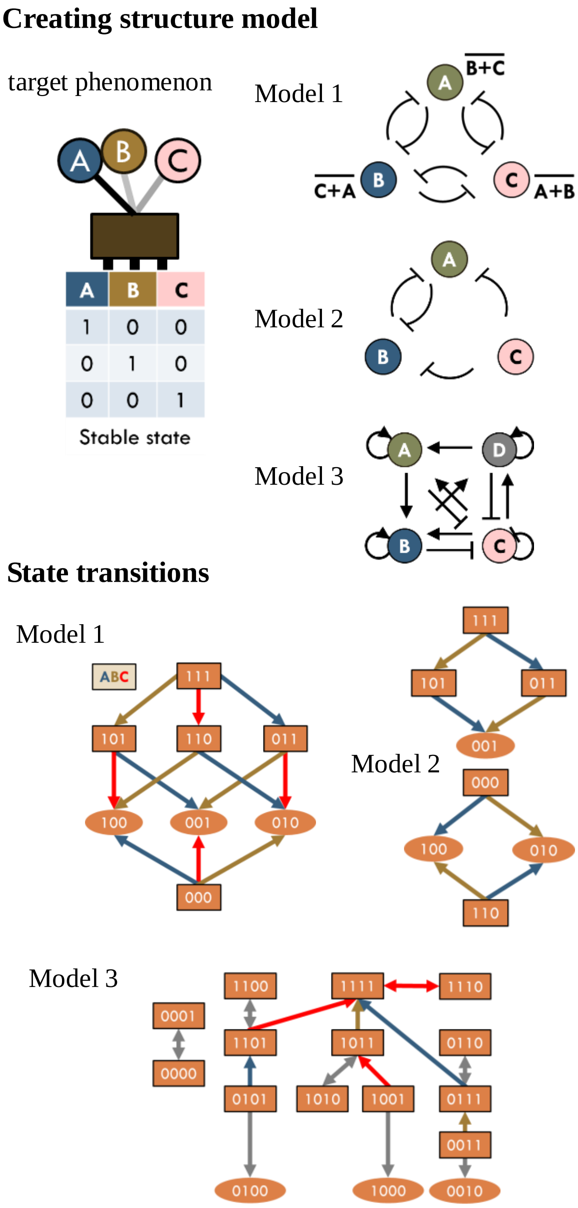


Fig.3a. Process of the proposed method

inductions and selection of expression vector need to be considered (Fig.3b, Creating biological model).

III. RESULTS

We extended a genetic toggle switch from 2-state to 3-state by our proposed procedure. 3-state toggle switch has three stable states as shown in Fig.3a (target phenomenon) and state transition graphs were produced using GINSim (Fig.3a, state transitions). Some of these candidates could be eliminated based on the complexity of state transition graph, in particular those that produced more than 3 stable states, or displayed cyclic behaviors. The selected candidates (Fig.4, structure) were sent to the next step, where dynamic parameters were incorporated in the structures. Fig.4 (state transition graph) shows the state transition graph of the selected candidate for 3-state genetic toggle switch. All states and transitions can be confirmed in this graph. Differential equations for 3-state genetic toggle switch were produced by extending the ones for 2-state genetic toggle switch (Fig.5). The dynamic behaviors of the obtained artificial genetic toggle switch model was analyzed on the system of these differential equations.

Nullclines of differential equations for 2-state genetic toggle switch are shown in Fig.2b (nullclines), showing two stable states as two cross points of two nullclines. These two cross points of nullclines corresponds to two stable states of 2-state genetic toggle switch. In the case of 3-state genetic toggle switch, nullclines were drawn as a three-dimensional diagram (Fig.6) and the diagram shows three stable states of the genetic toggle switch as three cross points of three nullclines in Fig.6a.

Sekine et al. suggested a tunable bistable genetic toggle switch [4]. They considered that their model can help to understand the design principles of the natural diversification. A state switch from ON or OFF must be important function on the genetic switch. We investigated the relation between nullcline diagram and state transition in order to consider a switching system on 3-state genetic toggle switch.

Fig.7 shows a relation of the transition states to the three dimensional diagram of nullclines. Three stable state points (001, 010, 100), transition states (111, 101, 110, 011, 000), separatrix (purple line) and transition regions are indicated in the diagram. The genetic toggle switch can be stabilized at one of the three stable state points in this diagram even if the genetic toggle switch

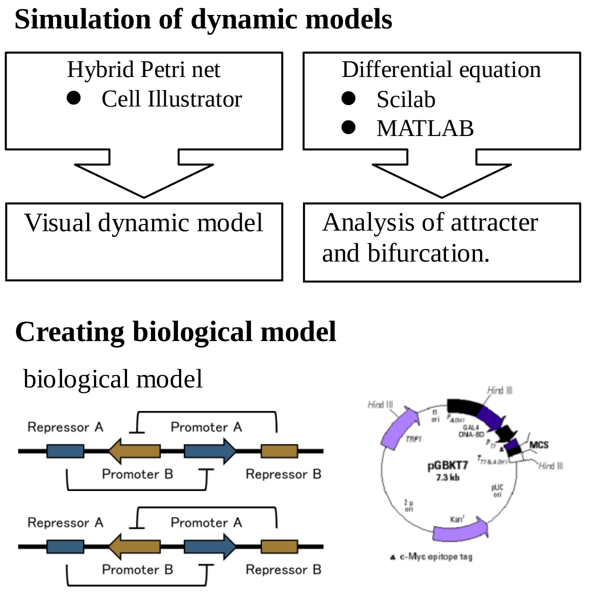


Fig.3b. Process of the proposed method

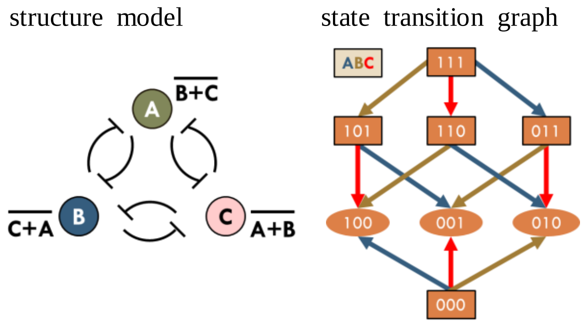


Fig.4. A structure model and state transition graph for extension to 3-state toggle switch

A $$\frac{du_1}{dt} = \frac{a_1}{1 + u_2{}^n + u_3{}^l} - u_1$$

B $$\frac{du_2}{dt} = \frac{a_2}{1 + u_3{}^l + u_1{}^m} - u_2$$

C $$\frac{du_3}{dt} = \frac{a_3}{1 + u_1{}^m + u_2{}^n} - u_3$$

Fig.5 Differential equations for 3-state genetic toggle switch

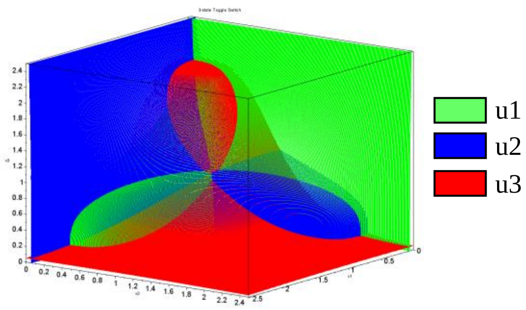


Fig.6a Nullclines of the 3-state genetic toggle

starts from any points. For instance, the genetic toggle switch can be stabilized at 001 stable point if it starts from the dark gray region (red arrow) and if it starts from light gray region, the genetic toggle switch can be stabilized at 010 stable point in the diagram (blue arrow). These state transitions correspond to the ones of the state transition graph in Fig.4.

VI. CONCLUSION

Our proposed method can evaluate the artificial genetic circuits and select the best candidate before incorporating dynamic parameters using state transition graph. Besides the 3-state toggle switch, many biological behaviors such as oscillations and flip-flop can be considered as the targets for artificial genetic circuits. In order to further evaluate of our proposed method, these behaviors will be studied in our future work.

REFERENCES

[1] Gardner, T.S., Cantor, C.R., Collins, J.J., Construction of genetic toggle switch in Escherichia coli, Nature 403:339-342, 2000.
[2] Fauré, A., Thieffry, D., Logical modeling of cell cycle control in eukaryotes: a comparative study, Molecular Biosystems, 5:1769-81, 2009.
[3] Cell Illustrator: http://www.cellillustrator.com/home
[4] Sekine, R., Yamaura, M., Ayukawa, S., Ishimatsu, K., Akama, S., Takinoue, M., Hagiya, M., Kiga, D., Tunable synthetic phenotypic diversification on Waddington’s landscape through autonomous signaling, PNAS,Vol.108, No.44:17969-17973,Nov 1,2011.

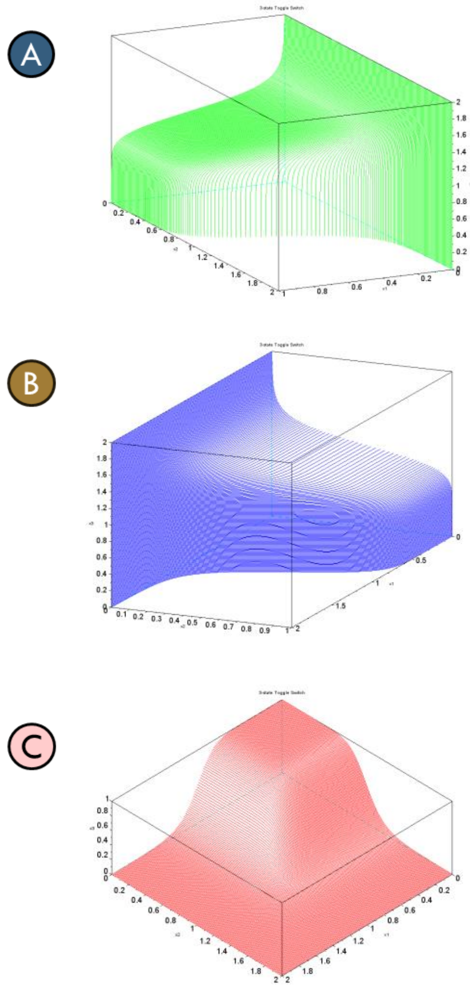


Fig.6b Nullcline diagrams of each differential equation of the 3-state genetic toggle switch

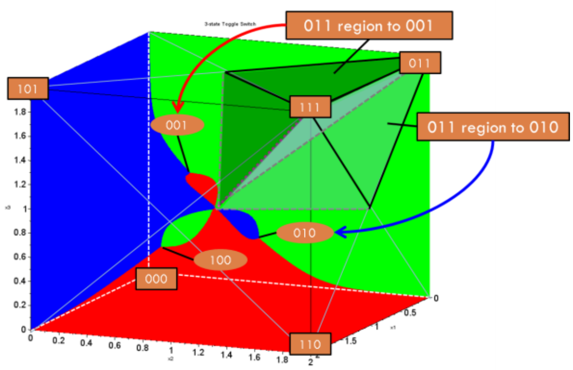


Fig.7. Relation to the transition states

Hierarchy Based on Neighborhood Template about k-Neighborhood Template $\mathcal{A}$-Type Three-Dimensional Bounded Cellular Acceptor

Makoto SAKAMOTO[1], Makoto NAGATOMO[1], Xiaoyang FENG[1], Tatsuma KUROGI[1], Tuo ZHANG[1], Takao ITO[2], Yasuo UCHIDA[2], Tsunehiro YOSHINAGA[3], Satoshi IKEDA[1], Masahiro YOKOMICHI[1], and Hiroshi FURUTANI[1]

[1] *University of Miyazaki, Miyazaki 889-2192, Japan.*
[2] *Ube National College of Technology, Ube, Yamaguchi 755-8555, Japan.*
[3] *Tokuyama College of Technology, Shunan, Yamaguchi 745-8585, Japan.*

Abstract: Recently, due to the advance in dynamic image processing, computer animation, augmented reality (AR), and so both, it has become increasingly apparent that the study of four-dimensional pattern processing (three-dimensional pattern processing with time axis) should be very important. Thus, the study of four-dimensional automata as the computational model of four-dimensional pattern processing has been meaningful. From this point of view, we first introduced a four-dimensional automaton in 2002. In the multi-dimensional pattern processing, designers often use a strategy whereby features are extracted by projecting high-dimensional space on low-dimensional space. In this paper, from this viewpoint, we introduce a new computational model, *k-neighborhood template $\mathcal{A}$-type three-dimensional bounded cellular acceptor* (abbreviated as $\mathcal{A}$-$3BCA(k)$] on four-dimensional input tapes, and discuss hierarchy based on neighborhood template about $\mathcal{A}$-$3BCA(k)$.

Keywords: cellular acceptor, computational complexity, configuration-reader, converter, four-dimension, neighbor

I. INTRODUCTION

Due to the advances in many application areas such as computer animation, dynamic image processing, and so on, the study of four-dimensional pattern processing has been of crucial importance. Thus, the study of four-dimensional automata as the computational models of four-dimensional pattern processing has been meaningful. From this point of view, we first proposed four-dimensional automata as computational models of four-dimensional pattern processing in 2002 [4], and investigated their several accepting powers. By the way, in the multi-dimensional pattern processing, designers often use a strategy whereby features are extracted by projecting high-dimensional space on low dimensional space. So, from this viewpoint, we introduce a new computational model, k-neighborhood template $\mathcal{A}$-type three-dimensional bounded cellular acceptor (abbreviated as $\mathcal{A}$-$3BCA(k)$) on four-dimensional tapes in this paper, and discuss some basic properties. An $\mathcal{A}$-$3BCA(k)$ consists of a pair of a converter and a configuration-reader. The former converts the given four-dimensional tape to three-dimensional configuration. The latter determines whether or not the derived three-dimensional configuration is accepted, and concludes the acceptance or non-acceptance of given four-dimensional tape(see Fig.1). When an input four-

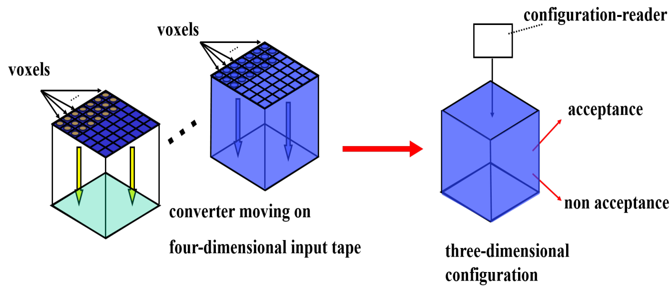


Fig 1: $\mathcal{A}$-$3BCA(k)$.

dimensional tape is presented to the $\mathcal{A}$-$3BCA(k)$, a three-dimensional cellular automaton as the converter first reads it to the future direction at unit speed (i.e., one three-dimensional rectangular array per unit time). From this process, the four-dimensional tape is converted to a configuration of the converter which is a state matrix of a three-dimensional cellular automaton. Second, three-dimensional automaton as the configuration-reader reads the configuration and determines its acceptance. We say that an input four-dimensional tape is accepted by the $\mathcal{A}$-3$BCA(k)$ if and only if the configuration is accepted by the configuration-reader. Therefore, the accepting power of the $\mathcal{A}$-3$BCA(k)$ depends on how to combine the converter and the configuration-reader. An $\mathcal{A}$-3$DBCA(k)$ ($\mathcal{A}$-3$NBCA(k)$) is called a k-neighborhood template $\mathcal{A}$-type three-dimensional deterministic bounded cellular acceptor (k-neighborhood template $\mathcal{A}$-type three-dimensional nondeterministic bounded cellular acceptor). This paper mainly investigates how the difference of the neighborhood template of the converter affects the accepting powers of $\mathcal{A}$-3$DBCA(k)$ ' s. In general, it is well known that two-dimensional digital pictures have 4- and 8-connectedness, and three-dimensional digital pictures have 6- and 26-connectedness. However, we include the remarkable pixel or voxel in neighbor. In other words, we deal with 5- and 9-connectedness in the two-dimensional case, and 7- and 27-connectedness in the three-dimensional case in this paper. In Addition, in our another paper, we mainly investigate how the difference of configuration-reader affects the accepting powers of $\mathcal{A}$-3$DBCA(k)$ ' s. When an input four-dimensional tape is presented to the $\mathcal{A}$-3$BCA(k)$, a three-dimensional cellular automaton as the converter first reads it to the future direction at unit speed (i.e., one three-dimensional rectangular array per unit time), and a two-dimensional cellular automaton as the converter next reads a converted three-dimensional configuration downward at unit speed (i.e., one plane per unit time). From this process, the four-dimensional tape is converted to a configuration of the converter which is a state matrix of a two-dimensional cellular automaton. Second, two-dimensional automaton as the configuration-reader, reads the configuration and determines its acceptance. We say that an input four-dimensional tape is accepted by the $\mathcal{A}$-3$BCA(k)$ if and only if the configuration is accepted by the configuration-reader. Therefore, the accepting power of the $\mathcal{A}$-3$BCA(k)$ depends on how to combine the converter and the configuration-reader. An $\mathcal{A}$-3$DBCA(k)$ ($\mathcal{A}$-3$NBCA(k)$) is called a k-neighborhood template $\mathcal{A}$-type three-dimensional deterministic bounded cellular acceptor (k-neighborhood template $\mathcal{A}$-type three-dimensional nondeterministic bounded cellular acceptor). A DA[1] (NA, DB[5], NB, DO[2], NO, DOP[3], NOP, DP[3], NP, DTM[4], NTM) is called a three-dimensional deterministic finite automaton (three-dimensional nondeterministic finite automaton, deterministic three-dimensional bounded cellular acceptor, nondeterministic three-dimensional bounded celular acceptor, three-dimensional deterministic on-line tessellation acceptor, three-dimensional nondeterministic on-line tessellation acceptor, deterministic three-way parallel/sequential array acceptor, nondeterministic three-way parallel/sequential array acceptor, deterministic four-way parallrel/sequential array acceptor, nondeterministic four-way parallel/sequential array acceptor, three-dimensional deterministic Turing machine, three-dimensional nondeterministic Turing machine). Let $T(M)$ be the set of four-dimensional tapes accepted by a machine M, and let $\mathcal{L}[\mathcal{A}\text{-}3DBCA(k)]$ = $\{T|T{=}T(M)$ for some $\mathcal{A}$-3$DBCA(k)$ $M\}$. $\mathcal{L}[\mathcal{A}\text{-}3NBCA(k)]$, etc. are defined in the same way as $\mathcal{L}[\mathcal{A}\text{-}3DBCA(k)]$.

Let $\sum$ be a finite set of symbols. A *four-dimensional tape* over $\sum$ is a four-dimensional rectangular array of elements of $\sum$. The set of all four-dimensional tapes over $\sum$ is denoted by $\sum^{(4)}$. Given a tape $x \in \sum^{(4)}$, for each integer $j (1 \leq j \leq 4)$, we let $l_j(x)$ be the length of x along the jth axis. The set of all $x \in \sum^{(4)}$ with $l_1(x) = n_1$, $l_2(x) = n_2$, $l_3(x) = n_3$, and $l_4(x) = n_4$ is denoted by $\sum^{(n_1,n_2,n_3,n_4)}$. When $1 \leq i_j \leq l_j(x)$ for each $j (1 \leq j \leq 4)$, let $x(i_1, i_2, i_3, i_4)$ denote the symbol in x with coordinates (i_1, i_2, i_3, i_4). Furthermore, we define $x[(i_1, i_2, i_3, i_4), (i'_1, i'_2, i'_3, i'_4)]$, when $1 \leq i_j \leq i'_j \leq l_j(x)$ for each integer $j (1 \leq j \leq 4)$, as the four-dimensional input tape y satisfying the follwing conditions:

(i) for each $j (1 \leq j \leq 4)$, $l_j(y) = i'_j - i_j + 1$;

(ii) for each $r_1, r_2, r_3, r_4 (1 \leq r_1 \leq l_1(y), 1 \leq r_2 \leq l_2(y), 1 \leq r_3 \leq l_3(y), 1 \leq r_4 \leq l_4(y)), y(r_1, r_2, r_3, r_4) = x(r_1 + i_1 - 1, r_2 + i_2 - 1, r_3 + i_3 - 1, r_4 + i_4 - 1)$. (We call $x[(i_1, i_2, i_3, i_4), (i'_1, i'_2, i'_3, i'_4)]$ the $[(i_1, i_2, i_3, i_4), (i'_1, i'_2, i'_3, i'_4)]$-*segment* of x.)

We let each sidelength of each input tape of these automata be equivalent in order to increase the theoretical interest.

II. MAIN RESULTS

This section investigates how the difference of the neighborhood template of converter affects the accepting powers of $\mathcal{A}$-3$BCA(k)$'s. First, we investigate the difference between the accepting powers of one-neighbor and five-neighbor.

Lemma 1. *Let* $T_1 = x{\in}\{0,1,2\}^{(4)} | \exists n \geq 1$ $[l_1(x) = l_2(x) = l_3(x) = l_4(x) = n$ & $\forall i(1{\leq}i{\leq}n)[x[(1,1,n,i), (n,n,n,i)] = x[(1,1,n,i), (n,n,i,n)]]]$. *Then,* (1) $T_1 \in \mathcal{L}[DA\text{-}3DBCA(7)] \cap \mathcal{L}[DB\text{-}2DBCA(7)] \cap \mathcal{L}[DO\text{-}2DBCA(7)]$, *and* (2) $T_1 \notin \mathcal{L}[TM\text{-}2NBCA(1)]$.

Proof: (1) The proof is omitted here since it is easy to prove. (If necessary, see the proof of Lemma 1(1) in [6].) (2) We can show $T_1 \notin \mathcal{L}[TM\text{-}3DBCA(1)]$. by using the same technique as in the proof of theorem 1 in [7]. Therefore, below, we show $T_1 \in \mathcal{L}[TM\text{-}3DBCA(1)]$. Suppose that there exists a TM-$3DBCA(1)$ $M = (R, B)$ accepting T_1, where R is a converter and B is a configuration-reader.

Let S be the number of states of each cell of R. For each $n \leq 1$, let

$V(n) = \{x{\in}\{0,1,\}^{(4)} | l_1(x) = l_2(x) = l_3(x) = l_4(x) = n+1$ & $\forall i(1{\leq}i{\leq}n)x[(1,1,1,i), (n+1,n+1,n,i)] \in \{0\}^{(3)}$, $x[(1,1,n+1,n+1), (n+1,n+1,n+1,n+1)] \in \{0\}^{(3)}\}$, $W(n) = V(n) \cap T_1$.

Also, for each $x \in V(n)$ and for each $i(1{\leq}i{\leq}n)$, let $\rho(x) \equiv$ the configuration of R just after needing x, $\rho_U(x) \equiv [(1,1,n+1,i),(n+1,n+1,n+1,i)]$-*segment* of $\rho(x)$, and $\rho_D(x) \equiv$ the $[(1,1,n+1,i),(n+1,n+1,n+1,i)]$-*segment* of $\rho(x)$. Further, for each $n \geq 1$, let $C(n) = \{\rho_D(x) \mid x{\in}W(n)\}$. Then, the following two propositions must hold.

Proposition 1. *For each* $i(1 \leq i \leq n)$,
(i) *For any two tapes* x, $y \in V(n)$ *such that their* $[(1,1,1,i),(n+1,n+1,n,i)]$*-segment are identical,* $\rho_U(x) = \rho_U(y)$,
(ii)*For any two tapes* x,$y \in V(n)$ *such that their* $[(1,1,1,i),(n+1,n+1,n,i)]$*-segments are identical,* $\rho_D(x) = \rho_D(y)$.

[**proof**: Since R is deterministic and is of one-neighbor, the proof is easy to see. □]

Proposition 2. *For any two different tapes* x, $y \in W(n)$, $\rho_D(x) \neq \rho_D(y)$.

[**Proof**: Suppose, on the contrary, that $\rho_D(x) = \rho_D(y)$. Consider the tapes $z \in V(n)$ satisfying the following two conditions for each $i(1 \leq i \leq n)$:

① $z[(1,1,1,i), (n+1,n+1,n,i)] = x[(1,1,1,i), (n+1,n+1,n,i)]$,

② $z[(1,1,n+1,i), (n+1,n+1,n+1,i)] = y[(1,1,n+1,i), (n+1,n+1,n+1,i)]$.

Clearly, $x \in W(n) \subseteq T_1$. Thus x is accepted by M. Therefore, $\rho(x)$ is a accepted by B.

On the other hand, it follows from ① and Proposition 1 (i) that $\rho_U(z) = \rho_U(x)$, and it follows from ② and Proposition 1 (ii) that $\rho_D(z) = \rho_D(y)$. Further, from the foregoing assumption $\rho_D(z) = \rho_D(y)$, it follows that $\rho_D(z) = \rho_D(x)$. From $\rho_U(z) = \rho_U(x)$ and $\rho_D(z) = \rho_D(x)$, it follows that $\rho(z) = \rho(x)$.
Since $\rho(x)$ is accepted by B, $\rho(z)$ is also accepted by B. Consequently, z is also accepted by M. This is a contradiction.(Note that $z \notin T_1$.) □]

Proof of Lemma 1(*continued*) **:** As is early seen, $|W(n)| = 2^{n(n+1)(n+1)}$ and $|C(n)| \leq s^{(n+1)(n+1)}$
Therefore, it follows for large n that $|W(n)| < |C(n)|$. Consequently, it follows for such large n that there must be two different tapes x,$y \in W(n)$ such that $\rho_D(x) \neq \rho_D(y)$. This contradicts Proposition 2. □

Theorem 1. *For each* $\mathcal{A} \in \{DA, NA, DB, NB, DO, NO, DOP, NOP, DP, NP, TM\}$ *and for each* $X \in \{D, N\}$, $\mathcal{L}[\mathcal{A}\text{-}3XBCA(1)] \subsetneq \mathcal{L}[\mathcal{A}\text{-}3XBCA(7)]$.

Proof: The inclusion relarion holds immediately definitions. Further, it is easily seen from Lemma 1 and Proposition 1 in [1] that the theorem holds. □

We investigate the difference between the accepting powers of seven-neighborhood and twenty-seven-neighbor. As shown later in Theorem 2 and 3, different situations emerge depending on whether the converter is deterministic or nondeterministic.
First, we consider the case when the converter is deterministic.

Theorem 2. *For each* $\mathcal{A} \in \{ DA, NA, DB, NB, DO, NO \}$, $\mathcal{L}[\mathcal{A}\text{-}3DBCA(7)] \subsetneq \mathcal{L}[\mathcal{A}\text{-}3DBCA(27)]$.

Proof : The inclusion relation holds immediately from definitions. Further, it is easily seen from Theorem 3 and Proposition 1 in [7] that the theorem holds. □

We conclude this action by investigating the difference between the accepting powers of seven-neighbor and twenty-seven-neighbor for the case when converter is nondeterministic.

Theorem 3. *For each* $\mathcal{A} \in \{ DA, NA, DB, NB, DO, NO, DOP, NOP, DP, NP, TM \}$, $\mathcal{L}[\mathcal{A}\text{-}3DBCA(7)] = \mathcal{L}[\mathcal{A}\text{-}3NBCA(27)]$.

Proof: For each $\mathcal{A} \in \{ DA, NA, DB, NB, DO, NO, DOP, NOP, DP, NP, TM \}$, it is obvious that $\mathcal{L}[\mathcal{A}\text{-}3NBCA(7)] \subseteq \mathcal{L}[\mathcal{A}\text{-}3NBCA(27)]$ from definitions. Below, we show that $\mathcal{L}[\mathcal{A}\text{-}3NBCA(7)] \supseteq \mathcal{L}[\mathcal{A}\text{-}3NBCA(27)]$.
a given $\mathcal{A} \in \{ DA, NA, DB, NB, DO, NO, DOP, NOP, DP, NP, TM \}$, let $M = (R, B)$ be an $\mathcal{A}$-$3NBCA(27)$. We now consider the $\mathcal{A}$-$3NBCA(7)$ M'

= (R', B') which acts as follows.

Suppose that a four-dimensional input tape x with each sidelength n ($n \geq 2$) is presented to M. Since R' has seven neighbors, each voxel of R' can refer directly to the states of north, south, east, west, up, down, neighbor voxels, but not other neighbor voxels. Therefore, by guessing the six states of neighbor voxels which cannot be referred to directly, and by checking whether or not this guess is convect, each voxel of R' simulates the action of the corresponding voxel of R. In fact, M' = (R', B') acts as follows.

(i) Action of the cnoverter R'

R' starts to act with the same initial configuration with one of R. That is, every voxel reading the boundary symbol is in $q_{\#}$, which is the boundary state of R, and all of the other voxels are in q_0, which is the initial state of R. Next, the seven states of neighbor voxels of R is easily simulated by R', and the twenty states of other neighbor voxels of R are guessed nondeterministically by R.

(ii) Action of the configuration-reader B'

B' accepts the configuration of R' just after reading x (say, $\rho(x)$), if and only if the following ①, ②, ③ are satisfied.
① For each i, j, k ($1 \leq i, j, k \leq n$), no (i, j, k)-voxel enters the dead state.
② The information guessed by any (i, j, k)-voxel on the last three-dimensional rectangular array of four-dimensional input tape is correct.
③ Let h be a mapping extracting one state, which is obtained by simulating the action of the corresponding voxel of R, from the states stored in the state of (i, j, k)-voxel. Let h be a projection which is obtained by extending the mapping h. Then $h(\rho(x))$ is accepted by B.
In ① and ② in the foregoing, check whether or not R' can correctly simulate the action of R. Therefore, it is clear thet $h(\rho(x))$ reflects the configuration of R just after reading x, if ① and ② are satisfying.

is easily seen that $T(M') = T(M)$ for $M = (R', B')$. Thus, $\mathcal{L}[\mathcal{A}\text{-}3NBCA(7)] \supseteq \mathcal{L}[\mathcal{A}\text{-}3NBCA(27)]$. This completes the proof of the theorem. □

III. CONCLUSIONS

In this paper, we investigated how the difference of neighborhood template of the converter affects the accepting powers of k-neighborhood template $\mathcal{A}$-type three-dimensional bounded cellular acceptor(abbreviated as $\mathcal{A}$-3$BCA(k)$). Generally speaking, when the converter is deterministic, the accepting power of the $\mathcal{A}$-3$BCA(k)$ tends to be more powerful as the number of neighborhood cells of the converter increases or the accepting power of the configuration-reader is more powerful. However, this tendency is not always true when the converter is nondeterministic.

We conclude this paper by giving two open problems.

(1) For each $\mathcal{A} \in \{DOP, NOP, DP, NP, DTM, NTM\}$, $\mathcal{L}[\mathcal{A}\text{-}3DBCA(7)] \subsetneq \mathcal{L}[\mathcal{A}\text{-}3DBCA(27)]$?

(2)$\mathcal{L}[NO\text{-}3NBCA(7)] \subsetneq \mathcal{L}[DOP\text{-}3NBCA(27)]$?

REFERENCES

[1] M.Blum and C.Hewitt, "Automata on a two-dimensional tape", IEEE Symposium on Switching and Automata Theory, pp. 155-160, 1967.

[2] K.Inoue and A.Nakamura, "An n-dimensional on-line tessellation acceptor", The Transactions of the IECE, Japan, J59-D(4), pp. 299-236, 1976.

[3] A.Rosenfeld and D.L.Milgram, "Parallel/sequential array acceptor", Information Processing Letters, 2, pp. 43-46,1976.

[4] M.Sakamoto, H.Okabe, S.Nagami, S.Taniguchi, T.Makino, Y.Nakama, M.Saito, M.Kono, and K.Inoue, "A note on four-dimensional finite automata", WSEAS Transactions on Computers, Issue 5, Vol.3, pp. 1651-1656, 2004.

[5] S.Seki, "Real-time recognition of two-dimensional tapes by cellular automata", Information Sciences, 19, pp. 179-198, 1979.

[6] H.Taniguchi, K.Fujii, K.Inoue and I.Takanami, "k-Neighborhood template $\mathcal{A}$-type two-dimansional bounded cellular acceptors (Part II)", Technical Report No. AL82-19, IECE, Japan, 1982.

[7] H.Taniguchi, K,Inoue and I.Takanami, "k-Neighborhood template A $\mathcal{A}$-type two-dimensional bounded cellular acceptor", The Transactions of the IECE, Japan, J69-D(3), pp. 291-301, 1986.

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

Hierarchy Based on Configuration-Reader about k-Neighborhood Template $\mathcal{A}$-Type Three-Dimensional Bounded Cellular Acceptor

Makoto SAKAMOTO[1], Tuo ZHANG[1], Tatsuma KUROGI[1], Makoto NAGATOMO[1], Xiaoyang FENG[1], Yasuo UCHIDA[2], Takao ITO[2], Tsunehiro YOSHINAGA[3], Satoshi IKEDA[1] , Masahiro YOKOMICHI[1], and Hiroshi FURUTANI[1]

[1] *University of Miyazaki, Miyazaki, Miyazaki 889-2192, Japan.*

[2] *Ube National College of Technology, Ube, Yamaguchi 755-8555, Japan.*

[3] *Tokuyama College of Technology, Shunan, Yamaguchi 745-8585, Japan.*

Abstract: Blum and Hewitt first proposed two-dimensional automata as computational models of two-dimensional pattern processing − two-dimensional finite automata and marker automata, and investigated their pattern recognition abilities in 1967. Since then, many researchers in this field have investigated the properties of automata on two- or three-dimensional tapes. On the other hand, the question of whether or not processing four-dimensional digital patterns is more difficult than processing two- or three-dimensional ones is of great interest from both theoretical and practical standpoints. Thus, the study of four-dimensional automata as the computational models of four-dimensional pattern processing has been meaningful. From this point of view, we are interested in four-dimensional computational models, In this paper, we introduce a new four-dimensional computational model, *k-neighborhood template $\mathcal{A}$-type three-dimensional bounded cellular acceptor* on four-dimensional input tapes, and investigate about hierarchy based on configuration-reader about this model.

Keywords : cellular acceptor, configuration-reader, converter, finite automaton, four-dimension, on-line tessellation acceptor, parallel/sequential array acceptor, Turing machine

I. INTRODUCTION AND PRELIMINARIES

In 2002, we first introduced a four-dimensional automaton, and investigated some properties [4]. In general, in the multi-dimensional pattern processing, designers often use a strategy whereby features are extracted by projecting high-dimensional space on low-dimensional space. In this paper, from this viewpoint, we introduce a new computational model, *k-neighborhood template $\mathcal{A}$-$type$ three-dimensional bounded cellular acceptor* (abbreviated as $\mathcal{A}$-3$BCA(k)$) on four-dimensional tapes, and discuss some basic properties. An $\mathcal{A}$-3$BCA(k)$ consists of a pair of a *converter* and a *configuration-reader*. The former converts the given four-dimensional tape to the three-dimensional configuration and the latter determines the acceptance or nonacceptance of given four-dimensional tape whether or not the derived three-dimensional configuration is accepted. When a four-dimensional input tape is presented to the $\mathcal{A}$-3$BCA(k)$, a three-dimensional cellular automaton as the converter first reads it to the future direction at unit speed (i.e., one three-dimensional rectangular array per unit time). From this process, the four-dimensional tape is converted to a configuration of the converter which is a state matrix of a three-dimensional cellular automaton. Second, three-dimensional automaton as the configuration-reader, reads the configuration and determines its acceptance. We say that a four-dimensional input tape is accepted by the $\mathcal{A}$-3$BCA(k)$ if and only if the configuration is accepted by the configuration-reader. Therefore, the accepting power of the $\mathcal{A}$-3$BCA(k)$ depends on how to combine the converter and the configuration-reader. An $\mathcal{A}$-3$DBCA(k)$ ($\mathcal{A}$-3$NBCA(k)$) is called a k-neighborhood template $\mathcal{A}$-type three-dimensional deterministic bounded cellular acceptor (k-neighborhood template $\mathcal{A}$-type three-dimensional nondeterministic bounded cellular acceptor). A DA[1] (NA, DB[5], NB, DO[2], NO, DOP[3], NOP, DP[3], NP, DTM[4], NTM) is called a three-dimensional deterministic finite automaton (three-dimensional nondeterministic finite automaton, deterministic three-dimensional bounded cellular acceptor, nondeterministic three-dimensional bounded celular acceptor, three-dimensional deterministic on-line tessellation acceptor, three-dimensional nondeterministic on-line tessellation acceptor, deterministic three-way parallel/sequential array acceptor, nondeterministic three-way parallel/sequential array acceptor, deterministic three-way parallrel/sequential array acceptor, nondeterministic four-way parallel/sequential array acceptor, three-dimensional deterministic Turing machine, three-dimensional nondeterministic Turing machine). Let $T(M)$ be the set of four-dimensional tapes accepted by a machine M, and let $\mathcal{L}$[$\mathcal{A}$-3$DBCA(k)$] = $\{T|T$=$T(M)$ for some $\mathcal{A}$-3$DBCA(k)$ $M\}$.

$\mathcal{L}[\mathcal{A}\text{-}3NBCA(k)]$, etc. are defined in the same way as $\mathcal{L}[\mathcal{A}\text{-}3DBCA(k)]$.

Let $\sum$ be a finite set of symbols. A *four-dimensional tape* over $\sum$ is a four-dimensional rectangular array of elements of $\sum$. The set of all four-dimensional tapes over $\sum$ is denoted by $\sum^{(4)}$. Given a tape $x \in \sum^{(4)}$, for each integer $j(1 \leq j \leq 4)$, we let $l_j(x)$ be the length of x along the jth axis. The set of all $x \in \sum^{(4)}$ with $l_1(x) = n_1$, $l_2(x) = n_2$, $l_3(x) = n_3$, and $l_4(x) = n_4$ is denoted by $\sum^{(n_1,n_2,n_3,n_4)}$. When $1 \leq i_j \leq l_j(x)$ for each $j(1 \leq j \leq 4)$, let $x(i_1, i_2, i_3, i_4)$ denote the symbol in x with coordinates (i_1, i_2, i_3, i_4), as shown in Fig.1. Furthermore, we define $x[(i_1, i_2, i_3, i_4), (i'_1, i'_2, i'_3, i'_4)]$, when $1 \leq i_j \leq i'_j \leq l_j(x)$ for each integer $j(1 \leq j \leq 4)$, as the four-dimensional input tape y satisfying the follwing conditions:

(i) for each $j(1 \leq j \leq 4)$, $l_j(y) = i'_j - i_j + 1$;

(ii) for each $r_1, r_2, r_3, r_4 (1 \leq r_1 \leq l_1(y), 1 \leq r_2 \leq l_2(y), 1 \leq r_3 \leq l_3(y), 1 \leq r_4 \leq l_4(y)), y(r_1, r_2, r_3, r_4) = x(r_1 + i_1 - 1, r_2 + i_2 - 1, r_3 + i_3 - 1, r_4 + i_4 - 1)$.

(We call $x[(i_1, i_2, i_3, i_4), (i'_1, i'_2, i'_3, i'_4)]$ the $[(i_1, i_2, i_3, i_4), (i'_1, i'_2, i'_3, i'_4)]$-*segment* of x.)

We let each sidelength of each input tape of these automata be equivalent in order to increase the theoretical interest.

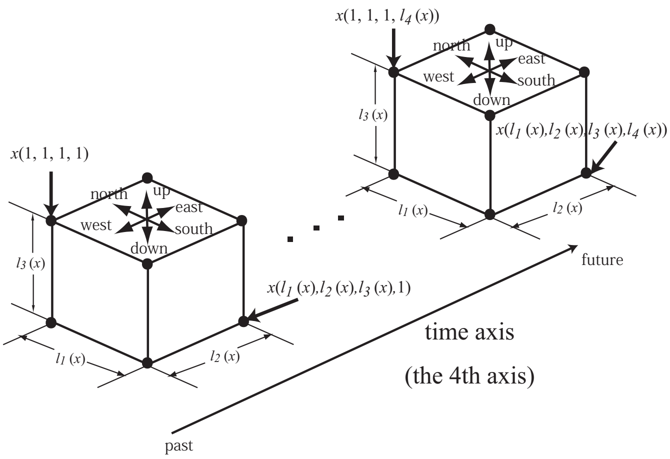


Fig. 1: Four-dimensional Input Tape.

II. MAIN RESULTS

This section investigates how the difference of configuratitu-reader affects the accepting powers of $\mathcal{A}$-$3BCA(k)$'s. First, we start to investigate the case when the converter is deterministic.

Lemma 1. Let T_1 = $\{x \in \{0,1,2\}^{(4)} | \exists n \geq 1$ $[l_1(x) = l_2(x) = l_3(x) = l_4(x) = n+1$ & $\exists i(1 \leq i \leq n)[x(i,n+1,n+1,n+1) = 2$ & (each symbol on the remaining parts is "0" or "1") & $x[(i,1,n+1,n+1), (i,n,n+1,n+1)] \neq x[(n+1,1,n+1,n+1), (n+1,n,n+1,n+1)]]]\}$. Then, (1) $T_1 \in \mathcal{L}[NA\text{-}3DBCA(1)]$, and (2) $T_1 \notin \mathcal{L}[DA\text{-}3DBCA(27)]$.

Proof: It is easily seen that there exists a nondeterministic two-dimensional finite automaton accepting the set of two-dimensional tapes which are obtained by extracting the bottom plane from the tape contained in T_1. Therefore, (1) holds. On the other hand, the proof of (2) is similar to that of Lemma 2(2) in [7]. □

Lemma 2. Let T_2 = $\{x \in \{0,1\}^{(4)} \mid \exists\ n \geq 1$ $[l_1(x) = l_2(x) = l_3(x) = l_4(x) = 2n$ & $[x(1,1,2n,2n),(2n,n,2n,2n) = x[(1,n+1,2n,2n),(2n,2n,2n,2n)]]]\}$. Then, (1) $T_2 \in \mathcal{L}[DOP\text{-}3DBCA(1)]$, and (2) $T_2 \notin \mathcal{L}[NO\text{-}3DBCA(27)]$.

Proof: (1) Note that there exists a deterministic one-way parallel sequential array acceptor accepting the set of two-dimensional tapes obtained by extracting the bottom plane of the first cube from the tape contained in T_2. It is easily seen from this fact that (1) holds. (2) The proof is similar to that of Lemma 2(2). Suppose that there exists an NO-$3DBCA(27)$ $M = (R, B)$ accepting T_2, where R is a converter and B is a configuration-reader. Let K be the set of each cell of $B \in NO$, and $|K| = s$.

For each $n \geq 1$, let $V(n)=\{x\{0,1\}^{(4)} \mid l_1(x)=l_2(x) = l_3(x) = l_4(x)= 2n$ & $x[(1,1,1,1), (2n, 2n, 2n, 2n\text{-}1)] \in \{0\}^{(4)}\}$, $V'(n) = V(n) \cap T_2$, $W(n) = \{w \in K^{(2)} \mid l_1(w) = 2n$ & $l_2(w) = 1\}$ ($K^{(2)}$ means the set of all two-dimensional tapes over $\sum$.). For each $x \in V(n)$, let $\rho(x) \equiv$ the configuration of R just after reading x, $\rho_W(x)$ = the west half of $\rho(x)$, and $\rho_E(x) \equiv$ the east half of $\rho(x)$. Further, for each $x \in V'(n)$, let Run$(x) = \{z \in K^{(2)} \mid z$ is a run of B on $\rho(x)$ whose lower right corner symbol is an accepting state of $B.\}$ and $r(x) = \{z[(1,n),(2n,n)] \mid z \in$ Run$(x)\} \subseteq W(n)$. Then, the following proposition must hold.

Proposition 1. For any two different tapes x and y in $V'(n)$, $r(x) \cap r(y) = \phi$.

[Proof: The proof is similar to that of Proposition 4 in [8]. □]

Proof of Lemma 2 ($continued$)**:** As is easily seen,

$$|V'(n)| = 2^{2n^2} \text{ and } |W(n)| \leq s^{2n}$$

Therefore, it follows for large n that $|V'(n)| > |W(n)|$.

Consequently, ii follows for such large n that there must be two different tapes x and y in $V'(n)$ such that $r(x) \cap r(y) \neq \phi$. This contradicts Proposition 1. □

Lemma 3. Let T_3 = $\{x\in\{0,1\}^{(4)} \mid \exists n\geq 1[l_1(x) = l_2(x) = l_3(x) = l_4(x) = 2n$ & $x[(1,1,2n,2n), (n,2n,2n,2n)] = x[(n+1,1,2n,2n), (2n,2n,2n,2n)]]\}$. Then, (1) $T_3 \in \mathcal{L}[DP\text{-}3DBCA(1)]$, and (2) $T_3 \notin \mathcal{L}[NOP\text{-}3DBCA(27)]$.

Proof: (1) Note that there exists a deterministic two-way parallel sequential array acceptor accepting the set of two-dimensional tapes which are obtained by extracting the bottom plane of the last cube from the tape contained in T_3. It is easily seen, from this fact, that (1) holds.
(2) Suppose that there exists an NOP-3$DBCA$(27) M = (R,B) accepting T_3. Let s be the number of states of each cell of $B\in NOP$. For each $n\geq 1$, let $V(n)$ = $\{x\in\{0,1\}^{(4)}\ l_1(x) = l_2(\mathrm{x}) = l_3(x) = l_4(x) = 2n$ & $x[(1,1,1,1), (2n,2n,2n,2n\text{-}1)]\in\{0\}^{(4)}\}$, $V'(n) = V(n) \cap T_3$. For each $x \in V(n)$, let $\rho(x) \equiv$ the configuration of R just after reading x, $\rho_N(x) \equiv$ the nouth half of $\rho(x)$, and $\rho_S(x) \equiv$ the south half of $\rho(x)$. Furthermore, for each $x \in V'(n)$, let conf(x) $\equiv$ the set of possible configuration of B just after $\rho_N(x)$ is read, when $\rho(x)$ is accepted by B. (Note that $\rho(x)$ is accepted by B since each tape in $V'(n)$ is accepted by M.) Then, the following two propositions must hold. (The proofs are omitted here. If necessary, see proofs of Lemmas 7 and 8 in [6].)

Proposition 2. (i) For any two tapes x and y in $V'(n)$ such that their [(1, 1, $2n$, $2n$), (n, $2n$, $2n$,$2n$)]-segments are identical, $\rho_N(x) = \rho_N(y)$, and (ii) For any two tapes x and y in $V'(n)$ such that their [(n + 1, 1, $2n$, $2n$), ($2n$, $2n$, $2n$, $2n$)]-segnents are identical, $\rho_S(x) = \rho_S(y)$.

Proposition 3. For any two different tapes x and y in $V'(n)$, conf(x) $\cap$ conf(y) = ϕ.

Proof of Lemma 3 (*continued*): As is easily seen,

$$|V'(n)| = 2^{2n^2}$$

Let $t(n)$ be the total number of different configurations of R just after reading north halves of configurations of R just after reading tapes in $V'(n)$. Clearly,

$$t(n)\leq s^{2n}$$

Therefore, it follows for large n that

$$|V'(n)| > t(n)$$

Consequently, it follows for such large n that there must be two different tapes x and y in $V'(n)$ such that conf(x)$\cap$conf(y) $\neq \phi$. This contradicts Proposition 3. □

Lemma 4. Let T_4 be the set of three-dimensional tapes described in Lemma 1 in [7]. Then, (1) $T_4 \in\mathcal{L}[NB\text{-}3DBCA(1)]$, and (2) $T_4 \notin\mathcal{L}[DOP\text{-}3DBCA(27)]$.

Proof: (1) It is easily seen that there exists a nondeterministic one-dimensional branded eellolar automaton accepting the set of two-dimensional tapes which are obtained hy attracting the bottom plane from the tape contained in T_4. Therefore, (1) holds. On the other hand, the proof of (2) is shown Lemma 1 in [7]. □

Lemma 5. Let T_5 = $\{x \in \{0,1\}^{(4)} \mid \exists\ n \geq 1\ [l_1(x) = l_2(x)=l_3(x)=l_4(x)=2n$ & $[x(1,1,2n,2n), (n,n,2n,2n)]\neq x[(n+1,n+1,2n,2n),(2n,2n,2n,2n)]]\}$. Then, (1) $T_5 \in \mathcal{L}[NB\text{-}3DBCA(1)]$, and (2) $T_5 \notin \mathcal{L}[NA\text{-}3DBCA(27)]$.

Proof: (1) It is easily seen that there exists a nondeterministic one-dimensional bounded cellular automaton accepting the set of two-dimensional tapes which are obtained by extracting the bottom plane from the tape contained in T_5. Therefore, (1) holds. On the other hand, the proof of (2) is similar to that of Lemma 2 in [7]. □

From the foregoing lemmas, we can obtain the following theorem when the converter is deterministic.

Theorem 1. For each $k \in \{1, 7, 27\}$,

(1) $\mathcal{L}[DA\text{-}3DBCA(k)] \subsetneq \mathcal{L}[\text{NA-}3DBCA(k)] \subsetneq \mathcal{L}[NB\text{-}3DBCA(k)] = \mathcal{L}[NO\text{-}3DBCA(k)] \subsetneq \mathcal{L}[NOP\text{-}3DBCA(k)] \subsetneq \mathcal{L}[NP\text{-}3DBCA(k)]$
(2) $\mathcal{L}[DB\text{-}3DBCA(k)] \subsetneq \mathcal{L}[NB\text{-}3DBCA(k)]$,
(3) $\mathcal{L}[DO\text{-}3DBCA(k)] \subsetneq \mathcal{L}[NO\text{-}3DBCA(k)]$,
(4) $\mathcal{L}[DB\text{-}3DBCA(k)] \subsetneq \mathcal{L}[DOP\text{-}3DBCA(k)] \subsetneq \mathcal{L}[DP\text{-}3DBCA(k)]$, and
(5) $\mathcal{L}[DO\text{-}3DBCA(k)] \subsetneq \mathcal{L}[DOP\text{-}3DBCA(k)] \subsetneq \mathcal{L}[NOP\text{-}3DBCA(k)]$.

Proof: It is clear from Proposition 1 in [7] that the inclusion relations hold. Therefore, below, we show that the proper inclusion relations held for each $k \in \{1,7,27\}$.
(1): It is obvious from Proposition l in [7] that $\mathcal{L}[NB\text{-}3DBCA(k)] = \mathcal{L}[NO\text{-}3DBCA(k)]$. From Lemma 1, $\mathcal{L}[DA\text{-}3DBCA(k)] \subsetneq \mathcal{L}\ [NB\text{-}3DBCA(k)]$ holds, and from Lemma 5, $\mathcal{L}\ [NB\text{-}3DBCA(k)] \subsetneq \mathcal{L}[NB\text{-}3DBCA(k)]$ holds. In addition, it is obvious from Proposition 1 in [7] that $\mathcal{L}[DOP\text{-}3DBCA(k)] \subseteq \mathcal{L}[NP\text{-}3DBCA(k)]$. It follows from this and Lemma 2 that $\mathcal{L}[NO\text{-}3DBCA(k)] \subsetneq \mathcal{L}[NOP\text{-}3DBCA(k)]$ holds. Further, it is also obvious from Proposicion 1 in [7] that $\mathcal{L}[DP\text{-}3DBCA(k)] \subseteq \mathcal{L}[NP\text{-}3DBCA(k)]$. It follows from this and Lemma 3 that $\mathcal{L}[NOP\text{-}3DBCA(k)] \subsetneq \mathcal{L}[NP\text{-}3DBCA(k)]$ holds.
(2) and (3) : These are easily proved from Lemma 4 and Proposition 1 in [7].

(4) and (5) : These are also easily proved from Lemmas 4,5,6 and Proposition 1 in [7]. □

Next, we investigate the case when the converter is nondeterministic.

Lemma 6. For each $k \in \{1,7,27\}$, (1) $\mathcal{L}[NO\text{-}3NBCA(k)] \subseteq \mathcal{L}[DA\text{-}3NBCA(k)]$, (2) $\mathcal{L}[NO\text{-}3NBCA(k)] \subseteq \mathcal{L}[DB\text{-}3NBCA(k)]$, and (3) $\mathcal{L}[NO\text{-}3NBCA(k)] \subseteq \mathcal{L}[DO\text{-}3NBCA(k)]$.

Proof: (1) We prove only $\mathcal{L}[NO\text{-}3NBCA(1)]$ (The other cases are proved similarly.) Let $M= (R, B)$ be an arbitrary $NO\text{-}3NBCA(1)$, and let K_R and K_B be the set of states of R and B, respectively. Further, let $M' = (R', B')$ be a $DO\text{-}3NBCA(1)$ which acts as follows for a given four-dimensional tape x with each sidelength is n ($n \geq 1$).

(i) Actions of the converter R'

At each time, each (i, j, k, l)-cell $(1 \leq i, j , k, l, \leq n)$ of R' simulates the action of the corresponding cell of R on x at the same time. In parallel to this action, the cell selects nondeterministically a state in K_B (we let $q(i, j, k, l)$ be the state) and stores the state in its state, when the cell reads a symbol on the top plane of the first cube of x. Here, $q(i,j,k,l)$ is a guessed state of B which the (i, j, k, l)- cell of B will enter by reading the configuration of R just after reading x; $q(i, j, k, l)$ will have been stored in the state of the cell until x is completed to read.

(ii) Actions of the configuration reader B'

For each $i, j, k, l (1 \leq i, j , k, l \leq n)$, let $\rho(i, j, k, l)$ be a state in K_R which the (i, j, k, l)-cell of R' continues to simulate the action of corresponding cell of R and enters. B' accepts a configuration of R' just after reading x if and only if the following two conditions are satisfied.

① For each $i, j, k, l (1 \leq i, j, k, l \leq n)$, the (i, j, k, l)-cell can enter $q(i, j, k, l)$ when it reads $\rho(i, j, k, l)$.

② $q(n,n,n,n)$ is an accepting state of B.

It is easily seen that $T(M') = T(M)$ for $M = (R',B')$. This completes the proof of the lemma. □

From Lemma 6 and from Proposition 1 in [7], we can obtain directly the following theorem when the converter is nondeterministic.(The proof is omittied here.) It is of great interest to compare the following Theorem 2 with Theorem 1 mentioned for the deterministic case.

Theorem 2. For each $k \in \{1, 7, 27\}$, $\mathcal{L}[DA\text{-}3NBCA(k)] = \mathcal{L}[NA\text{-}3NBCA(k)] = \mathcal{L}[DB\text{-}3NBCA(k)] = \mathcal{L}[NB\text{-}3NBCA(k)] = \mathcal{L}[DO\text{-}3NBCA(k)] = \mathcal{L}[NO\text{-}3NBCA(k)]$.

III. CONCLUSIONS

In this paper, we investigated how the difference of the configuration-reader affects the accepting powers of k-neighborhood template $\mathcal{A}$-type three-dimensional bounded cellular acceptor(abbreviated as $\mathcal{A}\text{-}3BCA(k)$). As the results, we showed that when the configuration-reader is deterministic, the $\mathcal{A}\text{-}3BCA(k)$ which is the converter is nondeterministic is more powerful than the $\mathcal{A}\text{-}3BCA(k)$ which is the converter is deterministic. However, this tendency is not always true when the configuration-reader is nondeterministic.

We conclude this paper by giving a few open problems.

(1) Accepting powers in the case of alternating version of configuration-reader.

(2) Closure properties of $\mathcal{A}\text{-}3BCA(k)$.

(3) Recognizability of topological four-dimensional input tapes by $\mathcal{A}\text{-}3BCA(k)$.

REFERENCES

[1] M.Blum and C.Hewitt, “Automata on a two-dimensional tape”, IEEE Symposium on Switching and Automata Theory, pp. 155-160, 1967.

[2] K.Inoue and A.Nakamura, “An n-dimensional on-line tessellation acceptor”, The Transactions of the IECE, Japan, J59-D(4), pp. 299-236, 1976.

[3] A.Rosenfeld and D.L.Milgram, “Parallel/sequential array acceptor”, Information Processing Letters, 2, pp. 43-46,1976.

[4] M.Sakamoto, H.Okabe, S.Nagami, S.Taniguchi, T.Makino, Y.Nakama, M.Saito, M.Kono, and K.Inoue, “A note on four-dimensional finite automata”, WSEAS Transactions on Computers, Issue 5, Vol.3, pp. 1651-1656, 2004.

[5] S.Seki, “Real-time recognition of two-dimensional tapes by cellular automata”, Information Sciences, 19, pp. 179-198, 1979.

[6] H.Taniguchi, K.Fujii, K.Inoue and I.Takanami, “k-Neighborhood template $\mathcal{A}$-type two-dimansional bounded cellular acceptors (Part II)”, Technical Report No. AL82-19, IECE, Japan, 1982.

[7] H.Taniguchi, K,Inoue and I.Takanami, “k-Neighborhood template A $\mathcal{A}$-type two-dimensional bounded cellular acceptor”, The Transactions of the IECE, Japan, J69-D(3), pp. 291-301, 1986.

[8] H.Taniguchi, K.Inoue and I.Takanami, “Hierarchical properties of the k-neighborhood template $\mathcal{A}$-type 2-dimensional bounded cellular acceptor”, The Transactions of the IECE, Japan, J69-D(7), pp.1025-1034, 1986.

Handicap of Othello Game

Yusuke KATO[1] Satoshi IKEDA[2] Takao ITO[3] Makoto SAKAMOTO[2]

[1] Graduate course in Computer Science and System Engineering, University of Miyazaki
[2] Department of Computer Science and System Engineering, University of Miyazaki
[3] Department of Business Administration, Ube National College of Technology

Tel : 81-985-58-7414; Fax : 81-985-58-7414
e-mail address: tf13002@student.miyazaki-u.ac.jp

Abstract: In the game of Go, it is known experientially that Black (=the first move) is advantageous. Thus, in the game of Go, White (=the defensive hand) has received a handicap for this compensation. There is Othello game well known as the Go. However, the handicap of White in Othello game has been hardly considered until now. This research considered quantitatively the handicap in Othello game from a viewpoint of points. As our conclusions of Othello game is the followings.

- The advantage of the defensive hand on board of 4 × 4 is -0.641 points on average.
- The advantage of the defensive hand on board of 4 × 6 is 0.778 points on average.

Keywords: Othello, handicap, handicap stones, first move advantage

1. Introduction

In the game of Go, how to put handicap has been discussed for a long time. Its handicap is given by means of compensation and handicap stones as known well.

It is known experientially that Black (= the first move) has a first move advantage. Thus, in the game Go, White (= the defensive hand) has received a handicap for this compensation. On a 19×19 goban, this predetermined compensation for Black's first move advantage, called "komi", means that White may receive somewhere on average of +6.5 points.

There are handicap stones as another handicap in the game of Go. The rank difference within a given amateur ranking system is one guide to how many handicap stones should be given to make the game a more equal contest. Generally, it is considered that one stone is 13-16 points in term of points, although this figure is not constant over levels.

There is a Reversi well known as the Go. Reversi is also called Othello game in general. It is thought be given a handicap similar to the Go in Othello game. However, the discussion of quantitative handicap was not in the game of Othello such as in the game of Go.

This research considered quantitatively the handicap about a first move advantage and a handicap stones in the Othello game from a viewpoint of points.

Throughout this paper, we assume that the handicap for Black's first move advantage is compensation given to White. Therefore a negative handicap means a situation with an advantageous to White.

2. Method and Analysis

In the game of Othello, regular board is 8 × 8, but in this research, we use some reduced boards such as 4 × 4, 4 × 6 and so on.

In this experiment, players take a next move from all possible choice at random each other. Firstly, we enumerate all the situation that a game was finished, and count the number of results according to the difference of points of Black and White. Next, we add all the values (obtained for each point difference) of the product of a point difference (= Black’s – White’s) and the number of situations to become its point difference. Finally, we divide the values obtained above by the total number of situations. We assume this acquired value to be a handicap for White.

And the location with a handicap stone assumes the circumference of a board, see Fig.1.

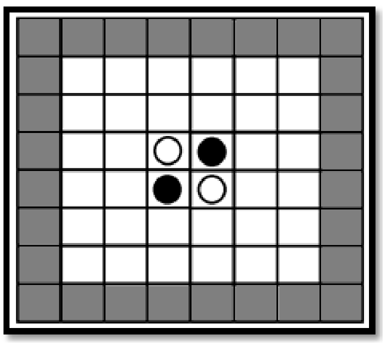


Figure.1: Location with a handicap stone

3. Results and Discussions

3.1 Board of 4 × 4

Fig.2 shows the number of situations of each point difference (Black's – White's) on 4 × 4 board.

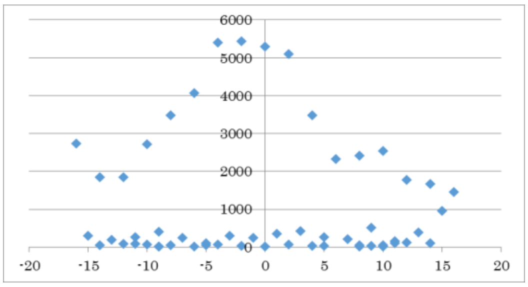


Figure.2: The number of situations of each point difference on 4 × 4 board

The advantage of Black is -0.642 points on average, which implies a first move disadvantage. Thus the handicap of White becomes -0.642. In fact, it has been found that white can certainly win on 4 × 4 board. Similarly, it has been found that White can win on 6 × 6 board [1]. Although the result of the 8 × 8 regular board has not been clear, many people believe that White can win. In this way, Othello game does not become clear whether the first move is advantageous.

For handicap stones, we display the location of the handicap stones as Fig.3.

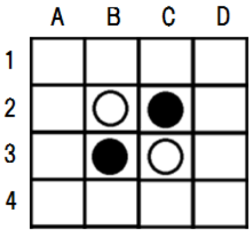


Figure.3: 4 × 4 board

The handicaps for White due to the color and location of a handicap stones are followings. (See Table.1.)

Table. 1: Handicap stones of 4 × 4 board

	Black stone	White stone
A1	+6.000	- 3.796
B1	- 0.689	+1.540
C1	- 2.322	+1.991
D1	+0.929	- 6.773

3.2 Board of 4 × 6

Fig.4 shows the number of situations of each point difference (Black's – White's) on 4 × 6 board.

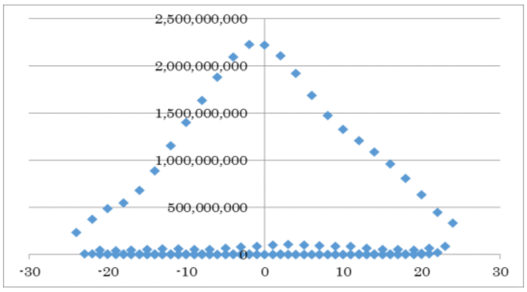


Figure.4: The number of situations of each point difference on 4 × 6 board

The handicap of White becomes +0.778 points, which implies a first move advantage.

According to an exhaustive search of 4 × 6 board, we found that Black can win. The fact that the value of the handicap for White of 4 × 6 board is positive is consistent with this result.

For handicap stones, we display the location of the handicap stones as Fig.5.

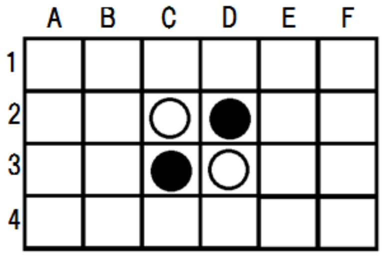


Figure.5: 4 × 6 board

The handicaps for White due to the color and location of a handicap stones are followings. (See Table.2.)

Table. 2: Handicap stones of 4 × 6 board

	Black stone	White stone
A1	+5.112	- 3.742
A2	- 0.119	+1.812
A3	+0.223	+1.934
A4	+5.080	- 2.878
B1	- 1.057	+2.267
C1	+1.740	- 0.381
D1	+2.118	+0.135
E1	- 0.564	+3.255

4. Conclusion

In this research, we tried to measure the handicap in Othello game quantitatively. By choosing a situation at random from all possible choice, we can estimate the handicaps for White (=the defensive hand) of 4 × 4 and

4 × 6 boards at -0.642 and +0.778, respectively. Sign of the handicap is decided whether the first move is advantageous. Unlike the game of Go that an advantage of the first move is confirmed empirically, the Othello game does not become clear whether the first move is advantageous. However, our results of handicap are consistent with results with an exhaustive search. This indicates that the evaluation method that we proposed is effective.

Moreover, since the defensive hand was believed to be advantageous in the game of Othello, it is a very interesting result that advantage and disadvantage of the first move is determined depending on the board. We have the following conjecture.

Conjecture: Suppose that m,n∈N={1,2,3,…}. For a given 2m×2n board in the game of Othello, we get the followings.
(1) Assume that m=1 or n=1, the game drow.
(2) Assume that m,n≧2 and n+m is odd number, the first move is advantageous.
(3) Assume that m,n≧2 and n+m is even number, the first move is disadvantageous.

For handicap stones, we found that the side that is placed in the corner is advantageous. However, we were not able to find the law except in the corner. To find the rules of general, experiments with a large board is required.

Reference

[1] Joel Feinstein (1993), Perfect Play in 6x6 Othello from two alternative starting positions, http://www.feinst.demon.co.uk/Othello/6x6sol.html

[2] Michael H. Albert, Richard J. Nowakowski and David Wolfe (2007), Lesson in Play: An Introduction to Combinatorial Game Theory

[3] Elwyn Berlekamp and David Wolf (1994), Mathematical Go: Chilling gets the Last Point

The International Conference on Artificial Life and Robotics (ICAROB 2014), Compal Hall, Oita, Japan, January 11-13, 2014

On Parameter Setting in Identifying the Same Languages Involved in Different Language Data

Ren Wu[1] and Hiroshi Matsuno[2]
[1]Yamaguchi Junior College, [2]Graduate School of Science and Engineering, Yamaguchi University
[1]1346-2, Oaza Daido, Hofu city, Yamaguchi prefecture, 747-1232, Japan
[2]1677-1, Yoshida, Yamaguchi city, Yamaguchi prefecture, 753-8511, Japan
Tel: [1]+81-835-32-0138, [2]+81-83-933-5697
Email: [1]wu@yamaguchi-jc.ac.jp, [2]matsuno@sci.yamaguchi-u.ac.jp

Abstract: We have proposed a method based on tree structure and string alignment technique for identifying the same languages involved in two language classification trees provided by different linguists. In this method, several kinds of similarity measure, such as language name similarity and language general similarity etc., and an algorithm are proposed. Several unknown parameters are used there and need to be set to constant value in order to calculate the similarities. This paper aims to determine all the values of these parameters and get the identification results in order to confirm the usefulness and effectiveness of our proposed similarity measures and the algorithm. As the result, we obtained reasonable good values for the parameters throughout the experiments.

Keywords: world languages tree, language name similarity, language classification similarity

I. INTRODUCTION

We have proposed a method based on tree structure and string alignment technique for identifying the same languages involved in two **world languages trees**, denoted by WLTs, which are two language classification trees provided by different linguists. We have named these two WLTs as T_Y and T_S.

In our previous work, we have quantified several kinds of similarity measure, such as **language name similarity**, **language classification similarity** and **language general similarity** and so forth. Language name similarity is defined as string similarity between two language names. On the other hand, language classification similarity is defined as a weighted average of three kinds of similarities (family name similarity, parent name similarity and brother name similarity) that are based on language name similarity. Furthermore, language general similarity is defined as a weighted average of language name similarity and language classification similarity. At the same time, we have developed an algorithm for finding out same language pairs involved in T_Y and T_S by applying above defined similarity measures. In this algorithm, several unknown parameters, (e, f, g), (a, b), Δ and threshold ρ, are used and need to be set to constant values firstly.

This paper aims to determine all the values of these parameters and then get the identification results in order to confirm that our proposed similarity measures and the algorithm are useful and effective.

This paper is organized as follows. First, we give the definitions of our proposed similarity measures and the algorithm for finding same language pairs. Then, we give the process for setting the parameters (e, f, g), (a, b), Δ and threshold ρ. Finally, we give the experimental results by applying the parameter values that are set.

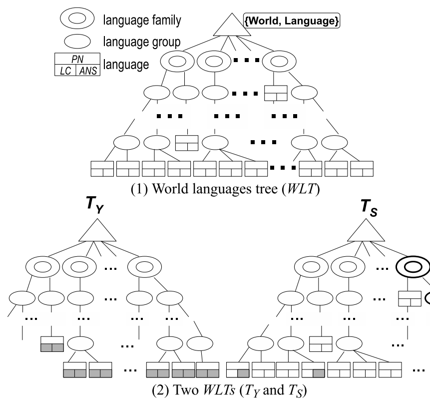


Fig. 1. Two WLTs (T_Y and T_S)

II. PRELIMINARY

A. Similarity Measure

Fig. 1(1) shows an image of WLT and Fig. 1(2) shows two WLTs called T_Y and T_S. A language y included in T_Y is denoted by $y \in V_{leaf}(T_Y)$. Language y and s are assumed to be the same language included in T_Y and T_S respectively.

Fig. 2 shows an example of the same language pair (y, s) involved in T_Y and T_S. The language names of y and s are "ARABIC, SUDANESE CREOLE" and "Arabic, Sudanese Creole" respectively and the same. Here, all alphabetical notations are not case-sensitive.

Finding out such same language pairs like (y, s) automatically is our purpose. In order to solve this problem, measure of language name similarity has been introduced and defined. In the following definition, a term **WORD** is used, which is a string consisting of only alphabets, such

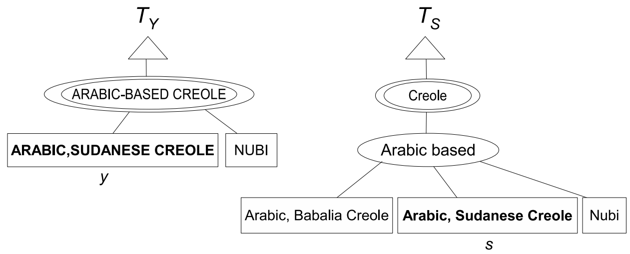


Fig. 2. The same language that needs to be identified

as "Arabic". Language name is a list of WORD(s), just like language names "ARABIC, SUDANESE CREOLE" and "Arabic, Sudanese Creole". If two WORDs are individually included in two language names and are most similar, for example "ARABIC" and "Arabic", they are called **WORD pair**.

[Definition 1] Let v and w be two WORDs. WORD Similarity between v and w, denoted by $sd\_w(v,w)$, is defined by

$$sd\_w(v,w) = \frac{l_A(v,w) - ed(v,w)}{l_A(v,w)}, \tag{1}$$

where $ed(v,w)$ and $l_A(v,w)$ respectively represent edit distance and length of optimal alignment between v and w, under the condition that the operations of insertions, deletions and substitutions all cost 1. □

[Definition 2] Let $\mathcal{L}_1$=$\{v_1, v_2, \cdots, v_m\}$ and $\mathcal{L}_2$ = $\{w_1, w_2, \cdots, w_n\}(m \geq n)$ be Language Names, (v_i, w'_i) be WORD pair. Note that for any $v_i \in \mathcal{L}_1$, if $\mathcal{L}_2$ does not contain w'_i that corresponds to v_i, then let w'_i=Null. Language Name Similarity, denoted by $sd\_ln(\mathcal{L}_1, \mathcal{L}_2)$, between $\mathcal{L}_1$ and $\mathcal{L}_2$ is defined by

$$sd\_ln(\mathcal{L}_1, \mathcal{L}_2) = \frac{\sum_{i=1}^{m} sd\_w(v_i, w'_i)}{m} \tag{2}$$

□

[Definition 3] Let $y \in V_{leaf}(T_Y)$ and $s \in V_{leaf}(T_S)$ be languages and $\mathcal{L}_y^{T_Y}$ and $\mathcal{L}_s^{T_S}$ be the primary language names of y and s. Let m be the number of alternate names with s, furthermore let all the alternate names be $\mathcal{A}_1^s, \mathcal{A}_2^s, \cdots, \mathcal{A}_m^s$ if $m > 0$. Then **language name similarity** of y and s, denoted by $sd\_ln_{node}(y,s)$, is defined by

$$sd\_ln_{node}(y,s) = \begin{cases} sd\_ln(\mathcal{L}_y^{T_Y}, \mathcal{L}_s^{T_S}) & (m{=}0) \\ \max\{sd\_ln(\mathcal{L}_y^{T_Y}, \mathcal{L}_s^{T_S}), sd\_ln(\mathcal{L}_y^{T_Y}, \mathcal{A}_1^s), \\ \quad sd\_ln(\mathcal{L}_y^{T_Y}, \mathcal{A}_2^s), \cdots, sd\_ln(\mathcal{L}_y^{T_Y}, \mathcal{A}_m^s)\} & (m{>}0) \end{cases} \tag{3}$$

□

Based on language name similarity, family name similarity denoted by $sd\_fn(y,s)$, parent name similarity denoted by $sd\_pn(y,s)$ and brother name similarity denoted by $sd\_bn(y,s)$ are defined as follows. These three similarities are designed for language classification similarity, which is intended to be a different similarity measure, from the aspect of language classification.

[Definition 4] Let $\mathcal{FN}_y$ and $\mathcal{FN}_s$ be the family names of languages $y \in V_{leaf}(T_Y)$ and $s \in V_{leaf}(T_S)$, respectively. Then **family name similarity** between y and s, denoted by $sd\_fn(y,s)$, is defined by

$$sd\_fn(y,s) = sd\_ln(\mathcal{FN}_y, \mathcal{FN}_s) \tag{4}$$

□

[Definition 5] Let $\mathcal{PN}_y$ and $\mathcal{PN}_s$ be the parent's names of languages $y \in V_{leaf}(T_Y)$ and $s \in V_{leaf}(T_S)$, respectively. Then **parent name similarity** between y and s, denoted by $sd\_pn(y,s)$, is defined by

$$sd\_pn(y,s) = sd\_ln(\mathcal{PN}_y, \mathcal{PN}_s) \tag{5}$$

□

Languages "NUBI" and "Nubi" shown in Fig. 2 are brothers of "ARABIC, SUDANESE CREOLE" and "Arabic, Sudanese Creole", respectively. A brother combination from two sets of brothers, such as ("NUBI", "Nubi"), is called **brother language pair**.

[Definition 6] Let x_1 (x_2) be a language included in T_Y or T_S (T_S or T_Y), BL_{x_1}=$\{bl_1^{x_1}, bl_2^{x_1}, \cdots\}$ and BL_{x_2}=$\{bl_1^{x_2}, bl_2^{x_2}, \cdots\}$ ($|BL_{x_1}| \geq |BL_{x_2}| > 0$) be respectively the sets of brother languages of x_1 and x_2, and further $BP(x_1, x_2)$ be the set of brother language pair. Then **brother name similarity** between x_1 and x_2, denoted by $sd\_bn(x_1, x_2)$, is defined by

$$sd\_bn(x_1,x_2) = \begin{cases} sd\_ln_{node}(x_1, x_2) & (m{=}n{=}0) \\ 0 & (m{>}0, n{=}0) \\ \frac{\sum_{(\mu,\nu) \in BP(x_1,x_2)} sd\_ln_{node}(\mu,\nu)}{\frac{m+n}{2}} & (m \geq n > 0) \end{cases} \tag{6}$$

□

Then language classification similarity, and further language general similarity that integrates language name similarity and language classification similarity, are defined as follows.

[Definition 7] language classification similarity between $y \in V_{leaf}(T_Y)$ and $s \in V_{leaf}(T_S)$, denoted by $sd\_lc(y,s)$, is defined by

$$sd\_lc(y,s) = e{*}sd\_fn(y,s) + f{*}sd\_pn(y,s) + g{*}sd\_bn(y,s), \tag{7}$$

where e, f, g satisfies $1 \geq e \geq 0$, $1 \geq f \geq 0$, $1 \geq g \geq 0$ and $e{+}f{+}g{=}1$. □

[Definition 8] language general similarity between $y \in V_{leaf}(T_Y)$ and $s \in V_{leaf}(T_S)$, denoted by $sd\_gen(y,s)$, is defined by

$$sd\_gen(y,s) = a * sd\_ln(y,s) + b * sd\_lc(y,s), \tag{8}$$

where a, b satisfies $1 \geq a \geq 0$, $1 \geq b \geq 0$ and $a + b{=}1$. □

B. Finding same language pairs

Same language pairs such as (y,s) shown in Fig. 2 can be found out by applying language name similarity and language general similarity as follows.

Two dummy WLTs with 3 languages in T_Y and 4 in T_Y are used to describe how to find them. (1) First, generate a

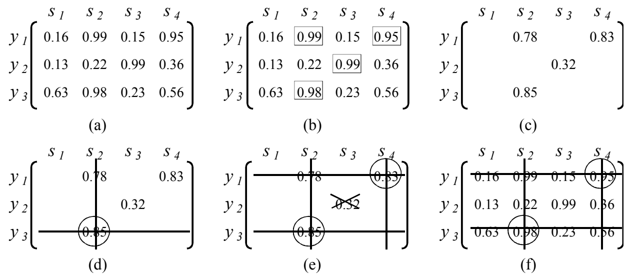


Fig. 3. Process of finding same language pairs

			Fact: Are two languages the same? Yes	Fact: Are two languages the same? No
Result	Are two languages outputted as the same language pair?	Yes	① TP (True Positive)	② FP (False Positive)
		No	③ FN (False Negative)	④ TN (True Negative)

Fig. 5. True-false judgment for outputted same language pairs

Algorithm: FSLV

Input : $V_{leaf}(T_Y)$, $V_{leaf}(T_S)$, e, f, g, a, b, Δ, ρ
Output : Same language pairs SLP for all $y{\in}V_{leaf}(T_Y)$
Procedure :

1° Set $SLP{\leftarrow}\phi$, $P{\leftarrow}\phi$, $\gamma{\leftarrow}1$. For all the pairs of $y{\in}V_{leaf}(T_Y)$ and $s{\in}V_{leaf}(T_S)$, do $\psi_{y,s}{=}sd\_ln_{node}(y, s)$ and generate language name similarity matrix $\Psi{=}(\psi_{y,s})$.

2° If $\gamma{\le}0$ goto 3°; otherwise do $P{\leftarrow}P{\cup}\{(y,s) \mid \gamma{\ge}\psi_{y,s} {\ge}\gamma{-}\Delta, \psi_{y,s}$: elements of $\Psi\}$ and do the following operations (i) $\sim$ (v):

- **(i)** For all $(y,s){\in}P$ do $\omega_{y,s}{=}sd\_gen(y,s)$ by using the coefficients e, f, g, a, b.
- **(ii)** Do $\omega_{max} = \max\{\omega_{y,s} \mid (y,s){\in}P\}$. If $\omega_{max} < \rho$, set $\gamma{\leftarrow}\gamma{-}\Delta$ and goto 2°.
- **(iii)** Select one pair $(y', s'){\in}P$ with $\omega_{y',s'} = \omega_{max}$ and do $SLP{\leftarrow}SLP{\cup}(y', s')$.
- **(iv)** Do $P{\leftarrow}P{-}\{(y', s')\}$ and delete y''s row and s''s column from Ψ.
- **(v)** If $P{\ne}\phi$ goto (ii); otherwise do $\gamma{\leftarrow}\gamma{-}\Delta$ and goto 2°.

3° Output SLP and stop.

Fig. 4. Algorithm: FSLV

matrix Ψ with $\psi_{y_i,s_j}{=}sd\_ln_{node}(y_i, s_j)$ as its element for $y_i{\in}V_{leaf}(T_Y)$ and $s_j{\in}V_{leaf}(T_S)$ as shown in Fig. 3 (a). (2) Then select all the combinations of y_i and s_j satisfying $\gamma{\ge}sd\_ln_{node}(y_i, s_j){>}\gamma{-}\Delta$. The selected combinations are boxed as shown in Fig. 3 (b). Here, the initial value of γ and Δ are set to 1 and 0.1, respectively. (3) Calculate the values of $sd\_gen(y_i, s_j)$ for the selected combinations as shown in Fig. 3 (c). Values of coefficients e, f, g and a, b will be determined by experiments. (4) Select the pair (y_3, s_2) with the highest language general similarity value from all the values of $sd\_gen(y_i, s_j)$. If we set ρ=0.5, this language pair (y_3, s_2) will be passed over and identified as the same language. Then the language general similarity values as shown Fig. 3 (d) are deleted. Similarly, (y_1, s_4) is identified as well as (y_3, s_2) as shown in Fig. 3 (e). (5) Delete the values related to the identified same language pairs as shown in Fig. 3 (f). (6) Update γ as $\gamma = \gamma{-}\Delta$ and repeat this process from (2) $\sim$ (5) till $\gamma{\le}0$. The algorithm of this processing is shown in Fig. 4.

III. Parameter setting

In algorithm FSLV shown in Fig. 4, parameters (e, f, g), (a, b), Δ and threshold ρ are used as input arguments, and are unknown. What should be done first is setting their values. In this section, we are going to give a solution for parameter setting.

A. Evaluation method and test data

Parameters (e, f, g) are used to calculate language classification similarity which is defined as a weighted average of family name similarity, parent name similarity and brother name similarity, and need to be set at first. These three parameters are individually the coefficients of the three similarities. Take (y, s) shown in Fig. 2 as the example, we consider that, if there are same language pairs such as ("NUBI", "Nubi") within the brothers, then it is most probably that y and s may also be the same language. Hence, we lay weight on g that is the coefficient of brother language similarity, and consider that $e{=}0.25$, $f{=}0.25$, $g{=}0.5$ should be appropriate.

On the other hand, parameters (a, b), Δ and threshold ρ will be determined throughout experiments.

We use test data of 200 languages which were selected out of all 2,869 languages included in T_Y in a random manner. In advance of setting parameters, we investigate the corresponding languages in T_S for all the languages of this test data. However, it does not mean that the corresponding same languages can always be found in T_S for all the languages.

For the test data, we repeat experiments according to algorithm FSLV shown in Fig. 4 by changing the values of parameters (a, b), Δ and threshold ρ, and then check the results with the investigated facts to get the values of TP, FP, TN and FN as shown in Fig. 5. Here, $TP + FP + TN + FN = 200$ should be satisfied.

We use evaluation measures $F{-}measure$, $Recall$, $Precision$ [5] to choose the best (a, b), Δ and threshold ρ. These measures are defined as follows.

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

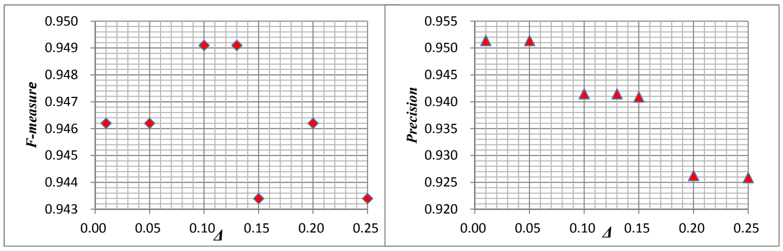
$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$F{-}measure = \frac{2{\cdot}Recall{\cdot}Precision}{Recall + Precision} \tag{11}$$

We consider that the values of parameters (a, b), Δ and threshold ρ, with the highest values of $F{-}measure$ and $Precision$, should be the best ones.

B. Process of parameter setting

Here, we determine the values of parameters (a, b), Δ and threshold ρ. First of all, we deal with parameters (a, b). Repeat executing algorithm FSLV with parameters $\Delta{=}0.01$ and $\rho{=}0.5$ by changing the values of parameters



(e=0.25, f=0.25, g=0.5, a=0.5, b=0.5, ρ=0.5)

(1) $F\!-\!measure$ (2) $Precision$

Fig. 6. Ups and downs of the evaluation values

(a, b) by steps of 0.1 as follows: $(1.0, 0.0)$, $(0.9, 0.1)$, $\cdots$, $(0.1, 0.9)$, $(0.0, 1.0)$. Then do the same thing as well by increasing the values of Δ from 0.01 to 0.25 by steps of 0.01 such as Δ=$0.02, 0.03, \cdots, 0.25$. Here, threshold ρ is set as ρ=0.5.

After tallying the values of TP, FP, TN and FN for every execution of algorithm FSLV, compute the values of $Recall$, $Precision$ and $F\!-\!measure$ and use them to estimate the trend of parameters (a, b). We get to notice that the values of $F\!-\!measure$ and $Precision$ have the best values when $a = 0.5$ and $b = 0.5$. So we set the values of parameters (a, b) to $(0.5, 0.5)$.

What need to be set continuously is the value of parameter Δ. As described above, we repeat executing algorithm FSLV by increasing the values of Δ from 0.01 to 0.25 by steps of 0.01. The values of $F\!-\!measure$ and $Precision$ have ups and downs due to the change of Δ. We show a part of values of $F\!-\!measure$ and $Precision$ in Fig. 6(1) and Fig. 6(2) respectively by setting Δ as $0.01, 0.05, 0.10, 0.15, 0.20, 0.25, 0.13$ under the condition of a=0.5, b=0.5, ρ=0.5. Fig. 6(1) indicates that $F\!-\!measure$ has a peak at Δ=0.10 and 0.13. At the two points, Fig. 6(2) has the same highest values. We cannot differentiate the effect between the cases when Δ=0.10 and Δ=0.13, So we set Δ=0.10 or Δ=0.13.

Finally, we set the value of threshold ρ. Under the condition of Δ=0.10, a=0.5, b=0.5 or Δ=0.13, a=0.5, b=0.5, we repeat execution of algorithm FSLV by increasing the value of ρ from 0.50 to 0.80 by steps of 0.05, then we found that $F\!-\!measure$ has a peak at ρ=0.55. So we repeat the execution additionally by setting ρ as 0.51, 0.52, 0.53 or 0.54 individually. $F\!-\!measure$ shows the same value at ρ=0.54 or ρ=0.55 as well as $Precision$. Under the condition of the same $F\!-\!measure$ and $Precision$, the threshold with lower value is better. Therefore, we set ρ=0.54.

Up to here, we have gotten the best values of parameters (a, b), Δ and threshold ρ under the condition of e=0.25, f=0.25, g=0.5 by using the test data. As the result, Δ=0.10, a=0.5, b=0.5, ρ=0.54 or Δ=0.13, a=0.5, b=0.5, ρ=0.54 are obtained.

TABLE 1. Experimental results

TN (total number of languages)	number of languages found out	*TP*	*FP*	*TP/TN*	*TP/(TP+FP)*
2,869	2,687	2,648	39	**92%**	98%

e=0.25, f=0.25, g=0.5, Δ=0.13, a=0.5, b=0.5, ρ =0.54

IV. Experimental Results

We have done experiment by applying the obtained parameters to original data T_Y and T_S to confirm the accuracy of parameter setting and the usefulness and effectiveness of our proposed method. By setting the parameters to (i) Δ=0.10, a=0.5, b=0.5, ρ=0.54 and (ii) Δ=0.13, a=0.5, b=0.5, ρ=0.54 (both under e=0.25, f=0.25, g=0.5), we have executed algorithm FSLV for T_Y and T_S and gotten the results shown in TABLE 1: in all the $2,869$ languages of T_Y, we have gotten $2,687$ same language pairs as the output of algorithm FSLV, in which $2,648$ languages (92%) are the true cases.

V. Concluding Remarks

We have given a way of parameter setting for our previously proposed method. Using the parameter values, about 92% languages of T_Y have been identified, and the precision of identification is up to 98%. These results imply that our parameter setting described above is reasonably good. And at the same time, we have confirmed that our proposed method previously is useful and effective. It should be noticed that the values of parameters (a, b), Δ and threshold ρ are set under the condition of e=0.25, f=0.25, g=0.5. As the future work, we need to do further tests to confirm the parameter setting.

VI. Acknowledgement

This work was supported by JSPS KAKENHI Grant-in-Aid for Challenging Explorotory Research Numbers 23650129.

References

[1] R. Wu and H. Matsuno (2010), Identifying Same Languages by Considering Similarities of Language Name and Language Classification, Proceedings of The 2010 International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2010), pp.516-519

[2] R. Wu, H. Inui and H. Matsuno (2012), New Measurement of Similarity of Language Classification, Proceedings of The 2012 International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2012), in CD-ROM(F-M2-05)

[3] http://www.ethnologue.com/web.asp

[4] G. Navarro (2001), A guided tour to approximate string matching, ACM Computing Surveys (CSUR), Vol.33, No.1, pp.31-88

[5] V. Vapnik (1995), *The nature of statistical learning theory*, Springer, New York