

ISSN : 2165-4069(Online)

ISSN : 2165-4050(Print)



IJARAI

International Journal of
Advanced Research in Artificial Intelligence

Volume 2 Issue 5

www.ijarai.thesai.org

A Publication of
The Science and Information Organization



INTERNATIONAL JOURNAL OF
ADVANCED RESEARCH IN ARTIFICIAL INTELLIGENCE



THE SCIENCE AND INFORMATION ORGANIZATION
www.thesai.org | info@thesai.org



Editorial Preface

From the Desk of Managing Editor...

"The question of whether computers can think is like the question of whether submarines can swim." — Edsger W. Dijkstra, the quote explains the power of Artificial Intelligence in computers with the changing landscape. The renaissance stimulated by the field of Artificial Intelligence is generating multiple formats and channels of creativity and innovation.

This journal is a special track on Artificial Intelligence by The Science and Information Organization and aims to be a leading forum for engineers, researchers and practitioners throughout the world.

The journal reports results achieved; proposals for new ways of looking at AI problems and include demonstrations of effectiveness. Papers describing existing technologies or algorithms integrating multiple systems are welcomed. IJARAI also invites papers on real life applications, which should describe the current scenarios, proposed solution, emphasize its novelty, and present an in-depth evaluation of the AI techniques being exploited. IJARAI focusses on quality and relevance in its publications.

In addition, IJARAI recognizes the importance of international influences on Artificial Intelligence and seeks international input in all aspects of the journal, including content, authorship of papers, readership, paper reviewers, and Editorial Board membership.

The success of authors and the journal is interdependent. While the Journal is in its initial phase, it is not only the Editor whose work is crucial to producing the journal. The editorial board members, the peer reviewers, scholars around the world who assess submissions, students, and institutions who generously give their expertise in factors small and large— their constant encouragement has helped a lot in the progress of the journal and shall help in future to earn credibility amongst all the reader members.

I add a personal thanks to the whole team that has catalysed so much, and I wish everyone who has been connected with the Journal the very best for the future.

Thank you for Sharing Wisdom!

Managing Editor
IJARAI
Volume 2 Issue 5 May 2013
ISSN: 2165-4069(Online)
ISSN: 2165-4050(Print)
©2013 The Science and Information (SAI) Organization

Editorial Board

Peter Sapaty - Editor-in-Chief

National Academy of Sciences of Ukraine

Domains of Research: Artificial Intelligence

Alaa F. Sheta

Electronics Research Institute (ERI)

Domain of Research: Evolutionary Computation, System Identification, Automation and Control, Artificial Neural Networks, Fuzzy Logic, Image Processing, Software Reliability, Software Cost Estimation, Swarm Intelligence, Robotics

Antonio Dourado

University of Coimbra

Domain of Research: Computational Intelligence, Signal Processing, data mining for medical and industrial applications, and intelligent control.

David M W Powers

Flinders University

Domain of Research: Language Learning, Cognitive Science and Evolutionary Robotics, Unsupervised Learning, Evaluation, Human Factors, Natural Language Learning, Computational Psycholinguistics, Cognitive Neuroscience, Brain Computer Interface, Sensor Fusion, Model Fusion, Ensembles and Stacking, Self-organization of Ontologies, Sensory-Motor Perception and Reactivity, Feature Selection, Dimension Reduction, Information Retrieval, Information Visualization, Embodied Conversational Agents

Liming Luke Chen

University of Ulster

Domain of Research: Semantic and knowledge technologies, Artificial Intelligence

T. V. Prasad

Lingaya's University

Domain of Research: Bioinformatics, Natural Language Processing, Image Processing, Robotics, Knowledge Representation

Wichian Sittiprapaporn

Maharakham University

Domain of Research: Cognitive Neuroscience; Cognitive Science

Yaxin Bi

University of Ulster

Domains of Research: Ensemble Learning/Machine Learning, Multiple Classification Systems, Evidence Theory, Text Analytics and Sentiment Analysis

Reviewer Board Members

- **Alaa Sheta**
Electronics Research Institute (ERI)
- **Albert Alexander**
Kongu Engineering College
- **Amir HAJJAM EL HASSANI**
Université de Technologie de Belfort-Monbéliard
- **Amit Verma**
Department in Rayat & Bahra Engineering College, Mo
- **Antonio Dourado**
University of Coimbra
- **B R SARATH KUMAR**
Lenora College of Engineering
- **Babatunde Opeoluwa Akinkunmi**
University of Ibadan
- **Bestoun S.Ahmed**
Universiti Sains Malaysia
- **Chien-Peng Ho**
Information and Communications Research Laboratories, Industrial Technology Research Institute of Taiwan
- **David M W Powers**
Flinders University
- **Dimitris Chrysostomou**
Democritus University
- **Dr.Dhananjay R Kalbande**
Mumbai University
- **Francesco Perrotta**
University of Macerata
- **Frank Ibikunle**
Covenant University
- **Grigoras Gheorghe**
"Gheorghe Asachi" Technical University of Iasi, Romania
- **Guandong Xu**
Victoria University
- **Haibo Yu**
Shanghai Jiao Tong University
- **Jatinderkumar R. Saini**
S.P.College of Engineering, Gujarat
- **Krishna Prasad Miyapuram**
University of Trento
- **Luke Liming Chen**
University of Ulster
- **Marek Reformat**
University of Alberta
- **Md. Zia Ur Rahman**
Narasaraopeta Engg. College, Narasaraopeta
- **Mokhtar Beldjehem**
University of Ottawa
- **Monji Kherallah**
University of Sfax
- **Nitin S. Choubey**
Mukesh Patel School of Technology Management & Eng
- **Rajesh Kumar**
National University of Singapore
- **Rajesh K Shukla**
Sagar Institute of Research & Technology-Excellence, Bhopal MP
- **Rongrong Ji**
Columbia University
- **Said Ghoniemy**
Taif University
- **Samarjeet Borah**
Dept. of CSE, Sikkim Manipal University
- **Sana'a Wafa Tawfeek Al-Sayegh**
University College of Applied Sciences
- **Saurabh Pal**
VBS Purvanchal University, Jaunpur
- **Shahaboddin Shamshirband**
University of Malaya
- **Shaidah Jusoh**
Zarqa University
- **Shrinivas Deshpande**
- **SUKUMAR SENTHILKUMAR**
Universiti Sains Malaysia
- **T C.Manjunath**
HKBK College of Engg
- **T V Narayana Rao**
Hyderabad Institute of Technology and Management
- **T. V. Prasad**

Lingaya's University

- **Vitus S.W Lam**
Domains of Research
- **VUDA Sreenivasarao**
St. Mary's College of Engineering &
Technology
- **Vishal Goyal**
- **Wei Zhong**
University of south Carolina Upstate
- **Wichian Sittiprapaporn**
Mahasarakham University
- **Yaxin Bi**

University of Ulster

- **Yuval Cohen**
The Open University of Israel
- **Zhao Zhang**
Deptment of EE, City University of Hong
Kong
- **Zne-Jung Lee**
Dept. of Information management, Huafan
University
- **Zhigang Yin**
Institute of Linguistics, Chinese Academy of
Social Sciences

CONTENTS

Paper 1: Rice Crop Quality Evaluation Method through Regressive Analysis between Nitrogen Content and Near Infrared Reflectance of Rice Leaves Measured from Near Field

Authors: Kohei Arai

PAGE 1 – 6

Paper 2: Sea Ice Concentration Estimation Method with Satellite Based Visible to Near Infrared Radiometer Data Based on Category Decomposition

Authors: Kohei Arai

PAGE 7 – 13

Paper 3: Sensitivity Analysis on Sea Surface Temperature Estimation Methods with Thermal Infrared Radiometer Data through Simulations

Authors: Kohei Arai

PAGE 14 – 21

Paper 4: Image Clustering Method Based on Self Organization Mapping: SOM Derived Density Maps and Its Application for Landsat Thematic Mapper Image

Authors: Kohei Arai

PAGE 22 – 31

Paper 5: Remote Sensing Satellite Image Database System Allowing Image Portion Retrievals Utilizing Principal Component Which Consists Spectral and Spatial Features Extracted from Imagery Data

Authors: Kohei Arai

PAGE 32 – 38

Paper 6: AMashup Based Content Search Engine for Mobile Devices

Authors: Kohei Arai

PAGE 39 – 43

Paper 7: The preliminary results of a force feedback control for Sensorized Medical Robotics

Authors: Duck Hee Lee, Seung Joon Song, Reza Fazel-Rezai, Jaesoon Choi

PAGE 44 – 48

Paper 8: Performance Analysis and Evaluation of Different Data Mining Algorithms used for Cancer Classification

Authors: Gopala Krishna Murthy Nookala, Bharath Kumar Pottumuthu, Nagaraju Orsu, Suresh B. Mudunuri

PAGE 49 – 55

Paper 9: Path Based Mapping Technique for Robots

Authors: Amiraj Dhawan, Parag Oak, Rahul Mishra, George Puthanpurackal

PAGE 56 – 62

Paper10: A New Optimization Algorithm For Combinatorial Problems

Authors: Azmi Alazzam, Harold W. Lewis III

PAGE 63 – 68

Paper 11: Gesture Controlled Robot using Image Processing

Authors: Harish Kumar Kaura, Vipul Honrao, Sayali Patil, Pravish Shetty

PAGE 69 – 77

Rice Crop Quality Evaluation Method through Regressive Analysis between Nitrogen Content and Near Infrared Reflectance of Rice Leaves Measured from Near Field

Kohei Arai¹

Graduate School of Science and Engineering
Saga University
Saga City, Japan

Abstract—Rice crop quality evaluation method through regressive analysis between nitrogen content in the rice leaves and near infrared reflectance measurement data from near field, from radio wave controlled helicopter is proposed. Rice quality dependency on nitrogen of chemical fertilizer and water supply condition is evaluated. Also homogeneity of the rice crop quality in the paddy fields is evaluated. Furthermore, influence due to shadow on near infrared reflectance of rice leaves measured from near field is taken into account in the rice crop quality evaluation processes.

Keywords—rice crop; regressive analysis; nitrogen content; reflectance measurement;

I. INTRODUCTION

Vitality monitoring of vegetation is attempted with photographic cameras [1]. Grow rate monitoring is also attempted with spectral reflectance measurements [2]. Bi-Directional Reflectance Distribution Function: BRDF is related to the grow rate for tealeaves [3]. Using such relation, sensor network system with visible and near infrared cameras is proposed [4]. It is applicable to estimate nitrogen content and fiber content in the tealeaves in concern [5]. Therefore, damage grade can be estimated with the proposed system for rice paddy fields [6]. This method is validated with Monte Carlo simulation [7]. Also Fractal model is applied to representation of shapes of tealeaves [8]. Thus the tealeaves can be assessed with parameters of the fractal model. Vitality of tea trees are assessed with visible and near infrared camera data [9].

One of the methods for monitoring the fields is to use remote sensing technology utilizing aircrafts, helicopters, hot air balloons, etc. with a wide field of view for monitoring relatively large scaled agricultural fields. In particular, there are remote sensing sensors which onboard radio controlled helicopters. Attitude stability of the radio controlled helicopters is getting well now a day. Field of view of the remote sensing sensors is good enough for relatively wide scale of agricultural fields.

One of the indexes which allows indicate quality of agricultural crops is nitrogen content in the agricultural crop leaves. The nitrogen content is proportional to the reflectance at

Near Infrared: NIR wavelength regions. Therefore, it is possible to estimate quality of agricultural crops with radio controlled helicopter based near infrared camera data.

Through experiments at the Saga Prefectural Agricultural Research Institute: SPARI for the period of rice crop growing, it is found that the proposed system works well for monitoring quality of the rice crops. Also it is found that the proposed method for nitrogen content estimation with near infrared camera data. Furthermore, it is capable to check rice crop quality distribution in the rice crop fields in concern. Then quality control which depends on location by location of the rice crop fields can be made with the quality monitoring results.

The following section describes the proposed monitoring system and nitrogen content estimation method based on the relation between nitrogen content in the rice crops and near infrared camera data followed by some experiments. Then conclusion is described together with some discussions.

II. PROPOSED METHOD

A. Regressive Analysis

Linear regressive equation is expressed in equation (1).

$$N = aR + b \quad (1)$$

where N , R denotes measured Nitrogen content in leaves, and measured Near Infrared: NIR reflectance, respectively while a and b denotes regressive coefficients. There is well known relation between nitrogen content and NIR reflectance. Therefore, regressive analysis based on equation (1) is appropriate.

B. Radio Controlled Helicopter Based Near Infrared Cameras Utilizing Agricultural Field Monitoring System

The helicopter used for the proposed system is "GrassHOPPER"¹ manufactured by Information & Science Techno-Systems Co. Ltd. The major specification of the radio controlled helicopter used is shown in Table 1. Also, outlook of the helicopter is shown in Figure 1. Canon Powershot S100²

¹ http://www.ists.co.jp/?page_id=892

² <http://cweb.canon.jp/camera/dcam/lineup/powershot/s110/index.html>

(focal length=24mm) is mounted on the GrassHOPPER. It allows acquire images with the following Instantaneous Field of View: IFOV at the certain altitudes, 1.1cm (Altitude=30m) 3.3cm (Altitude=100m) and 5.5cm (Altitude=150m) .



Fig. 1. Outlook of the GrassHOPPER

TABLE I. MAJOR SPECIFICATION OF GRASSHOPPER

Weight	2kg (Helicopter only)
Size	80cm × 80cm × 30m
Payload	600g

In order to measure NIR reflectance, standard plaque whose reflectance is known is required. Spectralon³ provided by Labsphere Co. Ltd. is well known as well qualified standard plaque. It is not so cheap that photo print papers are used for the proposed system. Therefore, comparative study is needed between Spectralon and the photo print papers.

The proposed system consist Helicopter, NIR camera, photo print paper. Namely, photo print paper is put on the agricultural plantations, tea trees in this case. Then farm areas are observed with helicopter mounted NIR camera. Nitrogen content in agricultural plants, rice crops in this case, is estimated with NIR reflectance.

C. Proposed Method for Rice Crop Quality Evaluation

Rice crop quality can be represented nitrogen content which is closely related to nitrogen content. Furthermore, it is well known that nitrogen content can be represented with NIR reflectance. Therefore, rice crops quality can be evaluated with measured NIR reflectance based on the equation (1).

The proposed method and tea farm area monitoring system with helicopter mounted NIR camera is based on the aforementioned scientific background.

D. Rice Crop Field at Saga Prefectura; Agricultural Research Institute: SPARI

Specie of the rice crop is Hiyokumochi⁴ which is one of the late growing types of rice species. Hiyokumochi is one of low amylase (and amylopectin rich) of rice species (Rice No.216).

Figure 2 and 3 shows layout of the test site of rice crop field at SPARI⁵ which is situated at 33°13'11.5" North, 130°18'39.6"East, and the elevation of 52feet.

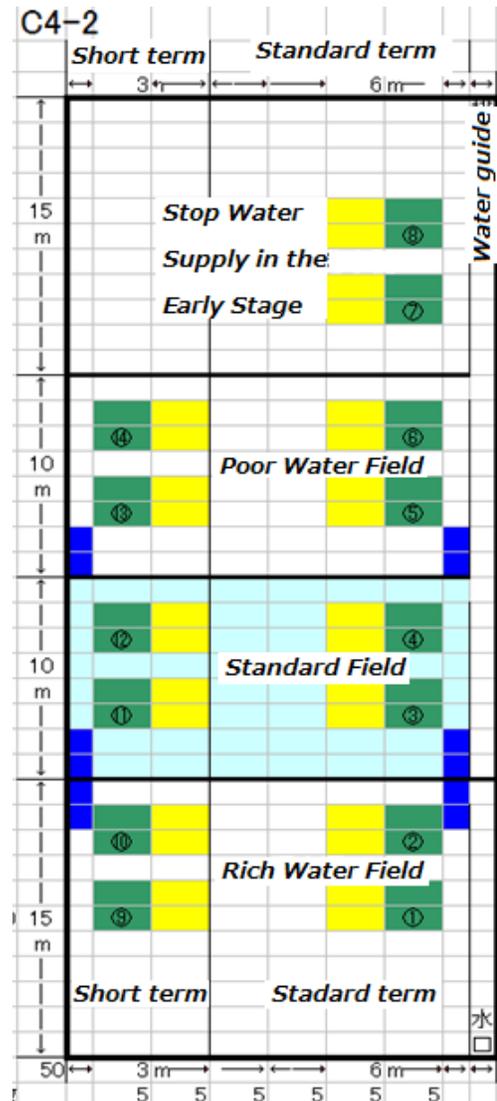


Fig. 2. Paddy field layout for investigation of water supply condition dependency on rice crop quality

The paddy field C4-2 is for the investigation of water supply condition on rice crop quality. There are 14 of the paddy field subsections of which water supply conditions are different each other.

There are two types of water supply scheduling, short term and standard term. Water supply is stopped in the early stage of

³ <https://www.google.co.jp/search?q=spectral+labsphere&hl=ja>

⁴ <http://ja.wikipedia.org/wiki/%E3%82%82%E3%81%A1%E7%B1%B3>
⁵ http://www.pref.saga.lg.jp/web/shigoto/_1075/_32933/ns-nouisetu/nouse/n_seika_h23.html

rice crop growing period for the short term water supply subsection fields while water supply is continued comparatively longer time period comparing to the short term water supply subsection fields. Meanwhile, there are three types of water supply conditions, rich, standard, and poor water supply subsection fields.

On the other hand, test sites C4-3 and C4-4 are for investigation of nitrogen of chemical fertilizer dependency on rice crop quality. There are two types of paddy subsections, densely and sparsely planted paddy fields. Hiyokumochi rice leaves are planted 15 to 20 fluxes per m² on June 22 2012. Rice crop fields are divided into 10 different small fields depending on the amount of nutrition including nitrogen ranges from zero to 19 kg/10 a/nitrogen.

Nitrogen of chemical fertilizer is used to put into paddy fields for five times during from June to August. Although rice crops in the 10 different small fields are same species, the way for giving chemical fertilizer are different. Namely, the small field No.1 is defined as there is no chemical fertilizer at all for the field while 9, 11, and 13 kg/ 10 a/ nitrogen of after chemical fertilizer are given for No.2 to 4, respectively, no initial chemical fertilizer though. Meanwhile, 9, 11, 13 kg/10 a/nitrogen are given as after chemical fertilizer for the small field No.5, 6, and 7, respectively in addition to the 3 kg/10 a/nitrogen of initial chemical fertilizer. On the other hand, 12, 14, and 16 kg/10 a /nitrogen are given for the small fields No.5, 6, 7, respectively as after chemical fertilizer in addition to the initial chemical fertilizer of 3 kg/ 10 a/ nitrogen for the small field No. 15, 17, 19, respectively. Therefore, rice crop grow rate differs each other paddy fields depending on the amount of nitrogen of chemical fertilizer.

III. EXPERIMENTS

A. Acquired Near Infrared Camera Imagery Data

Radio wave controlled helicopter mounted near infrared camera imagery data is acquired at C4-2, C4-3, C4-4 in SPARI on 24 September 2012 with the different viewing angle from the different altitudes. Figure 4 shows an example of the acquired near infrared image. There is spectralon of standard plaque as a reference of the measured reflectance in between C4-3 and C4-4. Just before the data acquisition, some of rice crops and leaves are removed from the subsection of paddy fields for inspection of nitrogen content. Using the removed rice leaves, nitrogen content in the rice leaves is measured based on the Dumas method⁶ (a kind of chemistry method) with Sumigraph NC-220F⁷ of instrument. The measured total nitrogen content is compared to the NIR reflectance.

The acquired near infrared camera imagery data is rectified and registered with geographic map of paddy field subsections as shown in Figure 5 (a) and (b).

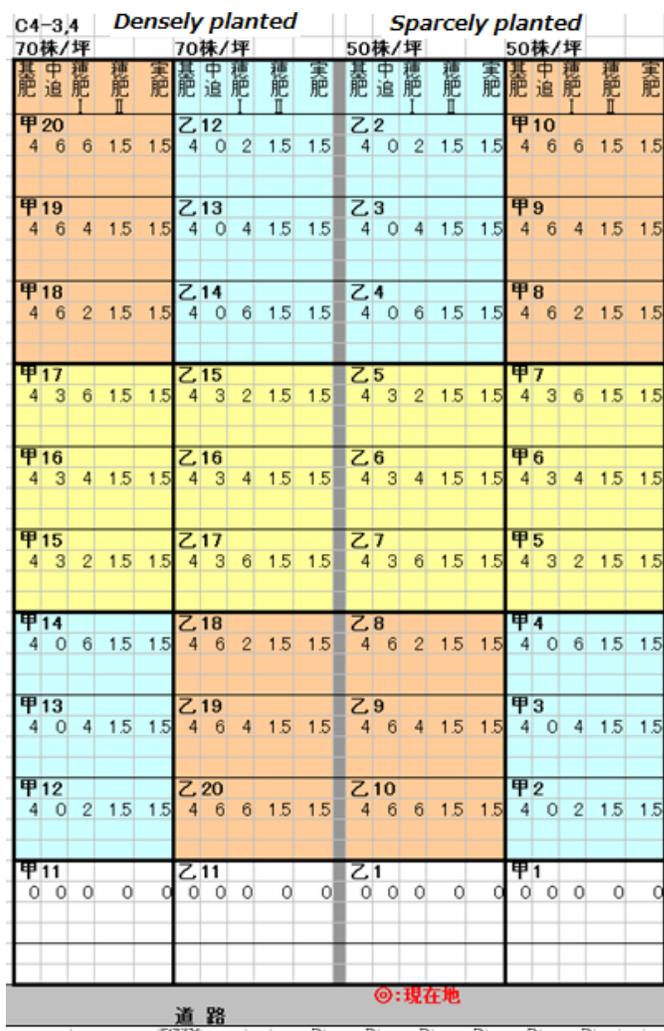


Fig. 3. Paddy field layout for investigation of nitrogen of chemical fertilizer dependency on rice crop quality

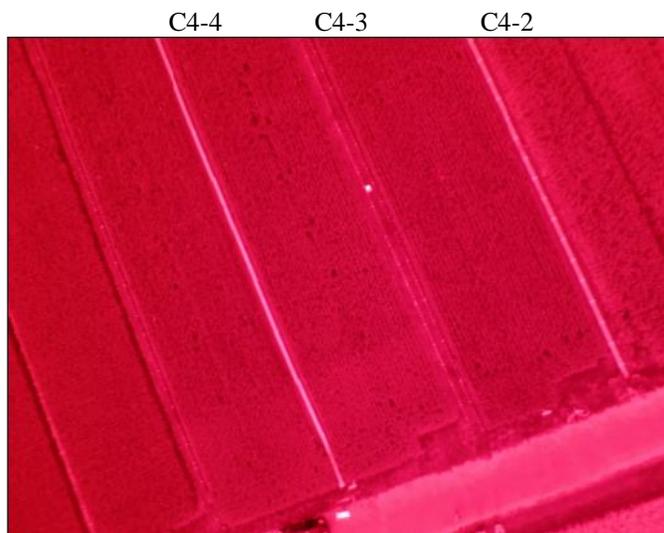
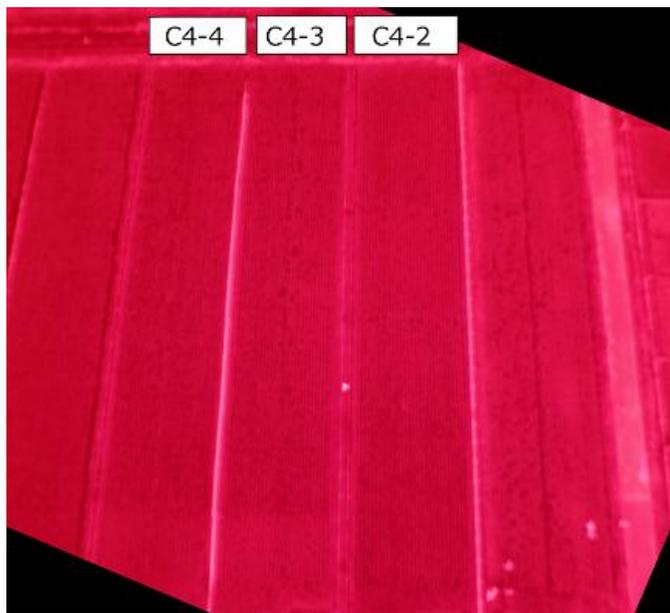


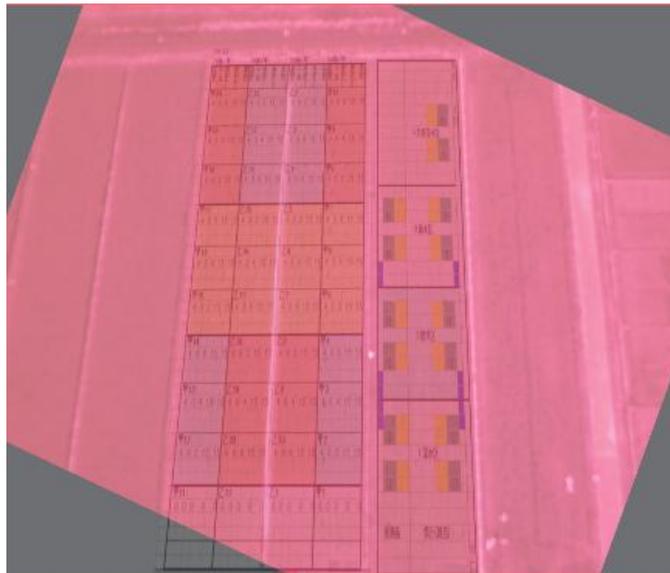
Fig. 4. Example of the acquired near infrared camera imagery data at SPARI on September 24 2012

⁶ <http://note.chiebukuro.yahoo.co.jp/detail/n92075>

⁷ http://www.scas.co.jp/service/apparatus/elemental_analyzer/sumigraph_nc-220F.html



(a)Rectified imagery data



(b)Superimposed with geographical map of subsections of the paddy fields

Fig. 5. Rectified and geometrically matched near infrared camera imagery data to the geographical map

B. Comparisons between Estimated Near Infrared Reflectance and Nitrogen Content Measured by Chemically

Table 2 shows the calculated near infrared reflectance of the subsections with the different conditions of water supply conditions and nitrogen of chemical fertilizer and measured nitrogen content in the rice leaves.

Through regressive analysis with these dataset, the relation between near infrared reflectance and nitrogen content is estimated. The results show so good correlation with 0.026 of R² value as shown in Figure 6.

TABLE II. ACTIVITY FOR 5 HIGHSET CATEGORY RESOLVED

Plot	NIR Reflectance	Nitrogen Content(%)
C4-2 1	0.7609	2.5913
C4-2 7	0.7939	2.5730
C4-2 11	0.7909	2.4011
C4-2 13	0.8093	2.4224
C4-3 甲1	0.7868	1.9056
C4-3 甲2	0.7781	2.2527
C4-3 甲3	0.7755	2.6008
C4-3 甲4	0.8038	2.5833
C4-3 甲5	0.7982	2.4630
C4-3 甲6	0.8184	2.6115
C4-3 甲7	0.8086	3.0306
C4-3 甲8	0.7992	2.9693
C4-3 甲9	0.7768	2.8576
C4-3 甲10	0.7581	2.7578
C4-4 甲11	0.7757	1.9511

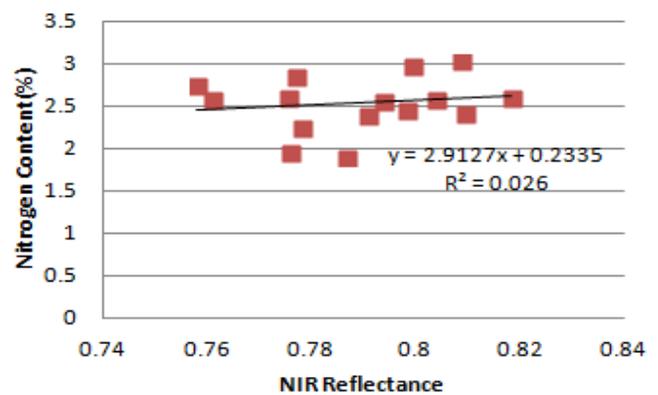


Fig. 6. Relation between Near Infrared: NIR reflectance and nitrogen content in the rice leaves

There are some influences due to shadow in the acquired near infrared images. When the pixels which do not affected by shadow are removed, then much nice relation between NIR reflectance and nitrogen content is obtained as shown in Figure 7. The results show much better correlation with 0.4971 of R² value.

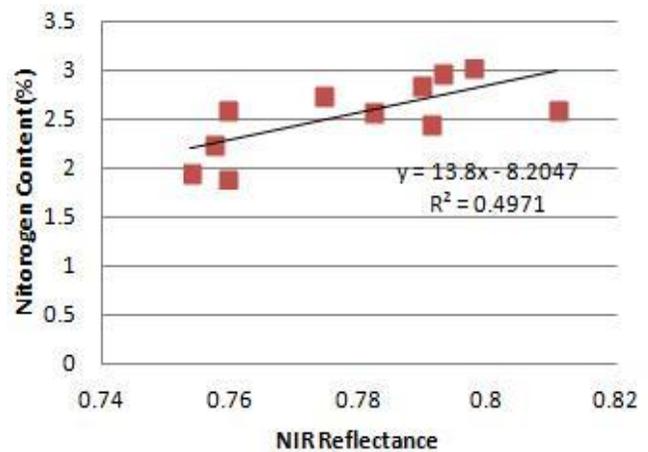


Fig. 7. Relation between Near Infrared: NIR reflectance and nitrogen content in the rice leaves after the pixel which affected by shadow is removed

Figure 8 shows example of the histogram of the NIR camera data of the subsection of rice paddy fields (Typically, histogram is bi-modal characteristics so that it is easy to distinguish between the pixels which are suffered from the shadow influence. In this paper, Ohtsu's method is applied to the data for discrimination of the pixels.

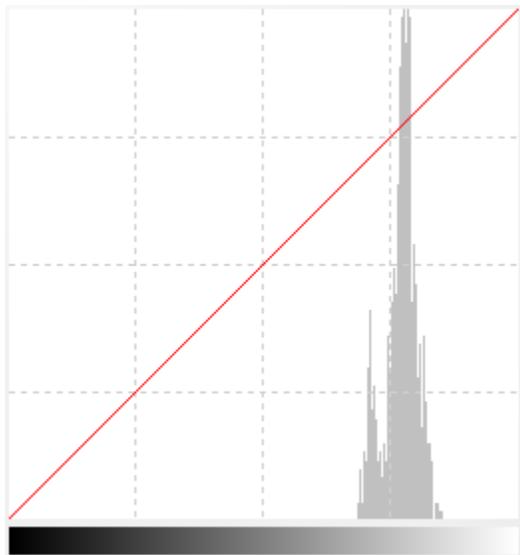


Fig. 8. Example of the histogram of the NIR camera data of the subsection of rice paddy fields (Typically, histogram is bi-modal characteristics so that it is easy to distinguish between the pixels which are suffered from the shadow influence.

Because we know the locations at where the sample of rice leaves are removed from the subsection of paddy fields, we could extract the pixels near the locations as shown in Figure 8. Dark portions in the paddy fields show the locations where the sample of rice leaves are removed from the subsection of paddy fields. The pixels near the location are extracted for the investigation of relations between NIR reflectance and nitrogen content.

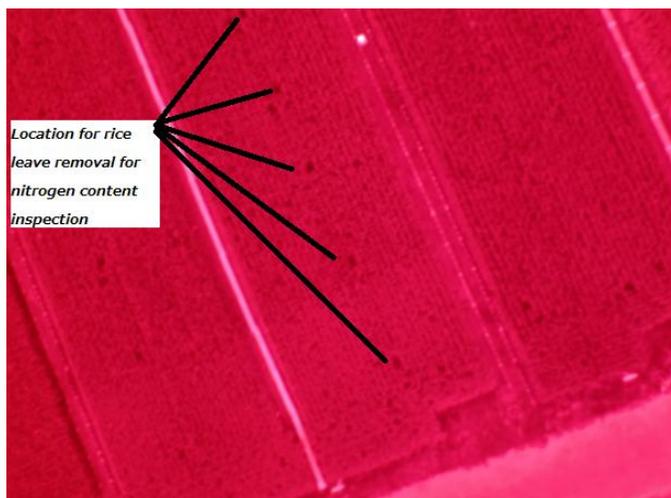


Fig. 9. Locations at where the sample of rice leaves are removed from the subsection of paddy fields

Then the relation between NIR reflectance and nitrogen content gets better up to 0.9674 of R^2 value as shown in Figure 9. The pair of NIR reflectance and nitrogen content is listed in Table 3.

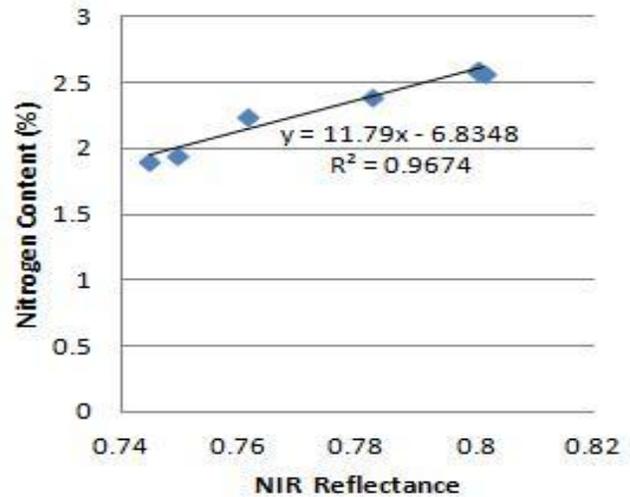


Fig. 10. Relation between nitrogen content in the rice leaves and near Infrared NIR reflectance of the rice leaves near the location at which the rice leaves are removed for nitrogen content measurements

TABLE III. ACTIVITY FOR 5 HIGHSET CATEGORY RESOLVED

Plot	NIR Reflectance	Nitrogen Content(%)
C4-3 甲1	0.7448	1.9056
C4-3 甲2	0.7613	2.2527
C4-3 甲3	0.8003	2.6008
C4-3 甲4	0.8015	2.5833
C4-4 甲11	0.7492	1.9511
C4-2 1	0.8001	2.5912
C4-2 11	0.7823	2.4011

C. Uniformity Evaluation with Radio Controlled Helicopter Based NIR Camera:

During the period from August 15 to September 24 2012, the rice crop fields are observed with radio controlled helicopter mounted NIR camera. Examples of the acquired images on September 28 are shown in Figure 9. C4-2 is situated on the right side of the photos while C4-3 is seen on the left hand side. In the middle of the photos, there is spectralon. It looks a small dots due to the fact that helicopter altitude is 30 m so that Instantaneous Field of View: IFOV is around 1.1 cm (Pixel size). Figure 9 (a) shows entire one shot of the acquired image with FOV of PowerShot of NIR camera while Figure 9 (b) shows enlarged portion of the acquired image. Meanwhile, Figure 9 (c) and (d) shows another shot of image at the different time on the same day. These show a good repeatability and reproduceability. NIR reflectance can be calculated by taking the ratio of the pixels value of the fields and that of Spectralon.

Uniformity in the small fields, C4-2, C4-3 are relatively good. Meanwhile, mean and variance are different by the small fields due to the fact that the given chemical fertilizers are different each other small fields.

IV. CONCLUSION

Rice crop quality evaluation method through regressive analysis between nitrogen content in the rice leaves and near infrared reflectance measurement data from near field, from radio wave controlled helicopter is proposed. Rice quality dependency on nitrogen of chemical fertilizer and water supply condition is evaluated.

Also homogeneity of the rice crop quality in the paddy fields is evaluated. Furthermore, influence due to shadow on near infrared reflectance of rice leaves measured from near field is taken into account in the rice crop quality evaluation processes.

Through experiments at test paddy fields in Saga Prefectural Agricultural Research Institute.

ACKNOWLEDGMENT

The author would like to thank Dr. Osamu Shigetomi, Yuko Miura of the Saga Prefectural Agricultural Research Institute: SPARI and Mr. Hideaki Munemoto of Saga University for their great effort to conduct the experiments.

REFERENCES

- [1] J.T.Compton, Red and photographic infrared linear combinations for monitoring vegetation, *Journal of Remote Sensing of Environment*, 8, 127-150, 1979.
- [2] C.Wiegand, M.Shibayama, and Y.Yamagata, Spectral observation for estimating the growth and yield of rice, *Journal of Crop Science*, 58, 4, 673-683, 1989.
- [3] Kohei Arai, Method for estimation of grow index of tealeaves based on Bi-Directional reflectance function:BRDF measurements with ground based network cameras, *International Journal of Applied Science*, 2, 2, 52-62, 2011.
- [4] Kohei Arai, Wireless sensor network for tea estate monitoring in complementally usage with Earth observation satellite imagery data

based on Geographic Information System(GIS), *International Journal of Ubiquitous Computing*, 1, 2, 12-21, 2011.

- [5] Kohei Arai, Method for estimation of total nitrogen and fiber contents in tealeaves with ground based network cameras, *International Journal of Applied Science*, 2, 2, 21-30, 2011.
- [6] Kohei Arai, Method for estimation of damage grade and damaged paddy field areas due to salt containing sea breeze with typhoon using remote sensing imagery data, *International Journal of Applied Science*, 2, 3, 84-92, 2011.
- [7] Kohei Arai, Monte Carlo ray tracing simulation for bi-directional reflectance distribution function and grow index of tealeaves estimation, *International Journal of Research and Reviews on Computer Science*, 2, 6, 1313-1318, 2011.
- [8] K.Arai, Fractal model based tea tree and tealeaves model for estimation of well opened tealeaf ratio which is useful to determine tealeaf harvesting timing, *International Journal of Research and Review on Computer Science*, 3, 3, 1628-1632, 2012.
- [9] K.Arai, H.Miyazaki, M.Akaishi, Determination of harvesting timing of tealeaves with visible and near infrared cameradata and its application to tea tree vitality assessment, *Journal of Japanese Society of Photogrammetry and Remote Sensing*, 51, 1, 38-45, 2012

AUTHORS PROFILE

Kohei Arai He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission-A of ICSU/COSPAR since 2008. He wrote 30 books and published 332 journal papers.

Sea Ice Concentration Estimation Method with Satellite Based Visible to Near Infrared Radiometer Data Based on Category Decomposition

Kohei Arai ¹

Graduate School of Science and Engineering
Saga University
Saga City, Japan

Abstract—Unmixing method for estimation of mixing ratio of the components of which the pixel in concern consists based on inversion theory is proposed together with its application to sea ice estimation method with satellite based visible to near infrared radiometer data. Through comparative study on the different unmixing methods with remote sensing satellite imagery data, it is found that the proposed inversion theory based unmixing method is superior to the other methods. Also it is found that the proposed unmixing method is applicable to sea ice concentration estimations.

Keywords—Unmixing; Inversion theory; Category decomposition; Remote Sensing

I. INTRODUCTION

Estimation of mixing ratio of the components of which the pixel in concern consists is called as “Unmixing”. Inversion theory based method is proposed together with its application to sea ice estimation method with satellite based visible to near infrared radiometer data. There are previously proposed methods such application oriented unmixing methods for hyperspectral instrument data, inversion theory based methods, etc.

On the other hands, there are many ice concentration estimation methods. Most of the methods utilize relations between brightness temperature and ice concentration. Estimation accuracy of the methods, however, is not good enough. In order to improve ice concentration estimation accuracy, inversion theory based unmixing method is introduced.

Through comparative study on the different unmixing methods with remote sensing satellite imagery data, it is found that the proposed inversion theory based unmixing method is superior to the other methods. Also it is found that the proposed unmixing method is applicable to sea ice concentration estimations.

The following section describes the proposed ice concentration estimation method followed by some experiments. Then conclusion is described with some discussions.

II. PROPOSED SEA ICE CONCENTRATION ESTIMATION METHOD

A. Conventional Method

The conventional method for sea ice concentration estimation method uses relation between brightness temperature which is acquired with microwave scanning radiometer onboard remote sensing satellite, T_B and sea ice concentration, C . Figure 1 shows an example of the relation. US Navy proposed this relation for sea ice concentration estimation. It is assumed empirically that brightness temperature of sea water is around 135K while that of sea ice is around 240K which is derived from NIMBUS-5/ESMR.

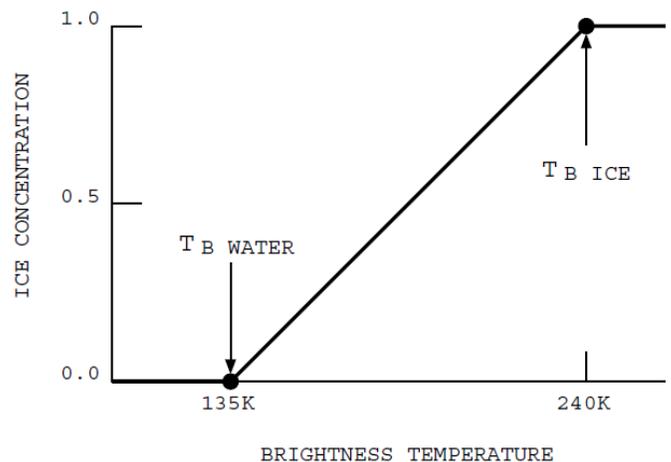


Fig. 1. Relation between brightness temperature and sea ice concentration

Gloesen et al., proposed the equation (1) for sea ice concentration.

$$C = \frac{T_b - 135}{\epsilon T_s - 135} \quad (1)$$

Where T_b and T_s denotes brightness temperature of sea ice and the ocean derived from NIMUS-7/ESMR and ϵ denotes emissivity of sea ice (around 0.92).

NASDA (Current JAXA) proposed the following equation for sea ice concentration estimation,

$$C = 4.17D - 220.83 \quad (2)$$

Where D denotes digital number which corresponds to the brightness temperature derived from vertical polarization of 37 GHz of frequency channel of MOS-1/MSR (Marine Observation Satellite-1/ Microwave Scanning Radiometer).

Meanwhile, NASA Team (SSIAWT) proposed the following method with horizontal and vertical polarization of 19 GHz channels and vertical polarization of 37GHz channel of digital count of data.

$$\begin{aligned} CT &= CF + CM \\ CF &= \frac{(C_1 + C_2PR + C_3GR + C_4PR * GR)}{D} \\ CM &= \frac{(C_9 + C_{10}PR + C_{11}GR + C_{12}PR * GR)}{D} \\ D &= C_5 + C_6PR + C_7GR + C_8PR * GR \\ PR &= \frac{TB(19, V) - TB(19, H)}{TB(19, V) + TB(19, H)} \\ GR &= \frac{TB(37, V) - TB(19, V)}{TB(37, V) + TB(19, V)} \end{aligned} \quad (3)$$

Where CT , CF , CM denotes total sea ice concentration, first year ice concentration, multiyear sea ice concentration. PR , GR denotes Polarization Ratio and Gradient Ratio, and D denotes digital number. Water vapor content in the atmosphere may affect to sea ice concentration estimations. In order to remove the influence due to water vapor, the following equation is utilizing,

$$GR > 0.05 \quad (4)$$

Then the following Comisso Bootstrap algorithm is widely used method for sea ice concentration estimation.

$$IC = \frac{\sqrt{(T_{B1} - T_{B1}^W)^2 + (T_{B2} - T_{B2}^W)^2}}{\sqrt{(T_{B1}^I - T_{B1}^W)^2 + (T_{B2}^I - T_{B2}^W)^2}} \quad (5)$$

The Bootstrap algorithm takes into account ocean area dependency (Antarctic and Arctic ocean areas are different treatments). Also open water is detected with the following equation,

$$WSLOPE * T_{B19V} + WINTRC > T_{B37V} \quad (6)$$

Seasonal changes and ocean area dependency are taken into account.

B. Inversion Theory Based Unmixing Method

Mixing ratio of component of which the observed pixel consists of the components, a_j can be estimated as follows,

$$\begin{aligned} P &= \sum_{j=1}^k a_j M_j \\ \sum_{j=1}^k a_j &= 1 \\ a_j &\geq 0 \end{aligned} \quad (7)$$

Where M_j denotes represented spectral characteristic of the component in concern, j and P denotes spectral characteristic of the observed pixel. Equation (7) can be rewritten by the following equation,

$$P = MA \quad (8)$$

where

$$\begin{aligned} P &= [p_1, p_2, \dots, p_n]^t \\ M &= [m_1, m_2, \dots, m_k] \\ A &= [a_1, a_2, \dots, a_k]^t \end{aligned}$$

If M is nonsingular matrix (square matrix), then mixing ratio vector can be estimated as follows,

$$A = M^{-1}P \quad (9)$$

Even if M is singular matrix (rectangle matrix), then mixing ratio vector can be expressed using the following Moore Penrose Inverse Matrix, M^+ .

$$A = (M^t M)^{-1} M^t P = M^+ P \quad (10)$$

It is possible to solve the equation with the following conditions,

$$\begin{aligned} |P - MA| &\rightarrow \min \\ u^t A &= 1 \quad (u = [1, 1, \dots, 1]^t) \\ a_j &\geq 0 \quad (j = 1, 2, \dots, k) \end{aligned} \quad (11)$$

Introducing Lagrange multiplier, then the equation can be solved as follows,

$$F(A, \lambda) = \frac{1}{2} |P - MA|^2 - \lambda(u^t A - 1) \quad (12)$$

Thus mixing ratio vector A can be estimated through solving the following equations,

$$\frac{\partial F}{\partial a_j} = -m_j^t (P - MA) - \lambda = 0 \quad (j = 1, 2, 3)$$

$$\frac{\partial F}{\partial \lambda} = -(u^t A - 1) = 0 \quad (13)$$

Equation (13) can be rewritten as follows, in the vector form,

$$-M^t (P - MA) - \lambda u = 0 \quad (14)$$

Thus

$$A = M^+P + \lambda(M^tM)^{-1}u \quad (15)$$

Where

$$\lambda = \frac{1 - u^tM^+P}{u^t(M^tM)^{-1}u} (M^tM)^{-1}u \quad (16)$$

Then A is rewritten from the equation (15) as follows,

$$A = M^+P + \frac{1 - u^tM^+P}{u^t(M^tM)^{-1}u} (M^tM)^{-1}u \quad (17)$$

Thus the mixing ratio vector can be estimated.

Although M denotes the representative observation vector, it contains error vector E as follows,

$$\begin{aligned} M &= M_0 + E \\ P &= M_0A_0 \end{aligned} \quad (18)$$

Namely, M is different from the true representative observation vector of M_0 . Then the true mixing ratio is expressed as follows,

$$A_1 = A_0 - M^+EA_0 + \frac{u^tM^+EA_0}{u^t(M^tM)^{-1}u} (M^tM)^{-1}u \quad (19)$$

Then estimation error of mixing ratio vector is expressed as follows,

$$\begin{aligned} |A_1 - A_0| &= \left| -M^+EA_0 + \frac{u^tM^+EA_0}{u^t(M^tM)^{-1}u} (M^tM)^{-1}u \right| \\ &\geq \left| \frac{u^tM^+EA_0}{u^t(M^tM)^{-1}u} (M^tM)^{-1}u \right| - |M^+EA_0| \\ &= \left| \frac{|\cos \beta|}{|\cos \alpha|} - 1 \right| |M^+EA_0| \end{aligned} \quad (20)$$

where α denotes the angle between u and $(M^tM)^{-1}u$ while β denotes the angle between u and M^+EA_0 . Then estimation error can be expressed as follows,

$$\begin{aligned} A' &= A_0 - M^+EA_0 \\ |A' - A_0| &= |M^+EA_0| \end{aligned} \quad (21)$$

Meanwhile the equation can be solved with the following conditions,

$$\begin{aligned} |M^+ - N| &\rightarrow \min \\ A &= NP \\ u^tA &= 1 \end{aligned} \quad (22)$$

These conditions are corresponding to the following conditions,

$$\begin{aligned} |A - M^+P| &\rightarrow \min \\ u^tA &= 1 \end{aligned} \quad (23)$$

Introducing Langrange multiplier, then

$$A = M^+P + \lambda u \quad (24)$$

Minimizing the error, M^+N , then

$$\begin{aligned} |V| &\rightarrow \min \\ u^t(M^+ - V)P &= 1 \end{aligned} \quad (25)$$

Introducing the following equation,

$$F(V, \lambda) = \frac{1}{2}|V|^2 - \lambda(u^t(M^+ - V)P - 1) \quad (26)$$

With the following conditional equations,

$$\frac{\partial F}{\partial V_{ji}} = V_{ji} + \lambda p_i = 0 \quad (i = 1, 2, \dots, n), (j = 1, 2, \dots, k)$$

$$\frac{\partial F}{\partial \lambda} = -(u^t(M^+ - V)P - 1) = 0 \quad (27)$$

Then V can be rewritten as follows in the vector representation,

$$V = -\lambda u P^+ \quad (28)$$

and then Langrange multiplier can be expressed as follows,

$$\lambda = \frac{1 - u^tM^+P}{|u|^2|P|^2} \quad (29)$$

Thus all unknown variables are estimated as follows,

$$\begin{aligned} V &= -\frac{1 - u^tM^+P}{|u|^2|P|^2} u P^+ \\ N &= M^+ + \frac{1 - u^tM^+P}{|u|^2|P|^2} u P^+ \\ A &= M^+P + \frac{1 - u^tM^+P}{|u|^2} u \end{aligned} \quad (30)$$

Estimated mixing ratio A_2 and estimation error, $|A_2 - A_0|$ are expressed as follows,

$$\begin{aligned} A_2 &= A_0 - M^+EA_0 + \frac{1 - u^tM^+EA_0}{|u|^2} u \\ |A_2 - A_0|^2 &= |M^+EA_0|^2 - \left(\frac{1 - u^tM^+EA_0}{|u|} \right)^2 \end{aligned} \quad (31)$$

M includes observation error so that P can be expressed with the following equation by considering the error ε ,

$$P = MA + \epsilon$$

$$M = \begin{pmatrix} p_1 & p_2 & \dots & p_n \\ m_{11} & m_{12} & \dots & m_{1k} \\ \vdots & \vdots & & \vdots \\ m_{n1} & m_{n2} & \dots & m_{nk} \end{pmatrix}^t$$

$$A = [a_1, a_2, \dots, a_k]^t$$

$$\epsilon = [\epsilon_1, \epsilon_2, \dots, \epsilon_k]^t \quad (32)$$

This method is referred to “Least Square Method”.

If the components of M are followed by normal distribution as follows,

$$N(m_{ij}^*, \sigma_{ij}^2) \quad (33)$$

where

$$m_i = m_i^* \cdot A, m_i^* = [m_{i1}^*, m_{i2}^*, \dots, m_{ik}^*]$$

$$\sigma_i^2 = A^t \cdot S_i \cdot A + \sigma_{e_i}^2, S_i = \text{diag}(\sigma_{i1}^2, \sigma_{i2}^2, \dots, \sigma_{ik}^2)$$

Then the probability of observation vector M can be expressed as follows,

$$Q(p_i) = \frac{1}{(2\pi \cdot \sigma_i^2)^{\frac{1}{2}}} \exp\left(-\frac{(p_i - m_i)^2}{2\sigma_i^2}\right) \quad (34)$$

It is rewritten in the vector form as follows,

$$Q(P) = \prod_{i=1}^n Q(p_i) \quad (35)$$

Best estimation of mixing ratio vector is to maximizing probability of observation vector so that,

$$R(P) = -\ln(Q(P))$$

$$\sum_{j=1}^k A_j = 1, A_j \geq 0, (j = 1, \dots, k) \quad (36)$$

All of the meshed points of equation (36) can be calculated then the maximum probability of the observation vector can be found which results in estimation of mixing ratio. This method is referred to “Maximum Likelihood Method”.

III. EXPERIMENTS

C. Data Used

ADEOS/AVNIR (Advanced Earth Observing Satellite/Advanced Satellite/Advanced Visible to Near Infrared Radiometer) imagery data of Saroma Lake in Hokkaido, Japan which is acquired on February 1997 is used for the experiments. AVNIR consist of four spectral bands, blue (B1), green (B2), red (B3) and near infrared (B4) wavelength regions.

From the original imagery data, 400 by 400 pixels of portion of image are extracted for the experiments. Instantaneous Field of View: IFOV is 16 meters. Figure 2 shows the averaged image among the four spectral bands of imagery data. The typical digital numbers of the spectral bands are shown in Table 1.

TABLE I. TYPICAL PIXEL VALUE OF THE COMPONENTS FOR ALL SPECTRAL BANDS OF ADEOS/AVNIR

	B1	B2	B3	B4
Open Water	40.6875	26.5312	19.3125	13.5625
Thin Ice	119.141	107.609	101.078	89.1250
Thick Ice	130.000	118.609	113.984	106.188

Figure 2 shows the ADEOS/AVNIR image for experiments. Figure 3 (a), (b), and (c) shows open water, thin sea ice, and thick sea ice estimated by Maximum Likelihood Method.

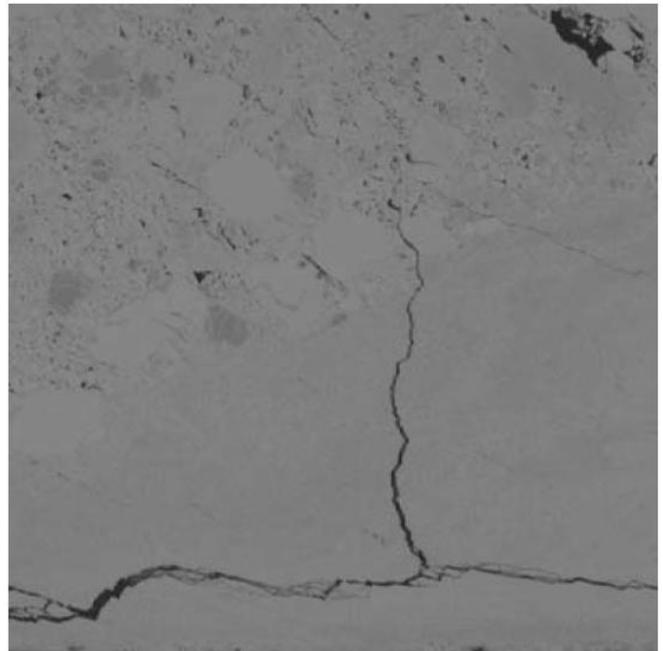
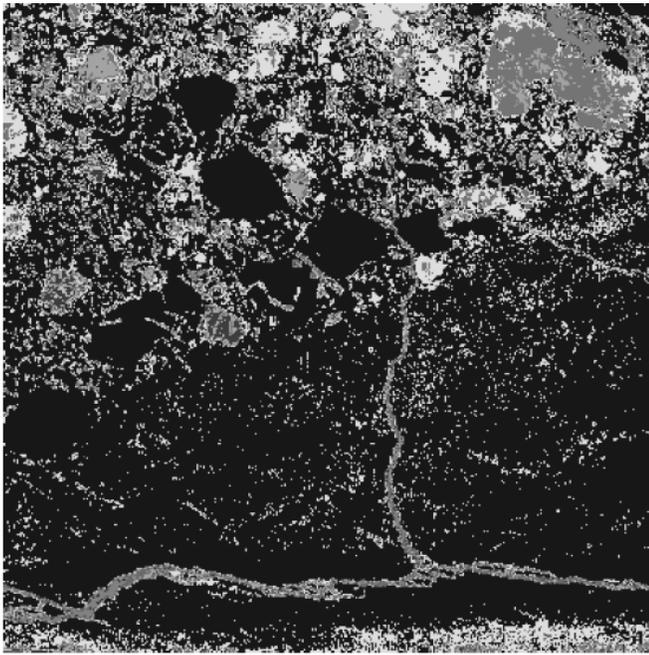
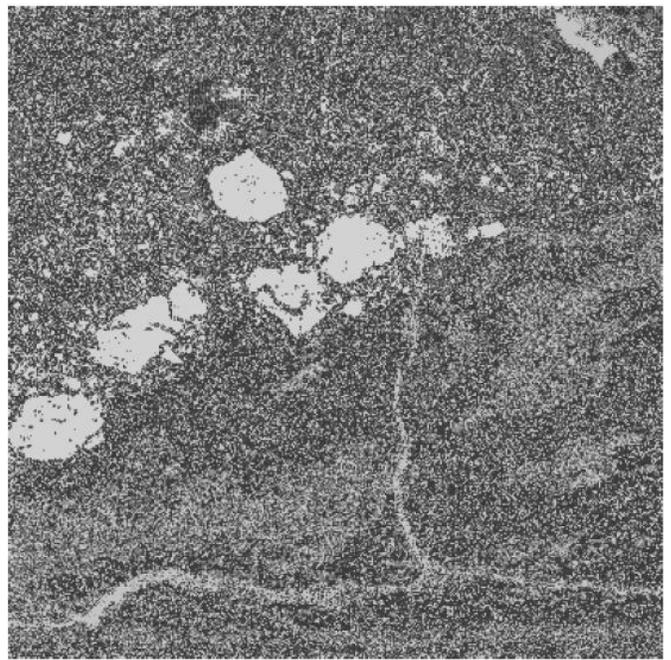


Fig. 2. ADEOS/AVNIR (Advanced Earth Observing Satellite/Advanced Visible to Near Infrared Radiometer) imagery data of Saroma Lake in Hokkaido, Japan which is acquired on February 1997 is used for the experiments. AVNIR consist of four spectral bands, blue (B1), green (B2), red (B3) and near infrared (B4) wavelength regions. This image is averaged image among the four spectral bands

On the other hands, Figure 4 (a), (b), and (c) shows three components estimated by Least Square Method. Through comparisons between Figure 3 and 4, Maximum Likelihood Method derived open water, thin and thick sea ice is more likely in comparison to the visual perception.

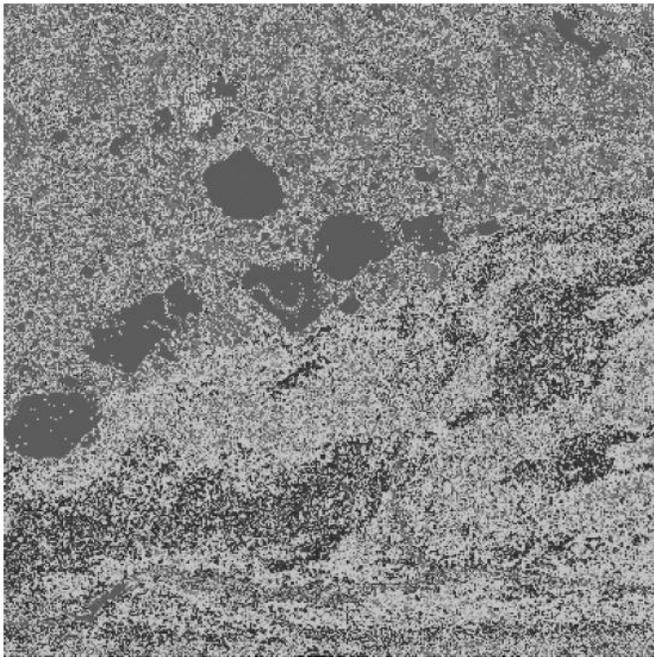


(a)Open Water



(c)Thick Sea Ice

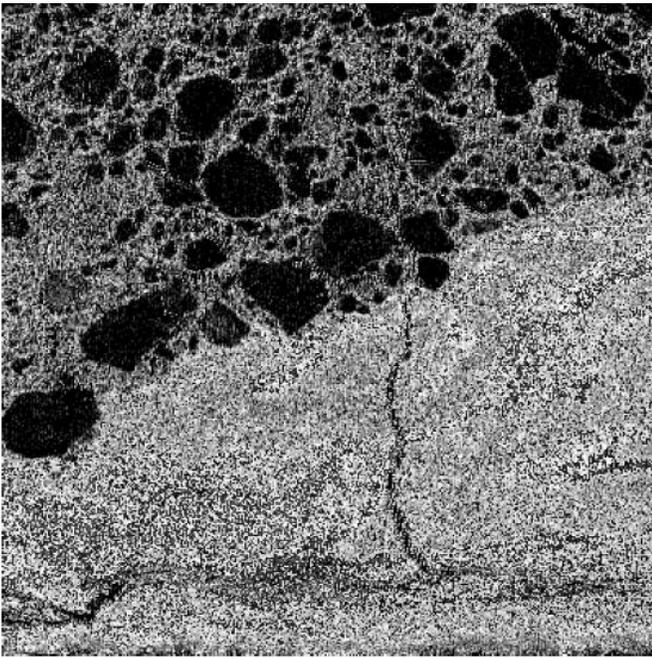
Fig. 3. Estimated three components, open water, thin and thick ice derived from Maximum Likelihood Method



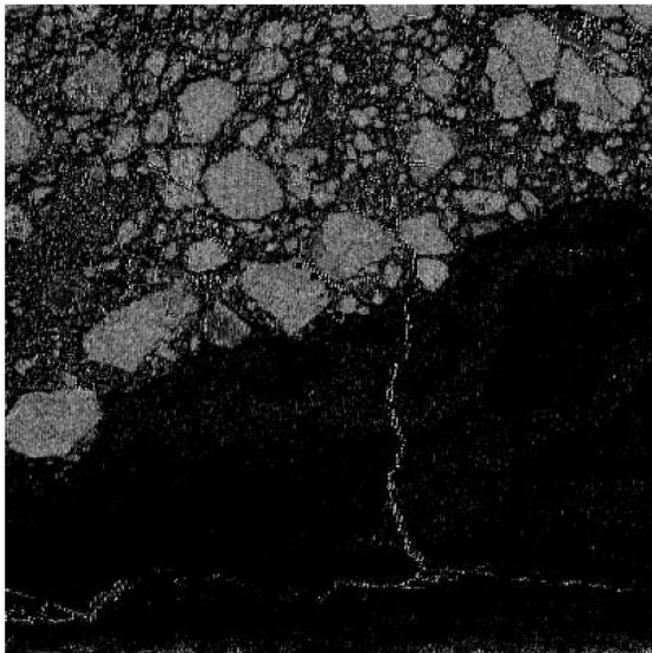
(b)Thin Sea Ice



(a)Open Water



(b)Thin Sea Ice



(c)Thick Sea Ice

Fig. 4. Estimated three components, open water, thin and thick ice derived from Least Square Method

Figure 5 shows the extracted “Lead” which is defined as breaking portion of sea ice. The percentage ratio of open water ranged from 7 to 50% is defined as Lead. By using unmixing method, mixing ratio of open water can be estimated.

Therefore, the Lead is extracted. The extracted lead, however, is disconnected. Therefore, some consideration of connectivity of the piece of the disconnected lead is required.

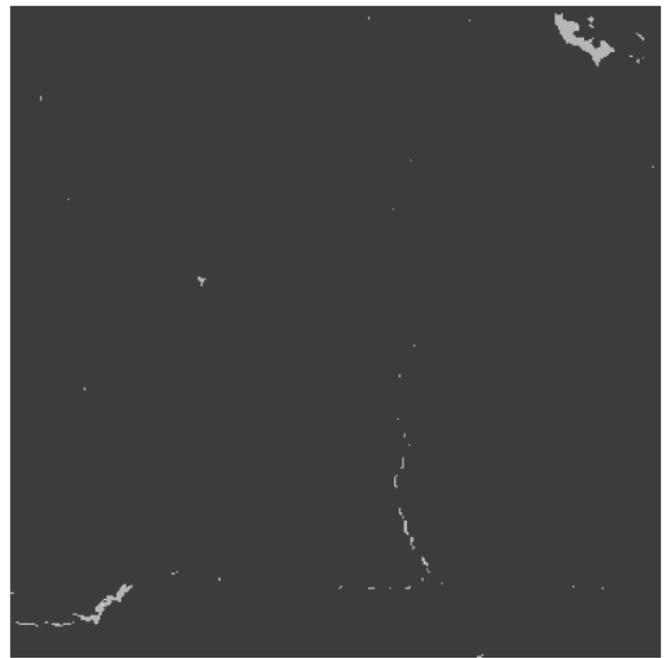


Fig. 5. Estimated “Lead” of which mixing ratio of open water ranges from 7 to 50% derived from Maximum Likelihood Method

ADEOS/AVNIR Band 1 image is shown in Figure 6. By using band 1 of imagery data, connectivity between disconnected portions of lead can be found. This information is called as contextual information and can be extracted by using 3 by 3 or 5 by 5 pixel windows. For instance, connectivity between the pixel in concern and the 8 surrounding pixels can be checked with 3 by 3 windows. Using the connectivity, it is possible to connect between disconnected portions of lead. Figure 7 shows the connected lead using contextual information.

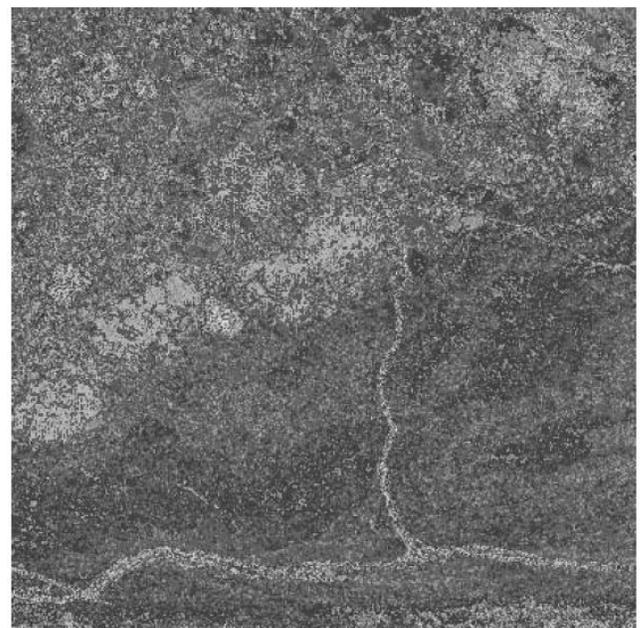


Fig. 6. ADEOS/AVNIR Band 1 image of Saroma, Hokkaido, Japan acquired on February 3 1997.

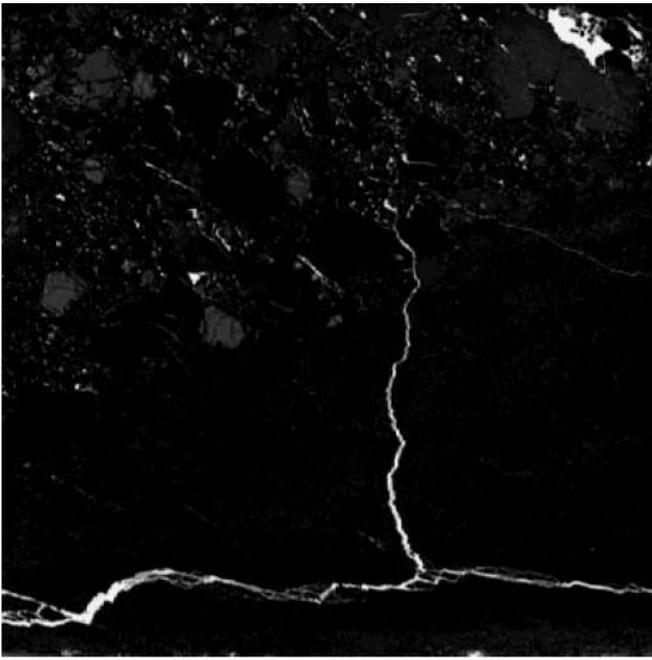


Fig. 7. Connected lead using contextual information.

IV. CONCLUSION

Unmixing method for estimation of mixing ratio of the components of which the pixel in concern consists based on inversion theory is proposed together with its application to sea ice estimation method with satellite based visible to near infrared radiometer data.

Through comparative study on the different unmixing methods with remote sensing satellite imagery data, it is found that the proposed inversion theory based unmixing method is superior to the other methods. Also it is found that the proposed unmixing method is applicable to sea ice concentration estimations. It is also found that contextual information is effective to connect disconnected portion of lead.

ACKNOWLEDGMENT

The author would like to thank Ms. Izumi Fujio for her effort to conduct the experiments.

REFERENCES

- [1] N. Keshava, J. Kerekes, D. Manolakis, and G. Shaw, "An Algorithm Taxonomy for Hyperspectral Unmixing," *SPIE* **4049**, 2000, pp. 42–63.
- [2] A. Tarantola and B. Valette, "Generalized Nonlinear Inverse Problems Solved Using the Least Squares Criterion," *Rev. Geophys. Space Phys.* **20** (2), 1982, pp. 219–232.

AUTHORS PROFILE

Kohei Aarai He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission-A of ICSU/COSPAR since 2008. He wrote 30 books and published 332 journal papers.

Sensitivity Analysis on Sea Surface Temperature Estimation Methods with Thermal Infrared Radiometer Data through Simulations

Kohei Arai¹

Graduate School of Science and Engineering
Saga University
Saga City, Japan

Abstract—Sensitivity analysis on Sea Surface Temperature: SST estimation with Thermal Infrared Radiometer: TIR data through simulations is conducted. Also Conjugate Gradient Method: CGM based SST estimation method is proposed. SST estimation error of the proposed CGM based method is compared to the conventional Split Window Method: SWM with a variety of conditions including atmospheric models. The results show that the proposed CGM based method is superior to the SWM.

Keywords—SST estimation; Split Window; Conjugate Gradient; MODTRAN; atmospheric model

I. INTRODUCTION

Sea Surface Temperature: SST estimation with thermal infrared radiometer onboard satellites is well known and widely used in a variety of research fields, in particular climate changes, global warming, etc. SST estimation methods are proposed [1]-[4]. Most of these are based on regressive analysis and use several spectral bands in Thermal Infrared: TIR wavelength region. The most dominant atmospheric factor is precipitable water. Using the different wavelength TIR bands whose influences due to water vapor are different, it is possible to reduce the influence. The most popular method is Multi Channel Sea Surface Temperature: MCSST [5]. Also previously proposed SST estimation methods are summarized by I. Barton [6]. In the same time, comparative study among the previously proposed methods is well reported [7].

Based on radiative transfer equation, inversion based SST estimation method is proposed [8]. Nonlinear radiative transfer equation is linearized then optimum combination of wavelength regions are selected [9]. Other than that, Geographic Information System: GIS based neural network is proposed for SST estimation method [10]. In this paper, sensitivity analysis results are described. SST estimation accuracy, in general, depends on relative humidity, air temperature, meteorological range, wind speed, aerosol type, and so on. Sensitivity of these factors on SST estimation accuracy is clarified in order to make clear that how does component influencing to SST estimation accuracy.

The following section describes the method for sensitivity analysis together with some theoretical background followed by some experiments. Then conclusion is described together with some discussions.

II. PROPOSED METHOD

A. Theoretical Background on SST Estimation with Thermal Infrared Radiometer Data

Radiation from a blackbody with physical temperature of T is expressed in equation (1)

$$B_{\nu}(T) = \frac{2hc^2}{\lambda^5(\exp(\frac{hc}{\lambda kT}) - 1)} [W \cdot cm^{-2} \cdot sr^{-1} \cdot \mu m^{-1}] \quad (1)$$

where

k : Boltzman constant [J/K]

h : Plank constant [J · s]

c : Light speed [m/s]

λ : Wavelength at wave number ν

The contribution from the atmosphere can be expressed as follows,

$$\tau(\theta, z_{\infty}, z) = \exp \left\{ - \int_z^{z_{\infty}} \frac{\rho(z)k(z)}{\cos(\theta)} dz \right\} \quad (2)$$

where

θ : Observation zenith angle

ρ : Density of atmospheric constituents

k : Volume extinction coefficient

and

$$\int_z^{z_{\infty}} \frac{\rho(z)k(z)}{\cos(\theta)} dz$$

is called as optical depth of the atmosphere.

B. At Sensor Radiance of Thermal Infrared Radiometer

For sea surface observation with TIR radiometers onboard remote sensing satellites, radiance includes three components, the contribution from sea surface, the contribution from the reflected radiance at sea surface and the contribution from the atmosphere.

$$I(\theta) = \int_{\lambda_1}^{\lambda_2} \Phi(\lambda) \left[\epsilon_{\lambda} B_{\lambda}(T_s) + (1 - \epsilon_{\lambda}) \int_{z_s}^{\infty} B_{\lambda}[T(z)] \frac{\partial \tau(\theta, z_{\infty}, z)}{\partial z} dz \right] \cdot \tau(\theta, z_{\infty}, z_s) + \int_{z_s}^{\infty} B_{\lambda}[T(z)] \frac{\partial \tau(\theta, z_{\infty}, z)}{\partial z} dz \Big] d\lambda \quad (3)$$

where

T_s : Sea surface temperature[K]

Φ : Spectral response function

ϵ : Emissivity

τ : Transparency

Spectral response function means spectral sensitivity function of spectral bands of TIR onboard satellites. In general, emissivity of sea surface in TIR wavelength region is almost 1. Therefore, the second term of the equation (3) can be neglected.

$$I(\theta) = \int_{\lambda_1}^{\lambda_2} \Phi(\lambda) \left\{ \epsilon_{\lambda} B_{\lambda}(T_s) \tau(\theta, z_{\infty}, z_s) + \int_{z_s}^{\infty} B_{\lambda}[T(z)] \frac{\partial \tau_{\lambda}(\theta, z_{\infty}, z)}{\partial z} dz \right\} d\lambda \quad (4)$$

C. Method for Sensitivity Analysis

By using MODTRAN of radiative transfer software code with six default atmospheric models, Tropic, Mid. Latitude Summer and Winter, Sub Arctic Summer and Winter, and 1976 US standard atmosphere, brightness temperature of the assumed spectral bands in Thermal Infrared wavelength regions can be estimated. Therefore Root Mean Square Error: RMSE of SST estimation error can be estimated for the assumed SST estimation method.

D. Assumed SST Estimation Method

Assuming spectral response function in the spectral wavelength region of spectral band is 1, then the equation (4) and be rewritten as follows,

$$I(\theta) = B_{\lambda}(T_s) \tau(\theta, z_{\infty}, z_s) + \int_{z_s}^{\infty} B_{\lambda}[T(z)] \frac{\partial \tau_{\lambda}(\theta, z_{\infty}, z)}{\partial z} dz \quad (5)$$

The second term of equation (5) can be approximated as follows,

$$\int_{z_s}^{\infty} B_{\lambda}[T(z)] \frac{\partial \tau_{\lambda}(\theta, z_{\infty}, z)}{\partial z} dz = [1 - \tau_{\lambda}(\theta, z_{\infty}, z)] I_{ai} \quad (6)$$

Where I_{ai} denotes representative of spectral band i of radiance. Atmospheric transparency can be rewritten as follows,

$$\begin{aligned} \tau_{\lambda}(u, \theta) &= c_{1i} \exp[-1(c_{2i} + c_{3i}m)u^{c_{4i} + c_{5i}m}] \\ &= c_{1i} \exp[-(c_{2i} + \frac{c_{3i}}{\cos \theta})u^{c_{4i} + \frac{c_{5i}}{\cos^p \theta}}] \\ m &\approx 1/\cos \theta \end{aligned} \quad (7)$$

where u denotes perceptible water while m denotes slant length between sea surface and TIR instrument onboard satellites. In the TIR wavelength region, perceptible water is major absorbing continuants in the atmosphere. Through simulation studies with radiative transfer code of MODTRAN with six atmospheric models (Tropic, Mid. Latitude Summer, Mid. Latitude Winter, Sub Arctic Summer, Sub Arctic Winter and 1976 US Standard Atmosphere), the coefficients are obtained as shown in Table 1. Then I_{ai} is calculated as follows,

$$I_{ai} = F_i(I_{ak}) = A_{1i} + A_{2i}I_{ak} \quad (8)$$

The coefficients in the equation (8) are calculated with MODTRAN in the same manner which is mentioned above. Table 2 shows the results.

TABLE I. COEFFICIENTS OF EQUATION (7) OBTAINED WITH MODTRAN OF ATMOSPHERIC SOFTWARE CODE

	c_{n1}	c_{n2}	c_{n3}
c_{1i}	0.8507924	0.9356485	0.9253728
c_{2i}	-0.0754923	-0.03505476	-0.03752114
c_{3i}	0.175898	0.08923810	0.1261287
c_{4i}	1.451688	1.739096	1.679308
c_{5i}	-0.2339985	-0.1563839	-0.1293923

TABLE II. COEFFICIENTS OF EQUATION (8) OBTAINED WITH MODTRAN OF ATMOSPHERIC SOFTWARE CODE

	A_{n1}	A_{n2}	A_{n3}
A_{1i}	-0.88610×10^{-6}	0.0	0.75270×10^{-6}
A_{2i}	0.62180	1.0	1.0590

Consequently, radiance of spectral band i can be expressed as follows,

$$I_i = B_i[T_s] \tau_i(u, \theta) + [1 - \tau_i(u, \theta)] F_i(I_{ak}) \quad (9)$$

In order to avoid divergence of the solution, the following conditional equation is introduced.

$$X = \frac{X_{max} + X_{min}}{2} + \frac{X_{max} - X_{min}}{\pi} \arctan \xi \quad (10)$$

The unknown factors are as follows,

$$\mathbf{x} = (T_s, u, I_{ak}) \quad (11)$$

Namely, sea surface temperature, perceptible water, and representative radiance. The following cost function is introduced,

$$J(\mathbf{x}) = \sum_{i=1}^3 (I_i - \hat{I}_i)^2 \quad (12)$$

Then iteration is stopped when the cost function is below the designated value,

$$J(\mathbf{x}) \leq \varepsilon \quad (13)$$

Radiance of the spectral band i can be rewritten as follows,

$$I_i = \frac{c1c_{1i}}{\lambda^3 \exp(\frac{c2}{\lambda T_i}) - 1} \exp \left\{ - \left(c_{2i} + \frac{c3i}{\cos \theta} \right) u^{c_{4i} + \frac{c5i}{\cos \theta}} \right\} + \left[1 - c_{1i} \exp \left\{ - \left(c_{2i} + \frac{c3i}{\cos \theta} \right) u^{c_{4i} + \frac{c5i}{\cos \theta}} \right\} \right] (C_{1i} + C_{2i} I_{ak}) \quad (14)$$

Then the following updating equation is introduced,

$$\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} + \beta^{(n)} A^{(n)} \nabla J[\mathbf{x}^{(n)}] \quad (15)$$

It is rewritten in matrix and vector as follows,

$$\begin{pmatrix} T_s \\ u \\ I_{ak} \end{pmatrix}^{(n+1)} = \begin{pmatrix} T_s \\ u \\ I_{ak} \end{pmatrix}^{(n)} + \beta^{(n)} \begin{pmatrix} \frac{\partial^2 J}{\partial T_s^2} & \frac{\partial^2 J}{\partial T_s \partial u} & \frac{\partial^2 J}{\partial T_s \partial I_{ak}} \\ \frac{\partial^2 J}{\partial u \partial T_s} & \frac{\partial^2 J}{\partial u^2} & \frac{\partial^2 J}{\partial u \partial I_{ak}} \\ \frac{\partial^2 J}{\partial I_{ak} \partial T_s} & \frac{\partial^2 J}{\partial I_{ak} \partial u} & \frac{\partial^2 J}{\partial I_{ak}^2} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial J}{\partial T_s} \\ \frac{\partial J}{\partial u} \\ \frac{\partial J}{\partial I_{ak}} \end{pmatrix}^{(n)} \quad (16)$$

where

$$\beta^{(n)} = 1/2^n \quad (17)$$

In general,

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, \dots, \frac{\partial f(\mathbf{x})}{\partial x_n} \right)^T \quad (18)$$

and

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix} \quad (19)$$

Therefore, the cost function can be rewritten as follows,

$$J = f(\mathbf{x}), H_n = \nabla^2 f(\mathbf{x}) \quad (20)$$

where

$$H_n = \begin{pmatrix} \frac{\partial^2 J}{\partial T_s^2} & \frac{\partial^2 J}{\partial T_s \partial u} & \frac{\partial^2 J}{\partial T_s \partial I_{ak}} \\ \frac{\partial^2 J}{\partial u \partial T_s} & \frac{\partial^2 J}{\partial u^2} & \frac{\partial^2 J}{\partial u \partial I_{ak}} \\ \frac{\partial^2 J}{\partial I_{ak} \partial T_s} & \frac{\partial^2 J}{\partial I_{ak} \partial u} & \frac{\partial^2 J}{\partial I_{ak}^2} \end{pmatrix} \quad (21)$$

is called Hessian or Hesse matrix.

Equation (1) can be rewritten as follows,

$$B_\nu(T_s) = \frac{c_i}{\exp(\frac{c_2}{T_s}) - 1}$$

$$B_\nu(T) = \frac{c1}{\lambda^3 (\exp(\frac{c2}{\lambda T}) - 1)} \quad (22)$$

where

$$c_i : 2hc^2/\lambda_i^5$$

$$c_{0i} : ch/\lambda_i k$$

$$c1 : 2hc = 1.191126^{-12}$$

$$c2 : ch/k = 1.43889$$

Then the unknown variables are estimated through iterations. Also Root Mean Square Error: RMSE can be evaluated.

$$RMSE = \sqrt{\frac{\sum_{k=1}^N (\epsilon_k)^2}{N}}$$

$$\epsilon_k = T_s - T'_s \quad (23)$$

The first derivatives of the cost function are expressed as follows,

$$\frac{\partial J}{\partial T_s} = -2 \sum_{i=1}^3 (I_i - \hat{I}_i) \frac{\partial \hat{I}_i}{\partial T_s}$$

$$\frac{\partial J}{\partial u} = -2 \sum_{i=1}^3 (I_i - \hat{I}_i) \frac{\partial \hat{I}_i}{\partial u}$$

$$\frac{\partial J}{\partial I_{ak}} = -2 \sum_{i=1}^3 (I_i - \hat{I}_i) \frac{\partial \hat{I}_i}{\partial I_{ak}} \quad (24)$$

Also the second derivatives are represented as follows,

$$\frac{\partial^2 J}{\partial T_s^2} = 2 \sum_{i=1}^3 \left\{ \frac{\partial \hat{I}_i}{\partial T_s} \frac{\partial \hat{I}_i}{\partial T_s} - (I_i - \hat{I}_i) \frac{\partial^2 \hat{I}_i}{\partial T_s^2} \right\}$$

$$\frac{\partial^2 J}{\partial T_s \partial u} = 2 \sum_{i=1}^3 \left\{ \frac{\partial \hat{I}_i}{\partial u} \frac{\partial \hat{I}_i}{\partial T_s} - (I_i - \hat{I}_i) \frac{\partial^2 \hat{I}_i}{\partial T_s \partial u} \right\}$$

$$\frac{\partial^2 J}{\partial T_s \partial I_{ak}} = 2 \sum_{i=1}^3 \left\{ \frac{\partial \hat{I}_i}{\partial I_{ak}} \frac{\partial \hat{I}_i}{\partial T_s} - (I_i - \hat{I}_i) \frac{\partial^2 \hat{I}_i}{\partial T_s \partial I_{ak}} \right\}$$

$$\frac{\partial^2 J}{\partial u \partial T_s} = 2 \sum_{i=1}^3 \left\{ \frac{\partial \hat{I}_i}{\partial T_s} \frac{\partial \hat{I}_i}{\partial u} - (I_i - \hat{I}_i) \frac{\partial^2 \hat{I}_i}{\partial u \partial T_s} \right\}$$

$$\frac{\partial^2 J}{\partial u^2} = 2 \sum_{i=1}^3 \left\{ \frac{\partial \hat{I}_i}{\partial u} \frac{\partial \hat{I}_i}{\partial u} - (I_i - \hat{I}_i) \frac{\partial^2 \hat{I}_i}{\partial u^2} \right\}$$

$$\frac{\partial^2 J}{\partial u \partial I_{ak}} = 2 \sum_{i=1}^3 \left\{ \frac{\partial \hat{I}_i}{\partial I_{ak}} \frac{\partial \hat{I}_i}{\partial u} - (I_i - \hat{I}_i) \frac{\partial^2 \hat{I}_i}{\partial u \partial I_{ak}} \right\}$$

$$\frac{\partial^2 J}{\partial T_s \partial T_s} = 2 \sum_{i=1}^3 \left\{ \frac{\partial \hat{I}_i}{\partial T_s} \frac{\partial \hat{I}_i}{\partial I_{ak}} - (I_i - \hat{I}_i) \frac{\partial^2 \hat{I}_i}{\partial I_{ak} \partial T_s} \right\}$$

$$\frac{\partial^2 J}{\partial I_{ak} \partial u} = 2 \sum_{i=1}^3 \left\{ \frac{\partial \hat{I}_i}{\partial u} \frac{\partial \hat{I}_i}{\partial I_{ak}} - (I_i - \hat{I}_i) \frac{\partial^2 \hat{I}_i}{\partial I_{ak} \partial u} \right\}$$

$$\frac{\partial^2 J}{\partial I_{ak}^2} = 2 \sum_{i=1}^3 \left\{ \frac{\partial \hat{I}_i}{\partial I_{ak}} \frac{\partial \hat{I}_i}{\partial I_{ak}} - (I_i - \hat{I}_i) \frac{\partial^2 \hat{I}_i}{\partial I_{ak}^2} \right\} \quad (24)$$

The first derivatives of radiance are expressed as follows,

$$\frac{\partial I_i}{\partial T_s} = \frac{c_i c_{0i} c_{1i}}{T_s^2 \{ \exp(\frac{c_{0i}}{T_s}) - 1 \}^2} \exp \left\{ \frac{c_{0i}}{T_s} - \alpha_1 u^{\alpha_2} \right\}$$

$$\frac{\partial I_i}{\partial u} = c_{1i} \alpha_1 \alpha_2 \alpha_3 u^{\alpha_2 - 1} \left\{ (C_{1i} + C_{2i} I_{ak}) - \frac{c_i}{\exp(\frac{c_{0i}}{T_s}) - 1} \right\}$$

$$\frac{\partial I_i}{\partial I_{ak}} = C_{2i} \{ 1 - c_{1i} \exp(-\alpha_1 u^{\alpha_2}) \} \quad (25)$$

Also the second derivatives of radiance is represented as follows,

$$\frac{\partial^2 I_i}{\partial T_s^2} = \left[c_i c_{0i} c_{1i} \exp \{ -\alpha_1 u^{\alpha_2} \} \right]$$

$$\left\{ \frac{-c_{0i} e^{\frac{c_{0i}}{T_s}} - 2T_s e^{\frac{c_{0i}}{T_s}}}{T^4 (e^{\frac{c_{0i}}{T_s}} - 1)^2} + \frac{2c_{0i} e^{\frac{c_{0i}}{T_s}}}{T^4 (e^{\frac{c_{0i}}{T_s}} - 1)^3} \right\}$$

$$\frac{\partial^2 I_i}{\partial T_s \partial u} = -\frac{c_i c_{0i} c_{1i}}{T_s^2 \{ \exp(\frac{c_{0i}}{T_s}) \}^2} \exp \left\{ \frac{c_{0i}}{T_s} - \alpha_1 u^{\alpha_2} \right\} \alpha_1 \alpha_2 u^{\alpha_2 - 1}$$

$$\frac{\partial I_i}{\partial T_s \partial I_{ak}} = 0$$

$$\frac{\partial^2 I_i}{\partial u \partial T_s} = -\frac{c_i c_{0i} c_{1i}}{T_s^2 \{ \exp(\frac{c_{0i}}{T_s}) - 1 \}^2} \alpha_1 \alpha_2$$

$$\exp \left\{ \frac{c_{0i}}{T_s} - \alpha_1 u^{\alpha_2} \right\} u^{\alpha_2 - 1}$$

$$\frac{\partial^2 I_i}{\partial u^2} = c_{1i} \alpha_1 \alpha_2 \left\{ (C_{1i} + C_{2i} I_{ak}) - \frac{c_i}{\exp(\frac{c_{0i}}{T_s}) - 1} \right\}$$

$$\frac{\partial^2 I_i}{\partial u \partial I_{ak}} = c_{1i} C_{2i} \alpha_1 \alpha_2 \alpha_3 u^{\alpha_2 - 1}$$

$$\frac{\partial^2 I_i}{\partial I_{ak} \partial T_s} = 0$$

$$\frac{\partial^2 I_i}{\partial I_{ak} \partial u} = c_{1i} C_{2i} \alpha_1 \alpha_2 \alpha_3 u^{\alpha_2 - 1}$$

$$\frac{\partial^2 I_i}{\partial I_{ak}^2} = 0 \quad (26)$$

where

$$\alpha_1 = c_{2i} + \frac{c_{3i}}{\cos \theta}$$

$$\alpha_2 = c_{4i} + \frac{c_{5i}}{\cos \theta}$$

$$\alpha_3 = \exp \{ -\alpha_1 u^{\alpha_2} \}$$

E. Assumed Spectral Bands

Spectral bands of ADEOS/OCTS (Advanced Earth

Observing Satellite / Ocean Color and Temperature Scanner) are assumed as typical spectral bands for SST estimation which are 10300-11360nm for Band 4, and 11360-12500nm for Band 5, respectively.

III. EXPERIMENTS

F. Simulation Conditions

The following parameters are set for the experiments with MODTRAN obtaining at sensor radiance of spectral TIR band data.

Atmospheric Model: Tropic, Mid.Latitude Summer,
Mid.Latitude Winter, SubArctic Summer, SubArctic
Winter, 1976 US Standard Atmosphere

Constraint: ± 2 [K]

Meteorological Range: $\pm 0\%$, $\pm 10\%$, $\pm 20\%$

Relative Humidity: Default x1.0, x1.1, x1.2, x0.9, x0.8

Air-Temperature: Default ± 0 , ± 3 [K]

Sea surface temperature: Default ± 0 , ± 3 [K]

Wind speed: 3.5, 7.0, 14.0 [m/s]

Aerosol Model: Navy Maritim, Maritim, Tropospheric, Desert

Observation Zenith Angle: 0, 30, 60 [deg]

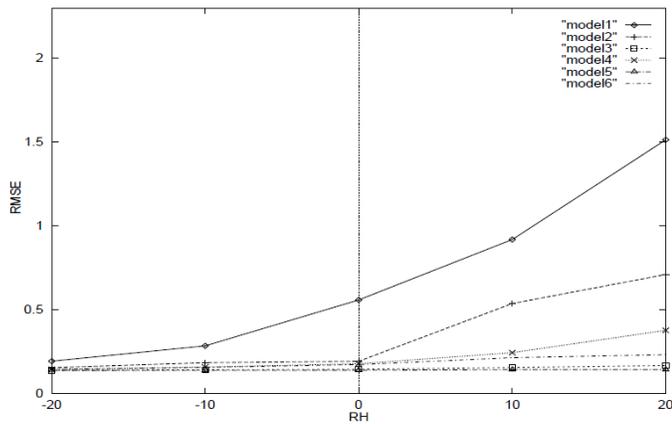
Then SST is estimated with the proposed method and the conventional split window method using the calculated at sensor radiance.

G. Evaluation of RMSE for the Proposed CGM

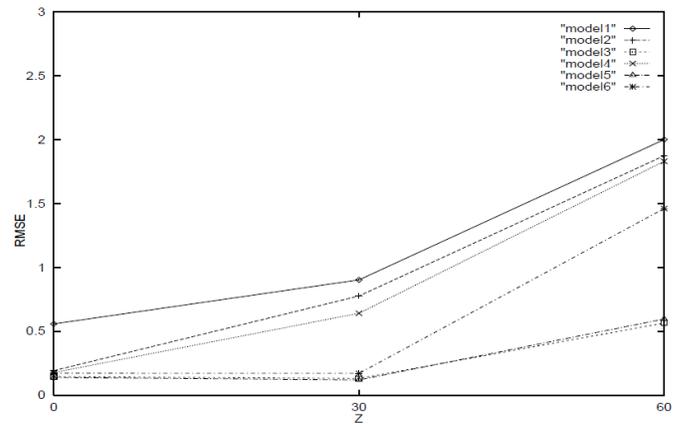
SST estimation error for the proposed CGM can be evaluated with RMSE which is expressed in equation (23) by using MODTRAN derived at sensor radiance of TIR bands.

Figure 1 show RMSE of CGM as functions of (a) relative humidity, (b) meteorological range, (c) air temperature, (d) observation zenith angle, (e) sea surface temperature, and (f) wind speed for six atmospheric models. In accordance with increasing of relative humidity, meteorological range, observation zenith angle, sea surface temperature, and wind speed, RMSE increased. RMSE is, on the other hands, decreases in accordance with increasing of air temperature. Meanwhile, Figure 2 (a) shows the relation between meteorological range and RMSE as parameters of different types of aerosol for the Mid.

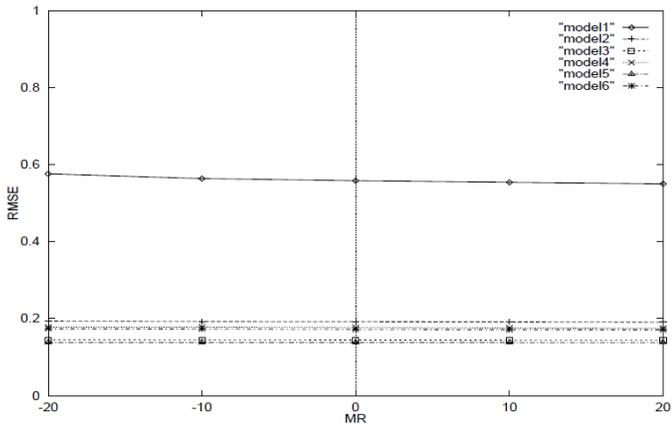
Latitude Summer of atmospheric model. RMSE for Navy Maritime shows the greatest followed by Maritime, Desert, and Troposphere aerosol. Figure 2 (b) shows RMSE for the six different aerosol types, Navy Maritime, Maritime, Urban, Desert, Rural, and Troposphere aerosol types at the meteorological range of 23 km. Figure 2 (c) shows RMSE as function of altitude for Desert aerosol. Figure 3 shows RMSE of the representative radiance from the atmosphere for the six different atmospheric models. It does not show monotonic relation between relative humidity and RMSE. Therefore, the representative radiance from the atmosphere has to be estimated precisely.



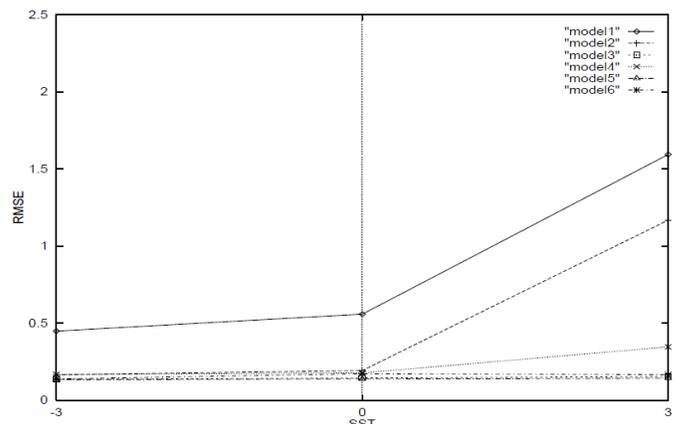
(a)Relative Humidity



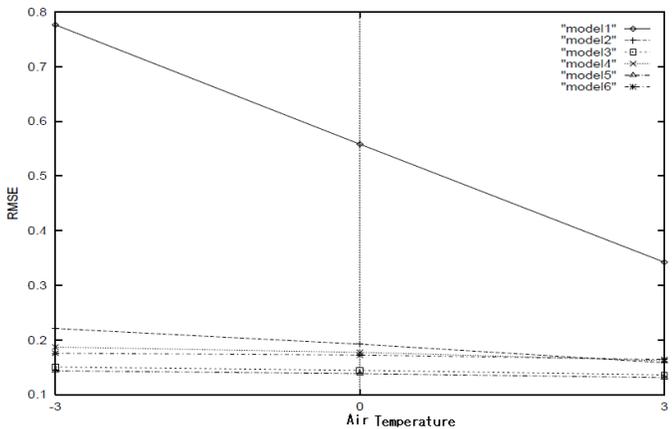
(d)Observation Zenith Angle



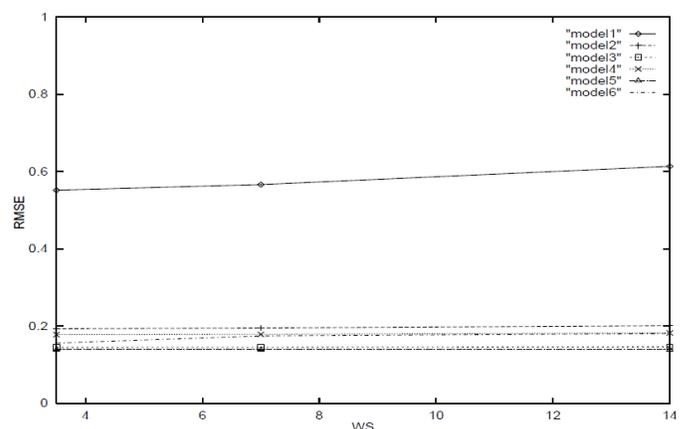
(b)Meteorological Range



(e)Sea Surface Temperature

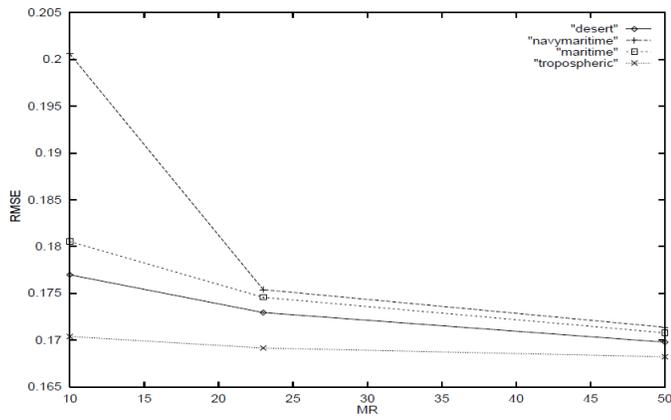


(c)Air Temperature

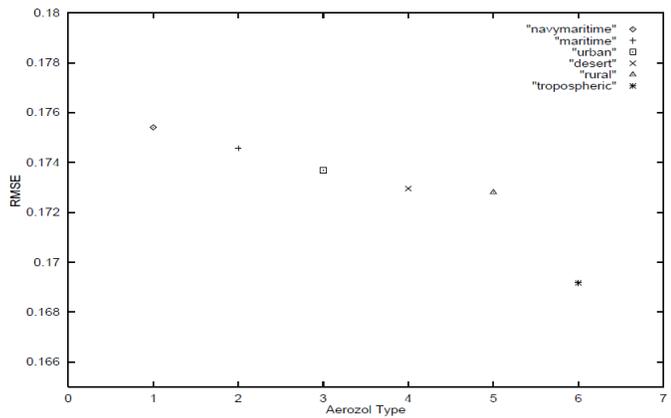


(f)Wind Speed

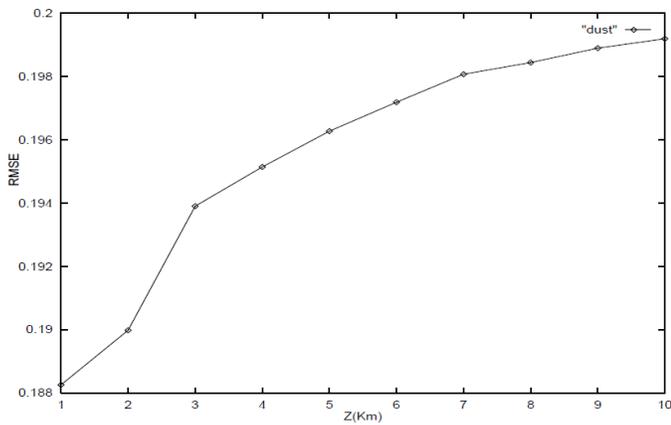
Fig. 1. RMSE of CGM as functions of (a) relative humidity, (b) meteorological range, (c) air temperature, (d) observation zenith angle, (e) sea surface temperature, and (f) wind speed for six atmospheric models



(a) Aerosol Type



(b) Aerosol Type



(c) Altitude (Desert type of aerosol)

Fig. 2. RMSE as function of aerosol types

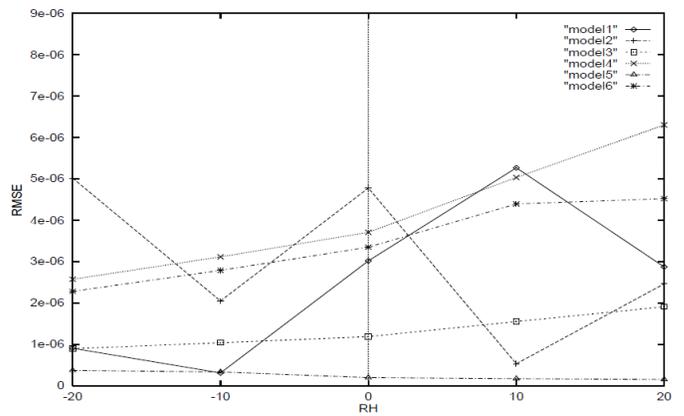
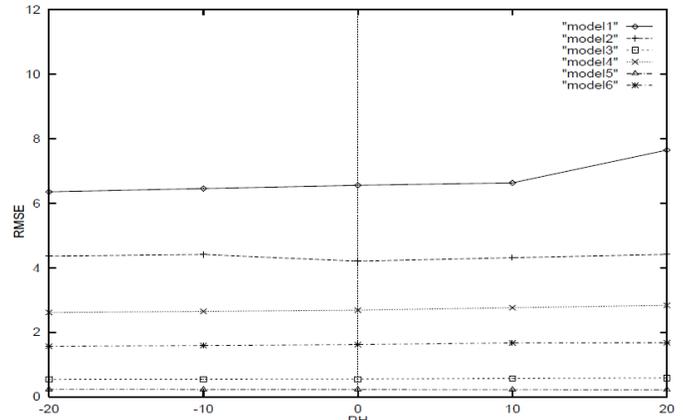


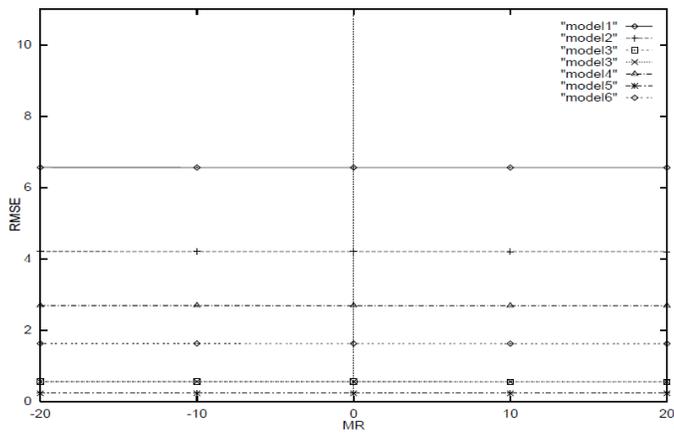
Fig. 3. RMSE of the representative radiance from the atmosphere for the six different atmospheric models

H. Evaluation of RMSE for the Conventional Split Window

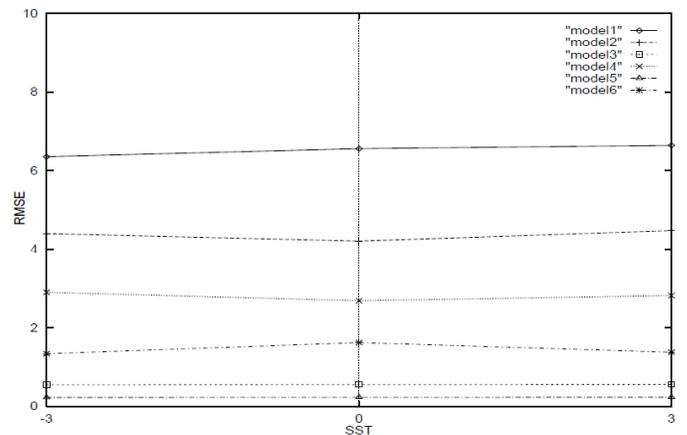
SST estimation error for the conventional split window can be evaluated with RMSE which is expressed in equation (23) by using MODTRAN derived at sensor radiance of TIR bands. Figure 4 shows RMSE of Split Window as functions of (a) relative humidity, (b) meteorological range, (c) air temperature, (d) observation zenith angle, (e) sea surface temperature, and (f) wind speed for six atmospheric models. RMSE of the conventional Split Window is much larger than that of CGM. In accordance with increasing of relative humidity, meteorological range, observation zenith angle, sea surface temperature, and wind speed, RMSE increased. RMSE is, on the other hands, decreases in accordance with increasing of air temperature.



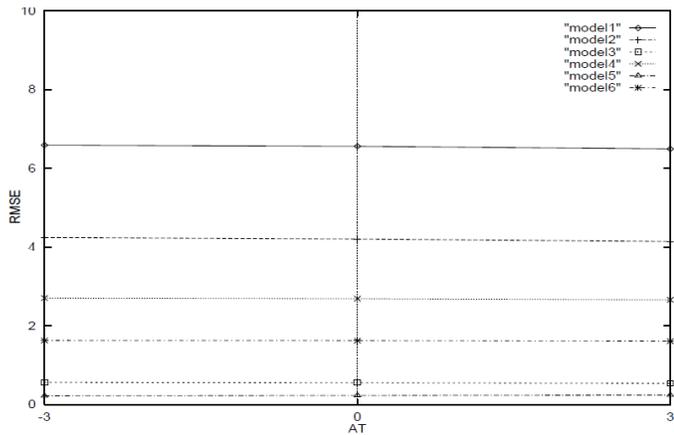
(a) Relative Humidity



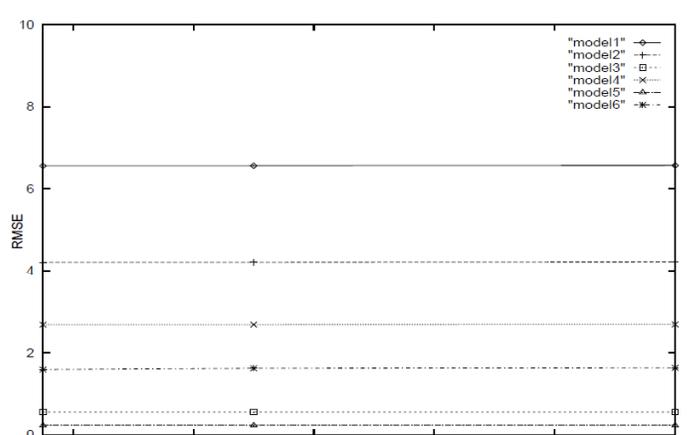
(b) Meteorological Range



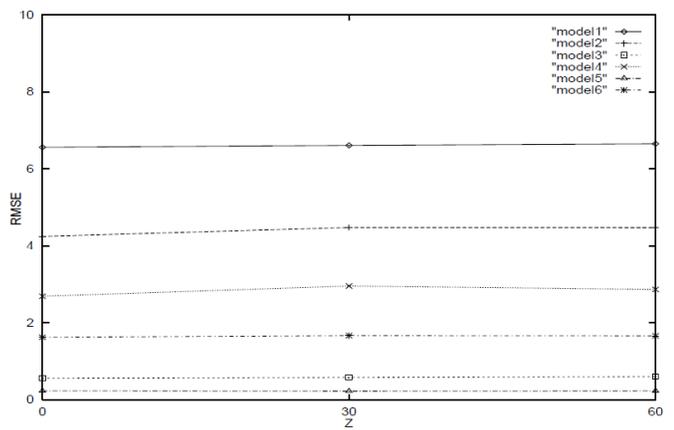
(e) Sea Surface Temperature



(c) Air Temperature



(f) Wind Speed



(d) Observation Zenith Angle

Fig. 4. RMSE of the conventional Split Window as functions of (a) relative humidity, (b) meteorological range, (c) air temperature, (d) observation zenith angle, (e) sea surface temperature, and (f) wind speed for six atmospheric models

I. Comparison of RMSE between Split Window and CGM

Overall RMSE of the conventional Split Window and the proposed CGM is shown in Table 3.

Although depending on the atmospheric model, RMSE between both are different, RMSE of the proposed CGM is lower than that of Split Window. Therefore, CGM is superior to Split Window.

TABLE III. OVERALL RMSE OF THE CONVENTIONAL SPLIT WINDOW AND THE PROPOSED CGM

	Split Window	Conjugate Gradient
Tropic	1.195	0.864
Mid. Latitude Summer	0.702	0.479
Mid. Latitude Winter	0.472	0.408
Sub Arctic Summer	0.641	0.596
Sub Arctic Winter	0.483	0.317
1976 US Standard	0.505	0.478
Average	0.726	0.559

J. Influence Due to Observation Noise

In order to evaluate influence due to observation noise on SST estimation accuracy, RMSE with and without of random number derived noise is evaluated. The normal distribution of random number with 10^{-6} of variance and with zero mean is generated by using Messene Twister. The random number is added to the at sensor radiance of the simulated TIR bands data. Then SST is estimated based on the proposed conjugate gradient method. Table 4 shows the result.

TABLE IV. INFLUENCE DUE TO OBSERVATION NOISE ON SST ESTIMATION ACCURACY

	Without noise	With 10^{-6} of noise
Tropic	0.4386	0.7214
Mid. Latitude Summer	0.1924	0.6012
Mid. Latitude Winter	0.1446	0.2871
Sub Arctic Summer	0.1774	0.5624
Sub Arctic Winter	0.1384	0.2601
1976 US Standard	0.1724	0.3631

IV. IV. CONCLUSION

Sensitivity analysis on Sea Surface Temperature: SST estimation with Thermal Infrared Radiometer: TIR data through simulations is conducted. Also Conjugate Gradient Method: CGM based SST estimation method is proposed. SST estimation error of the proposed CGM based method is compared to the conventional Split Window Method: SWM with a variety of conditions including atmospheric models. The results show that the proposed CGM based method is superior to the SWM.

ACKNOWLEDGMENT

The author would like to thank Mr. Noriyuki Takamatsu for his contributions to conduct experiments.

REFERENCES

- [1] Anding, D. and R. Kauth (1970): Estimation of sea surface temperature from space. *Remote Sens. Environ.*, **1**, 217–220.
- [2] McMillin, L. M. (1975): Estimation of sea surface temperatures from two infrared window measurements with different absorption. *J. Geophys. Res.*, **80**, 5113–5117.
- [3] Prabhakara, C., G. Dalu and V. G. Kunde (1974): Estimation of sea surface temperature from remote sensing in the 11- to 13- μ m window region. *J. Geophys. Res.*, **79**, 5039–5044.
- [4] Kohei Arai, Preliminary Assessment of Radiometric Accuracy for MOS-1 Sensors, International Journal of Remote Sensing, Vol.9, No.1, pp.5-12, Apr.1988.
- [5] Kidwell, K. B. (1991): NOAA Polar Orbiter Data Users Guide. NOAA/NESDIS/NCDC/SDSD, Washington, D.C., 192 pp.
- [6] Barton, I. J. (1995): Satellite-derived sea surface temperatures: Current status. *J. Geophys. Res.*, **100**, 8777–8790.
- [7] Moriyama, M., K. Mima and K. Arai, Comparison of the sea surface temperature estimation methods, Advances in Space Reserach, Vol.16, NO.10, pp.121-124, A1-044, July 1994.
- [8] K. Arai, Hidemasa Kobayashi, Masao Moriyama, Sea surface temperature estimation method based on Inversion and Linearized Inversion, Journal of Remote Sensing Society of Japan, 18, 3, 43-50, 1998.
- [9] K. Arai, Optimum band selection for sea surface teperature estimation based on regressive analysis, Journal of Japanese Remote Sensing and Photogrammetry, 39, 2, 74-81, 2000.
- [10] K. Arai, Sea Surface Temperature estimation method for ocean areas and seasons using Geographic Information System as a neural network, Sediment and Ecohydraulics, Proc.in Marine Science, 9, Elsevier ISBN978-0-444 53184-1, 2007

AUTHORS PROFILE

Kohei Arai He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission-A of ICSU/COSPAR since 2008. He wrote 30 books and published 332 journal papers.

Image Clustering Method Based on Self Organization Mapping: SOM Derived Density Maps and Its Application for Landsat Thematic Mapper Image Clustering

Kohei Arai¹

Graduate School of Science and Engineering
Saga University
Saga City, Japan

Abstract—A new method for image clustering with density maps derived from Self-Organizing Maps (SOM) is proposed together with a clarification of learning processes during a construction of clusters. Simulation studies and the experiments with remote sensing satellite derived imagery data are conducted. It is found that the proposed SOM based image clustering method shows much better clustered result for both simulation and real satellite imagery data. It is also found that the separability among clusters of the proposed method is 16% longer than the existing k-mean clustering. It is also found that the separability among clusters of the proposed method is 16% longer than the existing k-mean clustering. In accordance with the experimental results with Landsat-5 TM image, it takes more than 20000 of iteration for convergence of the SOM learning processes.

Keywords—SOM clustering; Density map; Boundary image; Labeling algorithm

I. INTRODUCTION

Clustering method is widely used for data analysis and pattern recognition [1]-[4]. Meanwhile, Self Organizing Map: SOM proposed by T. Kohonen is a neural network with two layers which allows use as un-supervised classification, or learning method [5] based on a similarity between separable data groups to be classified [6]. In other word, SOM is a visualization tool for multi-dimensional data rearranging the data in accordance with a similarity based on a learning process with the statistical characteristics of the data. It is used to be used for pattern recognition in combination with Learning Vector Quantization (LVQ¹). SOM is consists of m-dimensional input layer which represent as a vector and two dimensional output layer which is also represented as a vector connected each other nodes between input and output layers with weighting coefficients. In a learning process, winning unit is chosen based on the difference between input vector and weighting coefficients vector then the selected unit and surrounding units get closer to the input vector.

¹ http://en.wikipedia.org/wiki/Learning_Vector_Quantization

SOM is utilized for clustering [7]. After a learning process, a density map² is created in accordance with code vector density. Based on the density map, a pixel labeling³ can be done. This is the basic idea on the proposed image clustering method with SOM learning. Other than this, clustering methods with learning processes, reinforcement learning is also proposed for image retrievals [8] and rescue simulations [9]. Also probability density model for SOM is proposed.

The image clustering method with SOM learning based on density map is proposed in the following section followed by simulation study results and the experimental results with satellite remote sensing imagery data. Then finally, conclusions and some discussions are described.

II. PROPOSED CLUSTERING METHOD BASED ON DENSITY MAP DERIVED FROM LEARNING PROCESS OF SOM

A. SOM Learning Process

Firstly imagery data are mapped to a feature space. In parallel, SOM learning process creates a density map in accordance with a similarity between the mapped data in the feature space and density map or between input data in the feature space and two dimensional density maps. As a result of SOM learning process, code vector is obtained [10]. It is easy to recognize the density of the code vector visually. Although code vector density map represent cluster boundaries, it is not easy that neither to determine a boundary nor to put a label to the pixel in concern by using the density map. The method proposed here is to use density map for finding boundaries among sub-clusters then some of sub-clusters which have a high similarity are to be merged in the following procedure,

1) Create density map based on SOM learning

² <http://books.google.co.jp/books?id=wxvQoFy1YBgC&pg=SA1-PA210&lpg=SA1-PA210&dq=density+map+SOM&source=bl&ots=sU95Gi28ug&sig=uZBXSATAqYaXPJtkmrGHts7uoqU&hl=ja&sa=X&ei=hijYT7L0ClibiQfn0NSTAw&ved=0CGkQ6AEwBA#v=onepage&q=density%20map%20SOM&f=false>
³ <http://books.google.co.jp/books?id=jJad-0gh8YwC&pg=PA69&dq=pixel+labeling&hl=ja&sa=X&ei=ZinYT4CpFYjUmAWW1cCfAw&ved=0CDUQ6AEwAA#v=onepage&q=pixel%20labeling&f=false>

- 2) Binary image is generated from the density map
- 3) Define sub-clusters in accordance with the separated areas of the binary image
- 4) Calculate similarities of the sub-clusters
- 5) Merge the sub-clusters which show the highest similarity
- 6) Process (4) and (5) until the number of clusters reaches the desired number of clusters

Representing input vector, $x(t)$ and reference (or output) vector, $m(t)$, neural network proposed by T. Kohonen is expressed as follows,

$$m(t+1)=m(t)+h_i(t)[x(t)-m(t)] \quad (1)$$

Where $h(t)$ denotes neighboring function or weighting function including learning coefficients.

$$h_i(t)=a(t), \text{ when } i \in N(t) \\ =0, \text{ when } i \notin N(t) \quad (2)$$

Where $N(t)$ denotes the number or size of neighboring units. $a(t)$ is called learning coefficient and ranges from 0 to 1 as is expressed as follows,

$$a(t)=a_0(1-t/T) \quad (3)$$

Where a_0 is an initial value and T denotes the number of total learning number or the number of update. In the equation (1), $[x(t)-m(t)]$ implies cost function⁴ which should be minimized, and if

$$c=\underset{i}{\operatorname{argmin}}. \|x-m_i\| \quad (4)$$

is obtained then such m_i unit is called winning unit. The neighboring unit is defined around m_i unit. The size of the neighboring unit, $N(t)$ is a variable which starts with a relatively large then is getting small reaching to the winning unit only after the SOM learning process.

$$N(t)=N(0)(1-t/T) \quad (5)$$

The SOM learning process is illustrated in Figure 1.

B. Conventional Clustering Method

The existing clustering algorithm such as k-means clustering algorithm⁵ is similar to the SOM learning process. If m_i is redefined as mean vector of cluster i , then the cost function defined in the k-means clustering is expressed as follows,

$$J=\sum \|x(t)-m_i(x(t))\|^2 \quad (6)$$

Therefore, the mean vector of each cluster is determined to minimize the equation (6) of cost function. Let $I(x(t))$ be a binary function and is equal to 1 if the $x(t)$ belongs to the cluster i and is 0 if the $x(t)$ does not belong to the cluster i , then the cost function can be rewritten as follows,

$$J'=\sum\sum I(x(t)) \|x(t)-m_i(x(t))\|^2 \quad (7)$$

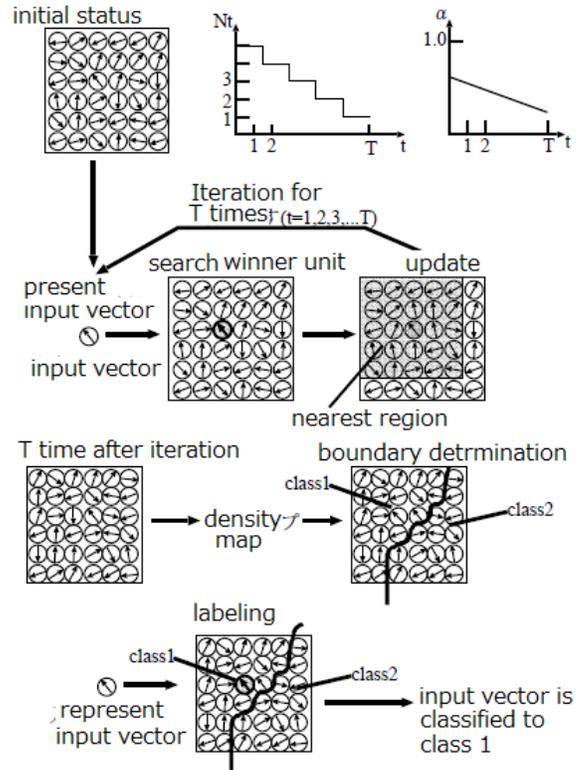


Fig.1. Illustrative view of the SOM learning process

Meanwhile $m_i(x(t))$ is updated as follows,

$$m_i(x(t+1))=m_i(x(t))+\lambda I(x(t)) \|x(t)-m_i(x(t))\| \quad (8)$$

It is because of the following equation.

$$\partial J/\partial m_i(x(t))=-2\sum I(x(t)) \|x(t)-m_i(x(t))\| \quad (9)$$

The k-means clustering algorithm can be rewritten as follows,

- (1) Set initial status of mean vectors of k clusters, $m_i(x(0))$, $i=1,2,\dots,k$, then
- (2) Iteration of the following two steps for $t=k+1, k+2,\dots,N$,

$$I_i(x(t))=1, \text{ when } \|x(t)-m_i(x(t))\|\leq\|x(t)-m_j(x(t))\| \forall j \\ =0, \text{ elsewhere} \quad (10)$$

$$m_i(x(t+1))=m_i(x(t))+I(x(t)) \|x(t)-m_i(x(t))\| / \sum_{t'=1}^t I(x(t')) \quad (11)$$

The equation (11) is identical to the equation (8) if λ is replaced to $1/\sum I(x(t'))$.

C. Density Map

The difference of input data is enhanced in the output layer unit through SOM learning so that similar code vector of the unit becomes formed.

⁴ <http://books.google.co.jp/books?id=AuY1PwAACAAJ&dq=cost+function&hl=ja&sa=X&ei=6ynYT6DxA8rxmAXQlsGNAw&ved=0CDUQ6AEwAA>
⁵ http://books.google.co.jp/books?id=WonHHAAACAAJ&dq=k-means+clustering&hl=ja&sa=X&ei=hirYT_DvF8PJmQWX8KGgAw&ved=0CD4Q6AEwAQ

Meanwhile, if the similar input data are separated in their location each other, it becomes neighboring units in the output layer unit. Density map $f(j,k)$ is defined as follows,

$$f(j,k) = \sum_{(l,n) \in D} (m_{j,k} - m_{j-l,k-n})^T (m_{j,k} - m_{j-l,k-n}) / D \quad (12)$$

Where D is neighboring unit, 8 neighbor unit centered the unit in concern in this paper. This density map has the relation among the input imagery data, feature space and SOM learning process as is illustrated in the Figure 2.

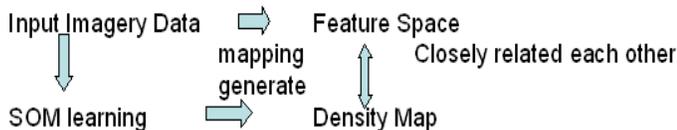


Fig.2. Relations among the input imagery data, feature space and density map generated through SOM learning.

D. Example of Density Map

This is an inverse function of the similar data concentration so that the density map obtained by a SOM learning process is quite similar to the distribution in the feature space mapped from the input data. An example of density map is illustrated in the Figure 3. In the figure, dark portion means dense of code vector meanwhile light portion is sparse of code vector and becomes boundary between the different clusters.

Figure 4 shows a preliminary result of density map, binarized density map and clustering result with increasing of the iteration number. In this case, initial variances of the two clusters are set at 0.03. In accordance with the number of iteration, density map becomes clear together with binarized density map. Furthermore, cluster result becomes ideal goal.



Fig.3. Example of density map as a result of SOM learning process.

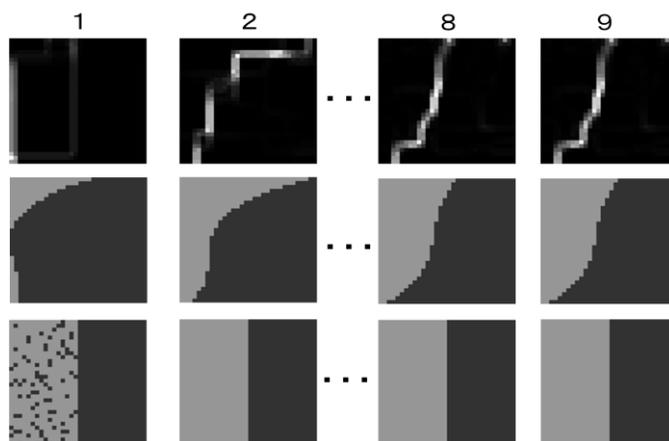


Fig.4. Example of preliminary result of density map, binarized density map and clustering result with increasing of the iteration number (multiplied by 512).

III. SIMULATION STUDIES AND THE EXPERIMENTS WITH REMOTE SENSING SATELLITE IMAGERY DATA

Simulation studies and experiments with real remote sensing satellite imagery data are conducted. In the simulation study, binarized density map is used for determination of boundary between two clusters while clusters are identified from the density map including candidate sub-clusters as shown in Figure 5.

A. Simulation Studies

Two clusters and two bands of imagery data are assumed. 32 by 32 pixels of 1024 of imagery data is created with random number generator of Messene Twister. 30 of image patterns are generated for simulation studies. 10 trials are conducted for 30 of image patterns. The maximum iteration number is set at 150,000.

Three different types of image pattern of datasets are prepared depending on separability as shown in Table 1. First dataset is very easy to separate while second dataset is a little bit difficult to separate. The third dataset is totally difficult to separate.

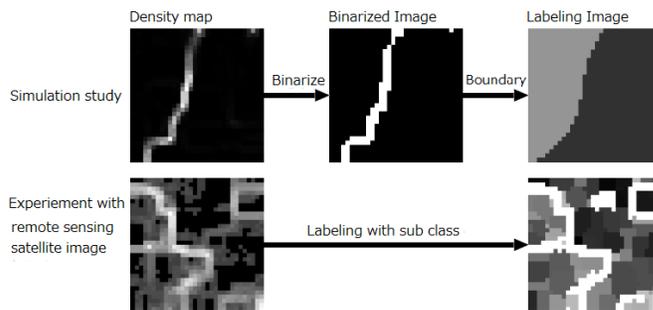


Fig.5. Methods for pixel labeling for simulation study and the experiments with remote sensing satellite images

TABLE I. PARAMETERS FOR SIMULATION DATA

Variance, σ	Distance between clusters
0.03	8σ
0.04	4σ
0.05	3σ

Two dimensional probability density function for No.1 dataset is shown in Figure 6 (a) while that for No.2 is shown in Figure 6 (b). Meanwhile, two dimensional probability density function for No.3 dataset is shown in Figure 6 (c).

Code vector distributions at the certain iteration numbers are shown in Figure 7 while density maps, boundary images and labeling result images for the certain iteration numbers are shown in Figure 8. These two figures are for No.1 of image dataset.

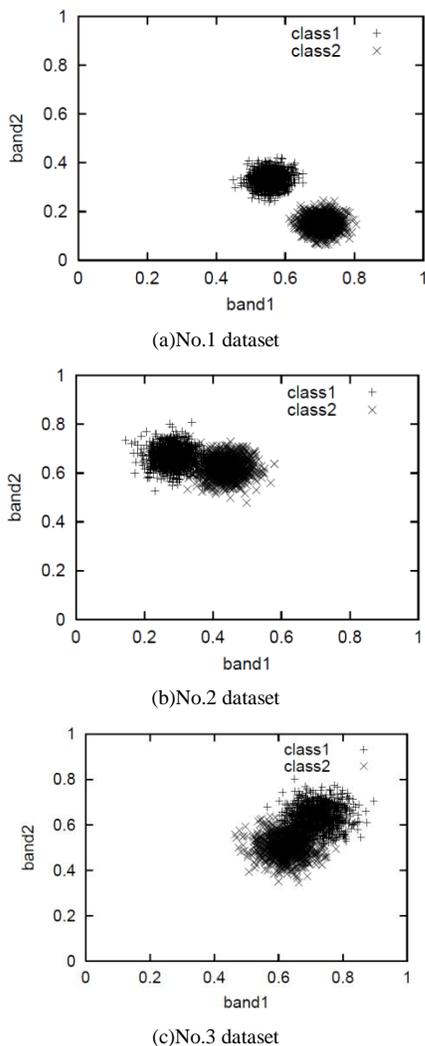


Fig.6. Two dimensional probability density function for No.1, 2, 3 dataset

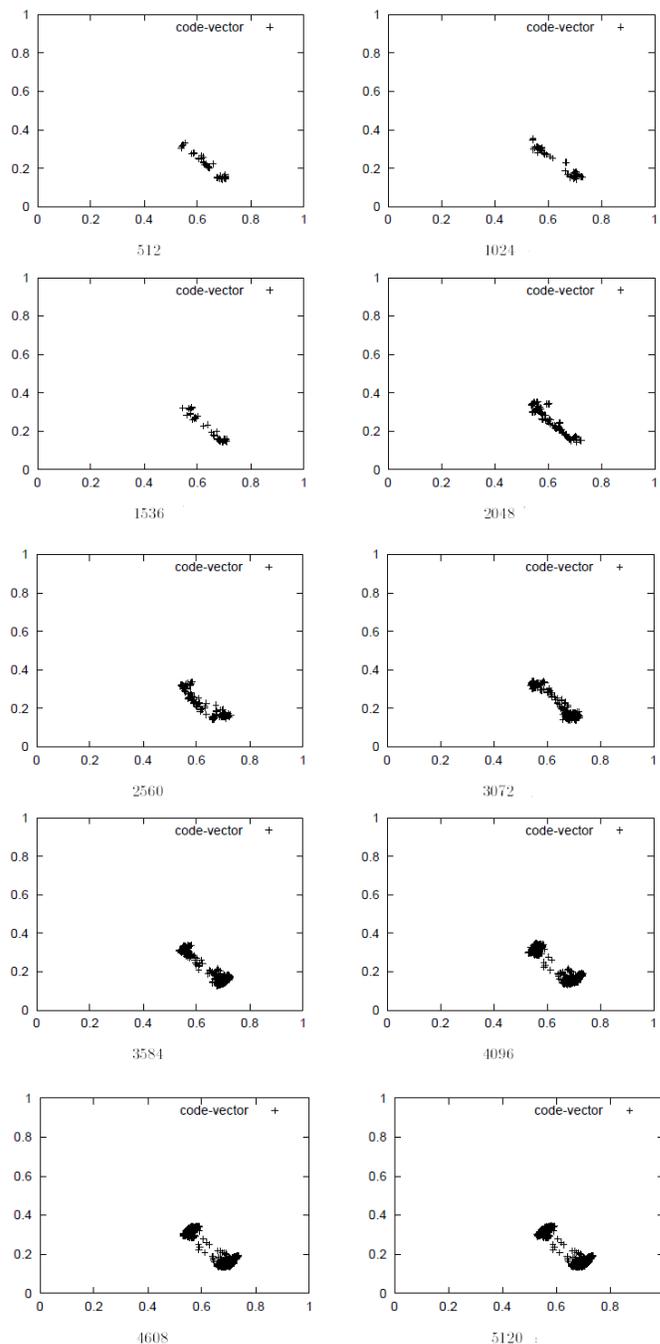


Fig.7. Code vector distributions at the certain iteration numbers for the dataset No.1.

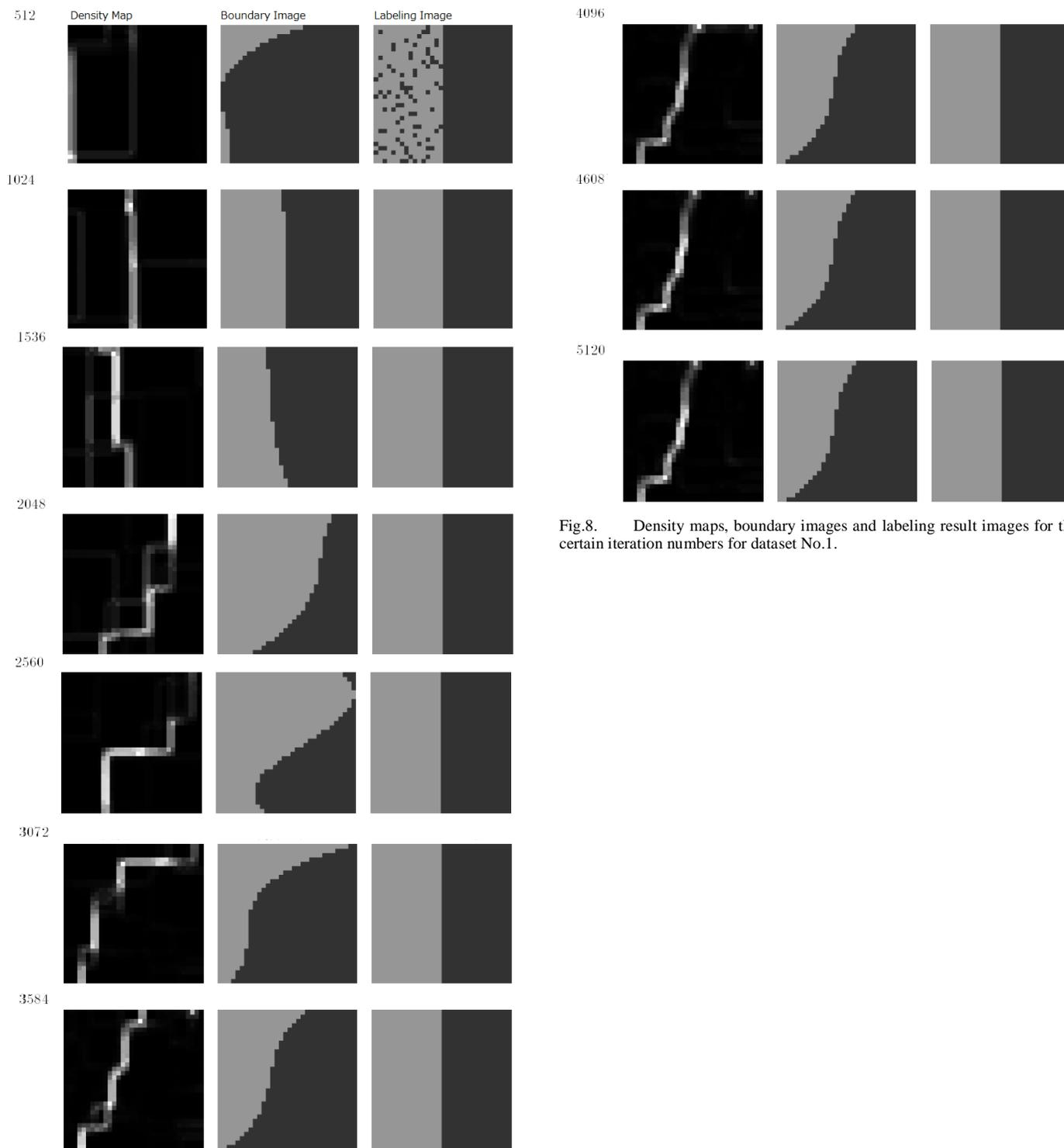


Fig.8. Density maps, boundary images and labeling result images for the certain iteration numbers for dataset No.1.

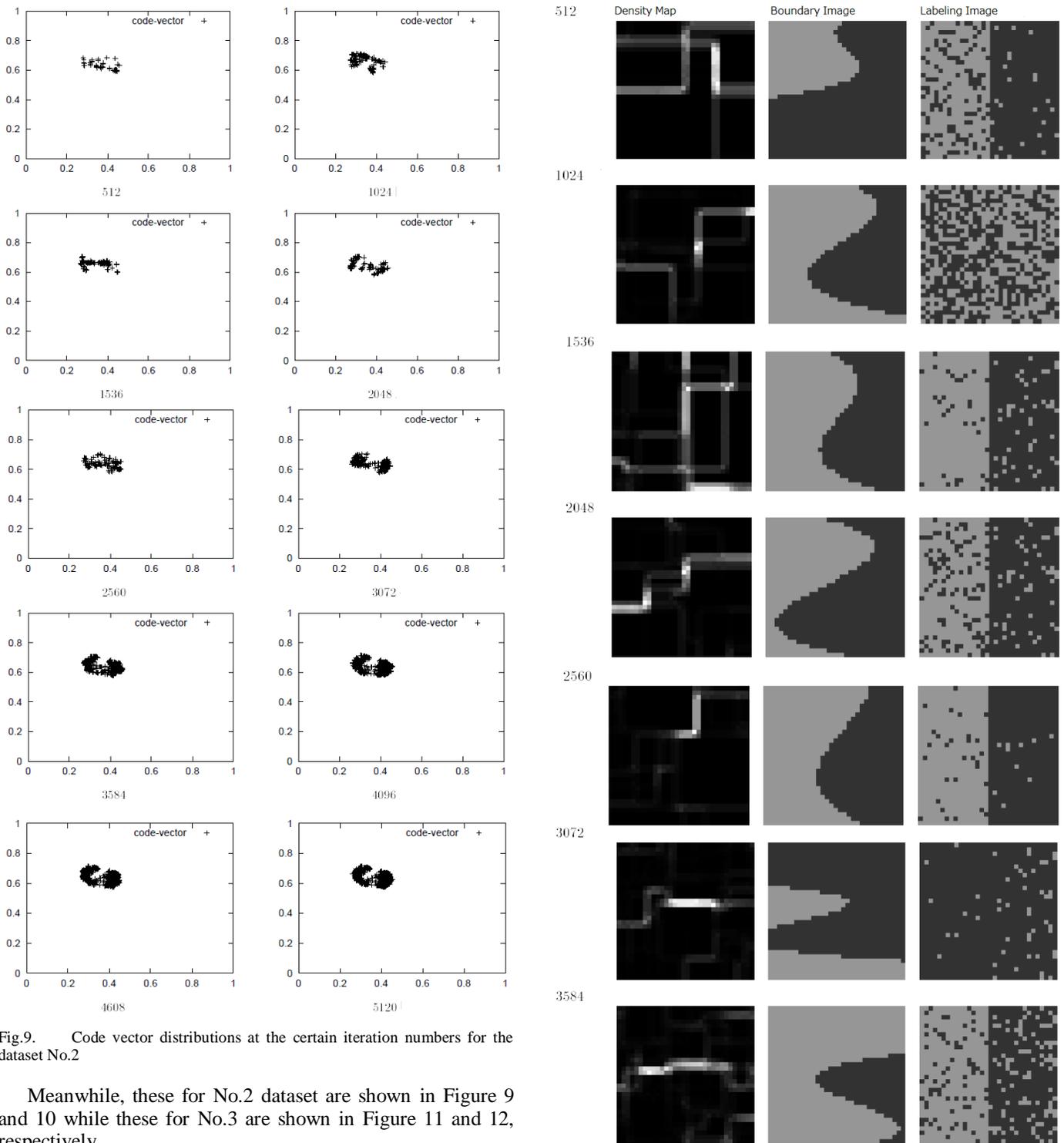


Fig.9. Code vector distributions at the certain iteration numbers for the dataset No.2

Meanwhile, these for No.2 dataset are shown in Figure 9 and 10 while these for No.3 are shown in Figure 11 and 12, respectively.

The code vector changes gradually and becomes separable situations. Density map, boundary image and labeling result image gets better and worth in accordance with increasing of iteration number and gradually become separable situation. These trends are different among the dataset No.1, 2, and 3.

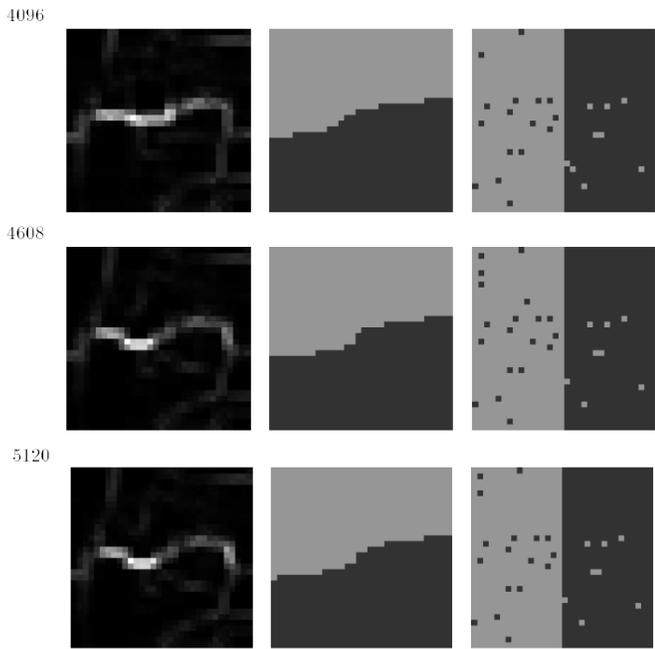


Fig.10. Density maps, boundary images and labeling result images for the certain iteration numbers for dataset No.2.

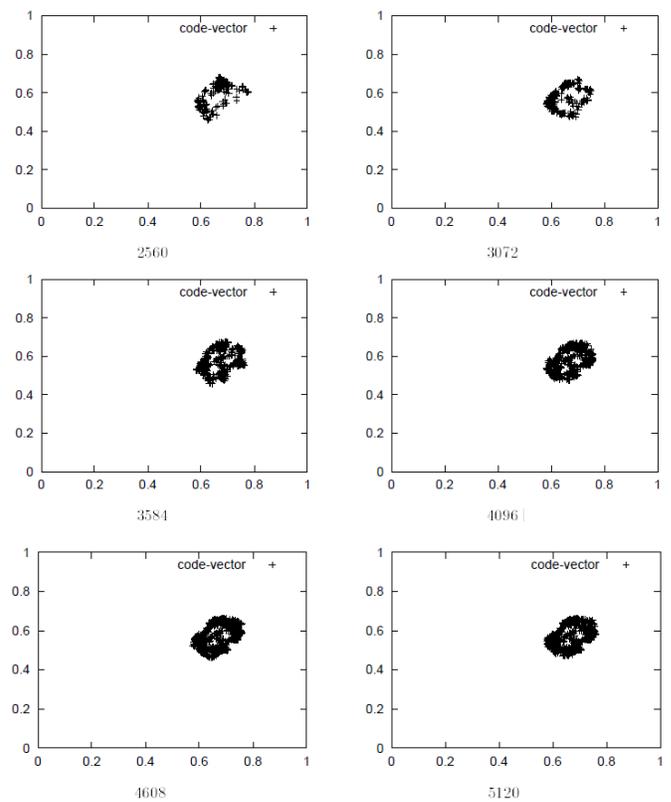
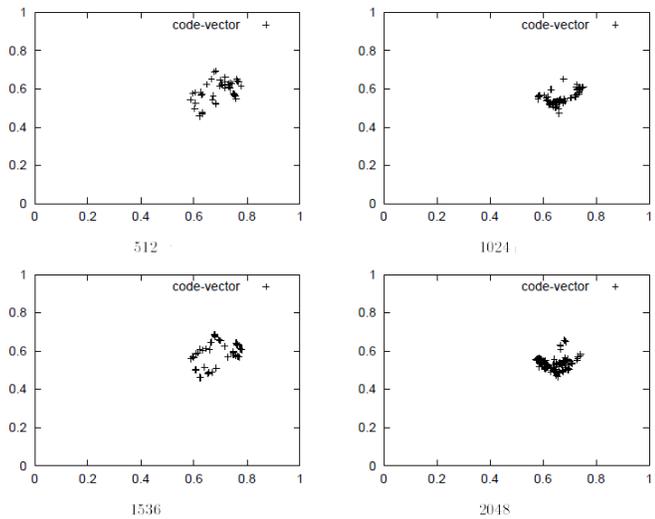


Fig.11. Code vector distributions at the certain iteration numbers for the dataset No.3

Difficulty of the convergence (converged iteration numbers) depends on within cluster variance and between cluster variance obviously. For instance, it is converged so quickly for image dataset No.1 while convergence speed is not fast for the image dataset No.2.

Meanwhile, convergence is very slow for the image dataset No.3. It is not converged yet for the iteration number is 5420 for the image datasets No.2 and 3. The relation between the iteration number and correct clustering ratio is shown in Figure 13.

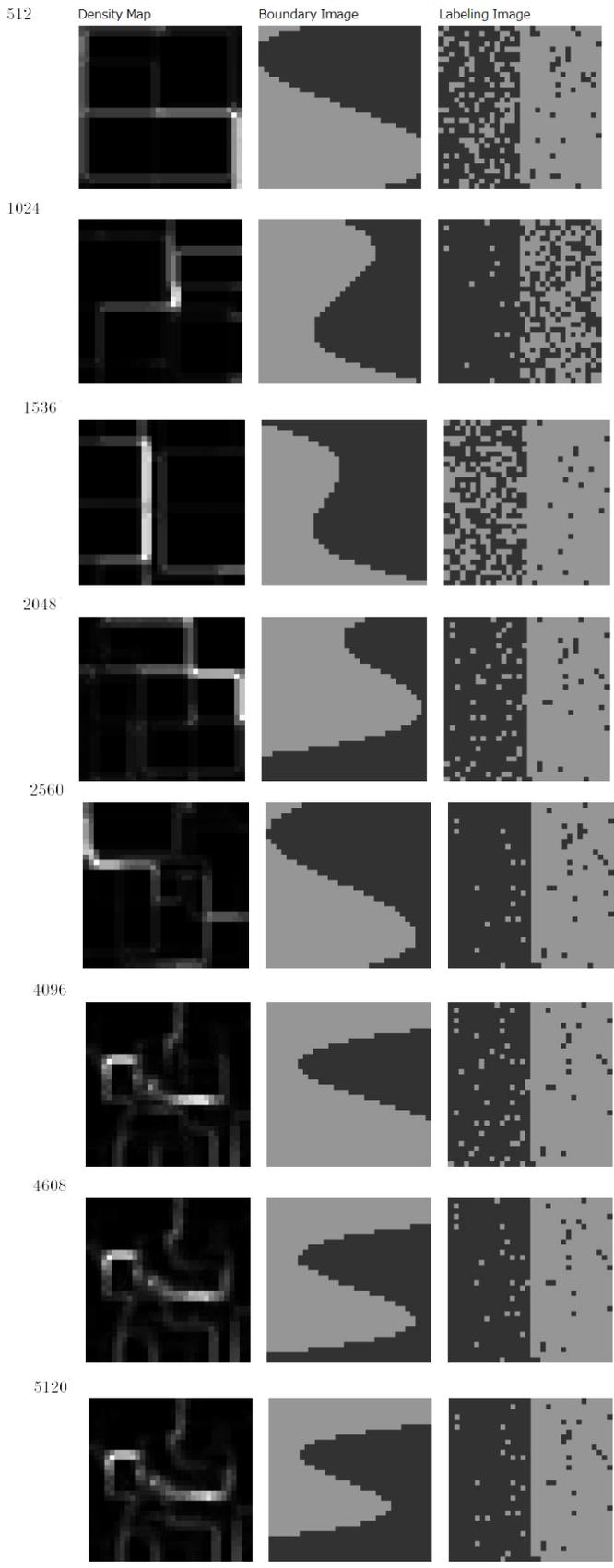


Fig.12. Density maps, boundary images and labeling result images for the certain iteration numbers for dataset No.3.

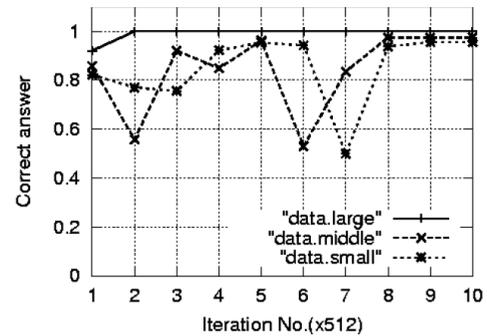


Fig.13. Correct clustering ratios as function of iteration number.

B. Experimental Studies with Remote Sensing Satellite Imagery Data

Landsat-5 TM data of Saga, Japan acquired on 15 May 1987 which is shown in Figure 14 is used. The meta data is as follows, Entity ID: LT51130371987135HAJ00, Acquisition Date: 15-MAY-87, Path: 113, Row: 37. A portion of band 1-5 and 7 of Landsat-5 TM image is shown in Figure 14 together with the topographic map of the corresponding area.

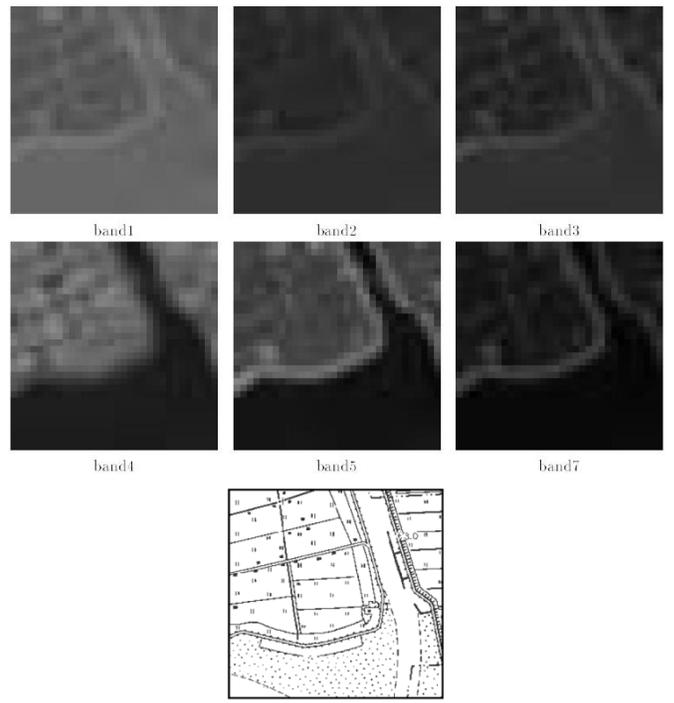


Fig.14. Landsat-5 TM data of Saga, Japan acquired on 15 May 1987 and topographic map of the corresponding area

Separability (ratio of between cluster variance to within cluster variance) is getting large in accordance with iteration number as shown in Figure 15. Density maps and labeling results are shown in Figure 16 in accordance with iteration number.

Figure 17 shows the density map and the labeling result image after the convergence. In accordance with iteration number, the labeling result image is getting resemble to the topographic map

Confusion matrix of the proposed SOM based clustering is compared to that of Maximum Likelihood: MLH classification as shown in Table 2.

In particular, bare soil tends to be clustered to water body. This fact can be confirmed in Figure 18.

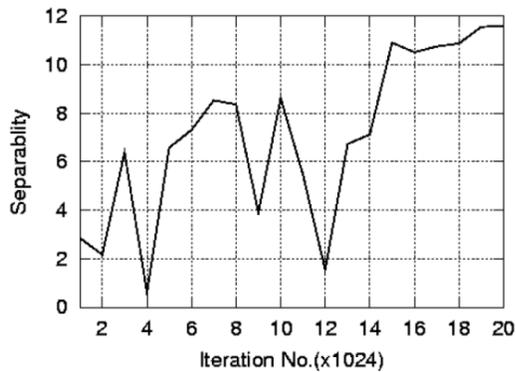


Fig.15. Relation between separability and iteration number

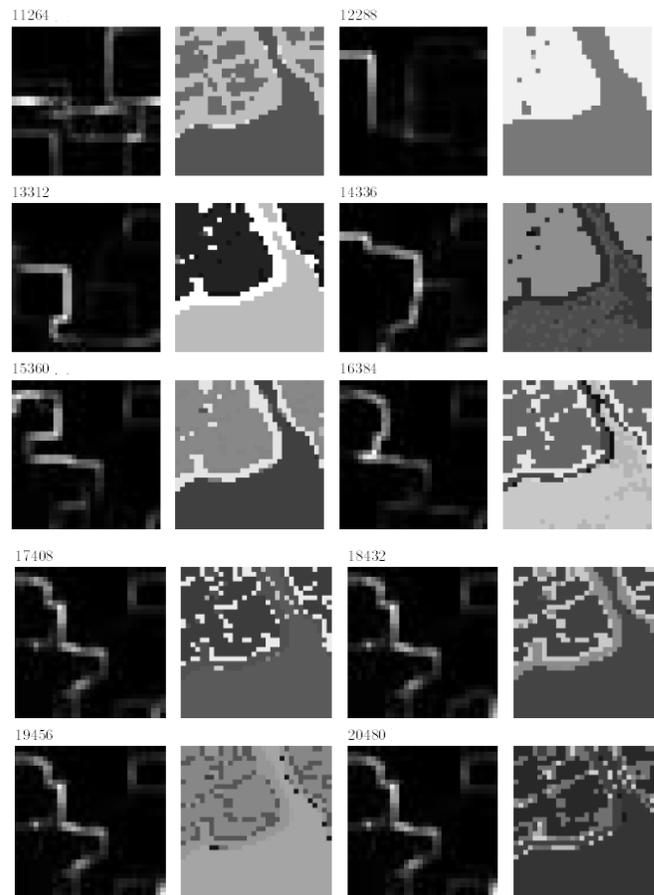


Fig.16. Density maps and labeling results

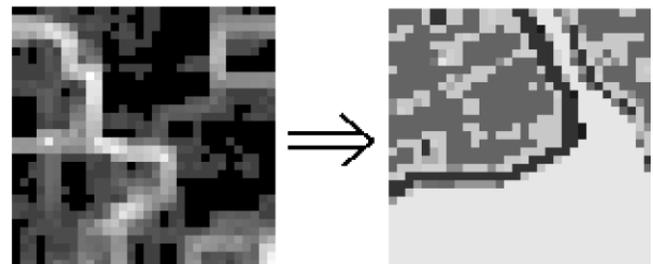
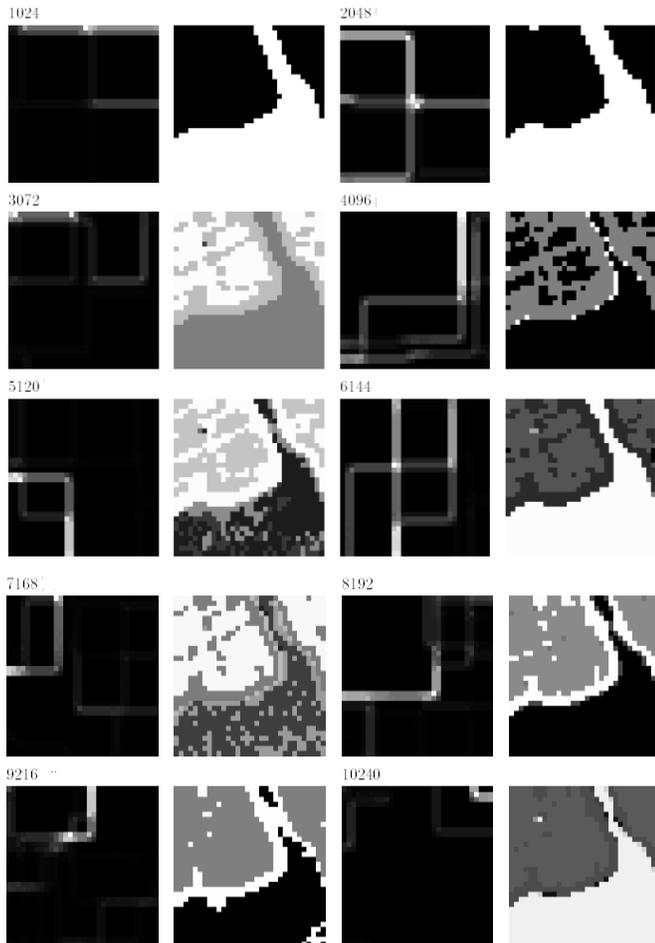


Fig.17. The density map and the labeling result image after the SOM learning process is converged.

TABLE II. COMPARISON OF CLUSTERED RESULT BY THE PROPOSED SOM BASED METHOD AND CLASSIFICATION RESULTS BY THE CONVENTIONAL MLH METHOD

		SOM				
		structure	road	paddy	soil	water
MLH	structure	94%	4%	0%	0%	0%
	road	1%	94%	5%	0%	0%
	paddy	0%	8%	92%	0%	0%
	soil	3%	0%	0%	64%	33%
	water	0%	0%	0%	0%	100%

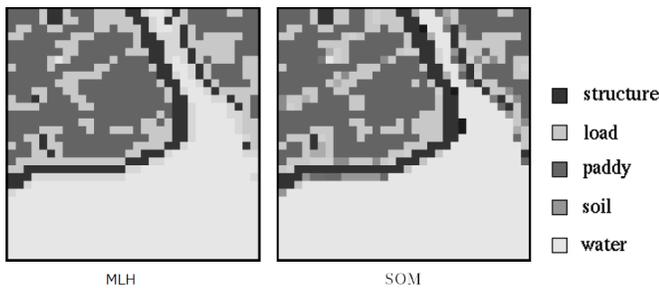


Fig.18. Clustered result by the proposed SOM based clustering and classified result by the conventional MLH method

IV. CONCLUSION

A new method for image clustering with density maps derived from Self-Organizing Maps (SOM) is proposed together with a clarification of learning processes during a construction of clusters. Simulation studies and the experiments with remote sensing satellite derived imagery data are conducted. It is found that the proposed SOM based image clustering method shows much better clustered result for both simulation and real satellite imagery data. It is also found that the separability among clusters of the proposed method is 16% longer than the existing k-mean clustering. It is also found that the separability among clusters of the proposed method is 16% longer than the existing k-mean clustering. In accordance with the experimental results with Landsat-5 TM image, it takes more than 20000 of iteration for convergence of the SOM learning processes

ACKNOWLEDGMENT

The author would like to thank Mr. Koichi Tateno for his effort to conduct simulation studies and the experiments with Landsat-5 TM data.

REFERENCES

[1] Mikio Takagi and Haruhisa Shimoda Edt. Kohei Arai et al., Image Analysis Handbook, The University of Tokyo Publishing Co. Ltd., 1991.,

[2] Kohei Arai, Fundamental Theory for Image Processing Algorithms, Gakujutu-Tosho-Publishing Co. Ltd., 1999.
[3] Kohei Arai, Fundamental Theory for Pattern Recognitions, Gakujutu-Tosho-Publishing Co. Ltd., 1999.
[4] Kohei Arai, Remote Sensing Satellite Image Processing and Analysis, Morikita Publishing Inc., 2001.
[5] T.Kohonen, Self-Organizing Maps, Springer Series in Information Sciences, Vol.30, 1995; Second edition, 1997; Third, extended edition, 2001.
[6] T.Kohonen, G.Barna and R.Chrisley, Statistical pattern recognition with neural networks: benchmarking studies, Proc. ICNN Vol.I, 61-68, 1988.
[7] Kohei Arai, Learning processes of image clustering method with density maps derived from Self-Organizing Mapping(SOM), Journal of Japan Photogrammetry and Remote Sensing, 43, 5, 62-67, 2004.
[8] Kohei Arai, XiangQiang Bu, Pursuit Reinforcement Learning based on-line clustering for image retrievals, Journal of Image Electronics and Engineering Society of Japan, 39,3,301-309,2010
[9] Kohei Arai, XiangQiang Bu, Pursuit Reinforcement Learning based on-line clustering with learning automaton for rescue simulations and its acceleration of convergence of learning processes, Journal of Image Electronics and Engineering Society of Japan, 40, 2, 361-168, 2011.
[10] Jouko Lampinen and Timo Kostiainen, Generative probability density model in the Self-Organizing Map. In U. Seiffert and L. Jain, editors, *Self-organizing neural networks: Recent advances and applications*. pages 75-94. Physica Verlag, 2002..

AUTHORS PROFILE

Kohei Arai He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission-A of ICSU/COSPAR since 2008. He wrote 30 books and published 332 journal papers.

Remote Sensing Satellite Image Database System Allowing Image Portion Retrievals Utilizing Principal Component Which Consists Spectral and Spatial Features Extracted from Imagery Data

Kohei Arai ¹

Graduate School of Science and Engineering
Saga University
Saga City, Japan

Abstract—Remote Sensing satellite image database which allows image portion retrievals utilizing principal component which consists of spectral and spatial features extracted from the imagery data is proposed. Through the experiments with actual remote sensing satellite images, it is found that the proposed image retrieval does work so well.

Keywords—Image retrieval; Spectral and spatial features; e

I. INTRODUCTION

In accordance with expansion in the multimedia technologies and the Internet, CBIR has been an active research topic since the first 1990's. The concept of content based retrieval (CBR) in image start from the first 1980s and Serious applications started in the first 1990s [1]. Retrieval from databases with a large number of images has attracted considerable attention from the computer vision and pattern recognition society.

Brahmi et al. mentioned the two drawbacks in the keyword annotation image retrieval. First, images are not always annotated and the manual annotation expensive also time consuming. Second, human annotation is not objective the same image may be annotated differently by different observers [2]. Unlike the traditional approach that using the keyword annotation as a method to search images, CBIR system performs retrieval based on the similarity feature vector of color, texture, shape and other image content. Comparing to the traditional systems, the CBIR systems perform retrieval more objectiveness [3].

Global features related to color or texture are commonly used to describe the image content in image retrieval. The problem using global features is this method cannot capture all parts of the image having different characteristics [4]. In order to capture specific parts of the image the local feature is used. The proposed method uses 2D Discrete wavelet transform with Haar base function, combined the two high sub-band frequency to make significant points and edge, choose any part of image by threshold the high coefficient value.

Image feature based image portion retrieval method is proposed in this paper. Principal component which consists of spectral and spatial features extracted from the image are used as features.

The following section describes the proposed image database system followed by some experiments. Then conclusion is described together with some discussions.

II. PROPOSED SATELLITE IMAGE DATABASE SYSTEM

A. Image Database

Database system has to have the following three functions,

- 1) *Image coding and archiving with attributes information*
- 2) *Database operations including image retrievals and updating database*
- (3) Database management including maintenance

Attributes can be used for meta search while meta search allows image retrievals. The meta database is information of image information. Therefore, it is useful for keyword search. For image search, image coding is important. If the image coding is performed perfectly, then image retrievals can be done easily. On the other hands, there is image portion retrievals which allows image portion search in the image in the image database. In the image portion retrievals, some measures for representation of image features are required. One of the simple features is similarity between the template image in the image database and the current query image portions. Other than these, principal component is used as feature. The extracted spectral and spatial features can be projected in principal component space as a vector. The vector can be interpreted as a feature of the image portion in concern and can be used as measure of the similarity between template and the query image portion.

B. Image Retrieval

Two image retrieval methods are supported by the proposed image retrieval system. One is keyword based search utilizing the following keywords,

- 1) Image file name
- 2) Satellite name
- 3) Sensor name
- 4) Band number
- 5) Area name
- 6) Observation date
- 7) Observation time
- 8) Area feature
- 9) Cloud content

Another image retrieval method is image feature based method. There are utilizing features for image retrievals.

C. Image Feature

Average, variance, skewness, kurtosis, energy, entropy, the number of edges, minimum/maximum pixel values, contrast, digital count at most frequent pixel values are selected as features. After calculation of features, principal component analysis is applied to the features. The first three principal components are extracted then the mapped features are divided into groups as shown in Figure 1.

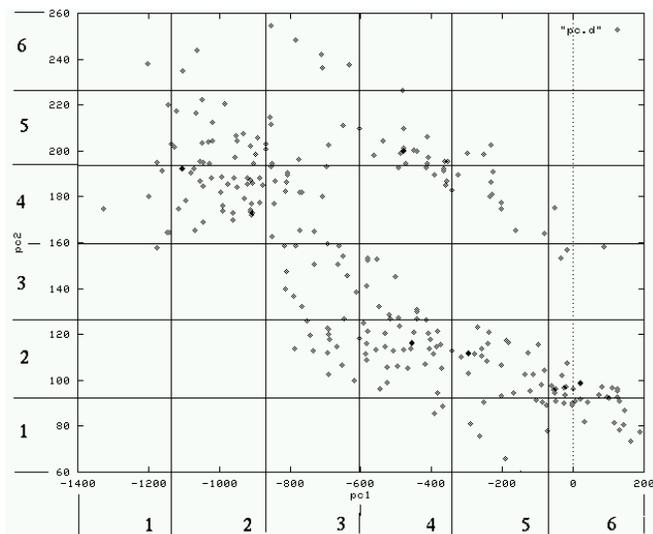


Fig. 1. The first three principal components are extracted then the mapped features are divided into groups

Thus the images in the image database are projected in the principal component space. Then the group number is used for one of attribute information.

D. Principal Component Analysis

Let $x_i, i=1, 2, \dots, p$ be features extracted from the image assuming correlation among the features. Then new variable, z is defined as follows,

$$z = \sum_{i=1}^p a_i x_i \quad (1)$$

Where a_i is determined as maximizing variance of z under

the condition of $1 = \sum_{i=1}^p a_i^2$. Such this z is called as the first

principal component. Also such these coefficients a_i are called as the first principal component vector. Rewriting the first principal component as z_j , and the first principal component vector a_i , then the arbitrary order, j of principal component z_j and principal component vector a_j can be written as follows,

$$a_j a_j^t = 1 \quad (2)$$

$$\sigma_i^2 \geq \sigma_j^2, i \geq j \quad (3)$$

$$\rho_{ij} = 0, i \neq j$$

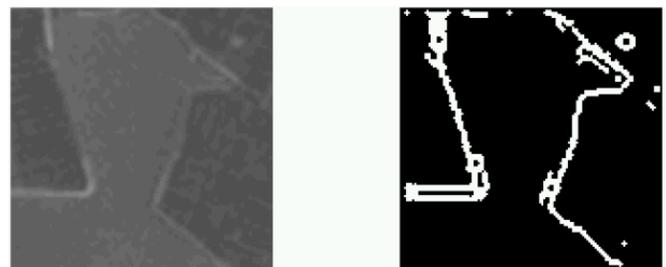
E. The Number of Edges

The following Sobel operator derived edges are extracted from the images for one of the features for image retrievals.

$$f_x : \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad f_y : \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

$$\begin{aligned} f_x(x, y) &= f(i+1, j+1) + 2f(i+1, j) \\ &\quad + f(i+1, j-1) - f(i-1, j+1) \\ &\quad - 2f(i-1, j) - f(i-1, j-1) \\ f_y(x, y) &= f(i-1, j-1) + 2f(i, j-1) \\ &\quad + f(i+1, j-1) - f(i-1, j+1) \\ &\quad - 2f(i, j+1) - f(i+1, j+1) \end{aligned} \quad (4)$$

Where f_x, f_y denotes pixel values of edge enhanced images in the horizontal and vertical directions. With appropriated threshold, the horizontal and vertical edges are extracted from the enhanced images. Figure 2 shows one of examples of edge detected images.



(a)Original image

(b)Edge detected image

Fig. 2. Edge detections

F. Texture Features

The following texture features, mean (μ) and variance (σ^2), skewness (SKW), kurtosis (KRT), energy (EGY), and entropy (ENT) are selected for representation of image features for search.

$$\begin{aligned} \mu &= \sum_{i=0}^{n-1} iP(i) \\ \sigma^2 &= \sum_{i=0}^{n-1} (i - \mu)^2 P(i) \\ SKW &= \frac{\sum_{i=0}^{n-1} (i - \mu)^3 P(i)}{\sigma^3} \\ KRT &= \frac{\sum_{i=0}^{n-1} (i - \mu)^4 P(i)}{\sigma^4} \\ EGY &= \sum_{i=0}^{n-1} P^2(i) \\ EPY &= -\sum_{i=0}^{n-1} P(i) \log P(i) \end{aligned} \quad (4)$$

Other than these, the following contrast is also calculated for image retrievals.

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \quad (5)$$

G. Principal Component for Image Retrievals

PCA component derived from satellite imagery data has different variance. In order to normalize the variance, the following equation is used,

$$x_{ik'} = \frac{x_{ik} - \bar{x}_i}{\sqrt{s_i}} \quad (i = 1, \dots, p), (k = 1, \dots, n) \quad (6)$$

Where x_{ij} denotes two dimensional imagery of pixel data while s_i denotes variance of the pixel data of band i . Correlation coefficient between band i and j is expressed as follows,

$$R_{ij} = \frac{s_{ij}}{\sqrt{s_i s_j}}$$

$$s_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) \quad (7)$$

Using the correlation coefficients, eigen value and eigen vector of correlation coefficient is estimated utilizing Jacob method. Let X, Λ be eigen vector and eigen value of A . Non singular matrix M is utilized, then,

$$\begin{aligned} AX &= \Lambda X \\ MAX &= \Lambda MX \\ y &= MX \\ MAM^{-1}y &= \Lambda y \end{aligned} \quad (8)$$

Thus it is found that eigen value and eigen vector is not going to be changed with the aforementioned operation with the non singular matrix, M . The followings are candidates of M ,

$$m_{ij} = \begin{cases} \cos \theta & (i = j = p, i = j = q) \\ \sin \theta & (i = p, j = q) \\ -\sin \theta & (i = q, j = p) \\ 0 & (i \neq j) \\ 1 & (i = j) \end{cases} \quad (9)$$

These are called as basic orthogonal matrices. Using the basic orthogonal matrix, power of A can be written as follows,

$$A^{(s+1)} = M' A^{(s)} M \quad (s = 1, 2, 3, \dots) \quad (10)$$

Namely,

$$\begin{aligned} a_{ij}^{(s+1)} &= a_{ij}^{(s)} \cos \theta + a_{ik}^{(s)} \sin \theta \\ a_{ik}^{(s+1)} &= -a_{ij}^{(s)} \sin \theta + a_{ik}^{(s)} \cos \theta \\ a_{jj}^{(s+1)} &= a_{jj}^{(s)} \cos^2 \theta + 2a_{jk}^{(s)} \cos \theta \sin \theta + a_{kk}^{(s)} \sin^2 \theta \\ a_{kk}^{(s+1)} &= a_{jj}^{(s)} \sin^2 \theta - 2a_{jk}^{(s)} \cos \theta \sin \theta + a_{kk}^{(s)} \cos^2 \theta \\ a_{jk}^{(s+1)} &= (a_{kk}^{(s)} - a_{jj}^{(s)}) \cos \theta \sin \theta + a_{jk}^{(s)} (\cos^2 \theta - \sin^2 \theta) \end{aligned} \quad (11)$$

If the maximum absolute value of the non diagonal elements, $a_{ij}^{(s)}$ is found, then

$$\begin{aligned} (a_{kk}^{(s)} - a_{jj}^{(s)}) \cos \theta \sin \theta + a_{jk}^{(s)} (\cos^2 \theta - \sin^2 \theta) &= 0 \\ (a_{jj}^{(s)} - a_{kk}^{(s)}) \sin 2\theta &= 2a_{jk}^{(s)} \cos 2\theta \\ \tan 2\theta &= \frac{2a_{jk}^{(s)}}{(a_{jj}^{(s)} - a_{kk}^{(s)})} \quad \left(-\frac{\pi}{4} \leq \theta \leq \frac{\pi}{4}\right) \end{aligned} \quad (12)$$

It is possible to find the θ as to equation (12) is satisfied. Then,

$$A^{(s)} = M_s' \dots M_1' A M_1 \dots M_s \quad (13)$$

Thus the eigen values are obtained,

$$A^{(s)} \rightarrow \Lambda \quad (14)$$

If we choose M_i as follows,

$$M = M_1 \dots M_s$$

Then,

$$\begin{aligned} M' A M &= \Lambda \\ A M &= M \Lambda \end{aligned}$$

Thus eigen values are determined followed by the corresponding eigen vector.

H. Implementation of Image Retrievals

Web design for image retrievals is conducted. Figure 3 shows the registration form. The aforementioned key words for key word image retrievals can be input to the database.



Fig. 3. Registration for which allows input keywords through dialog boxes

The registered image with attribution data can be confirmed as shown in Figure 4. Figure 5 shows top page of the image retrieval system. If user input query information through the web page which is shown in Figure 6, then search results can be obtained from the image retrieval system as shown in Figure 7. Keyword search can be done through the web page of Figure 6.

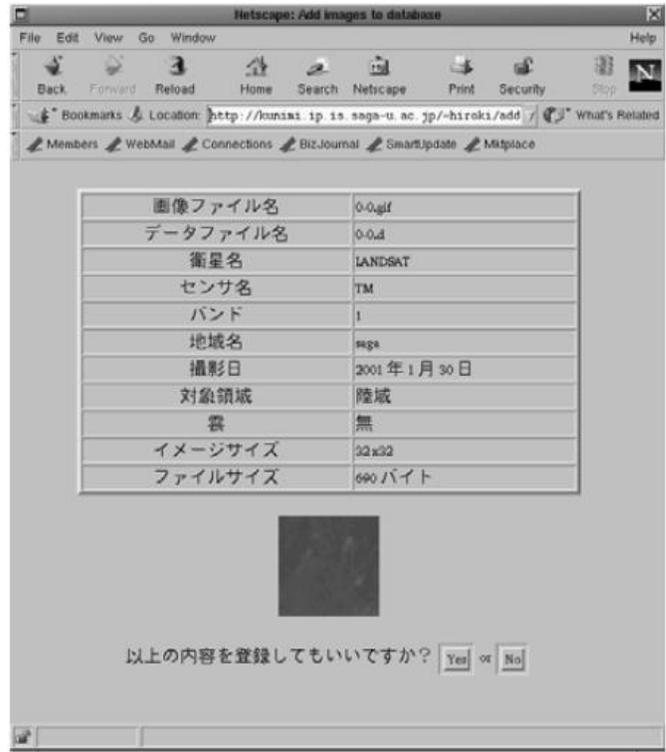


Fig. 4. Confirmation of the registered image and attribution data

Image feature based search is also done through the web page of Figure 8. Then the search results can be obtained as shown in Figure 9. Also the attribution information is obtained as shown in Figure 10.

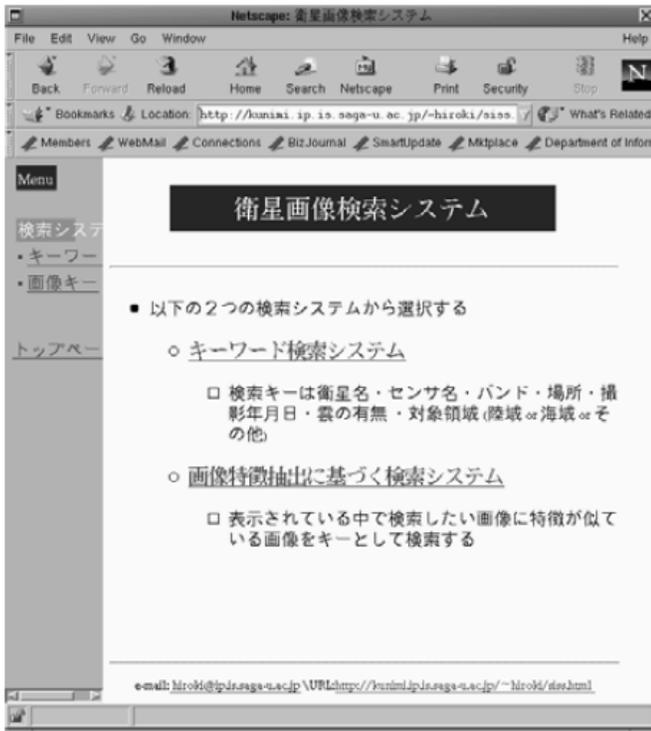


Fig. 5. page of the image retrieval system

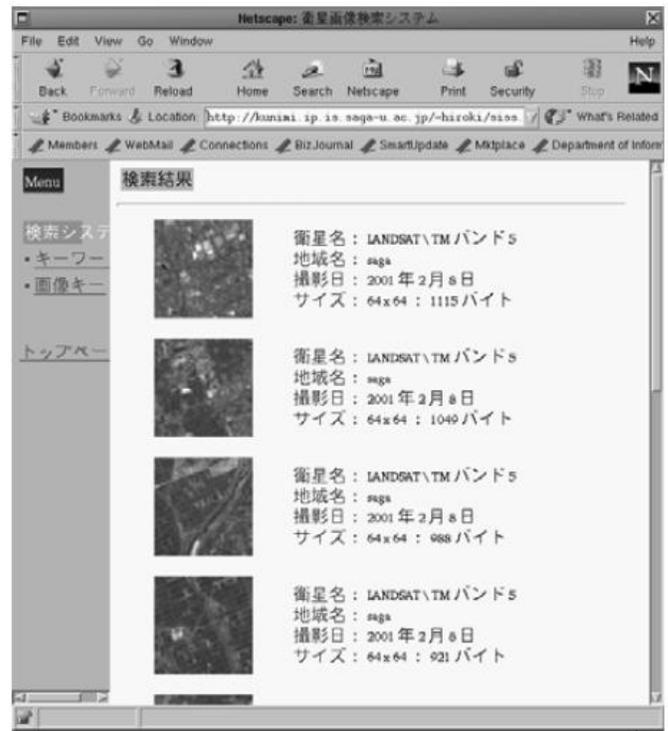


Fig. 7. Retrieved image are obtained as candidates



Fig. 6. Query information can be input through the web page

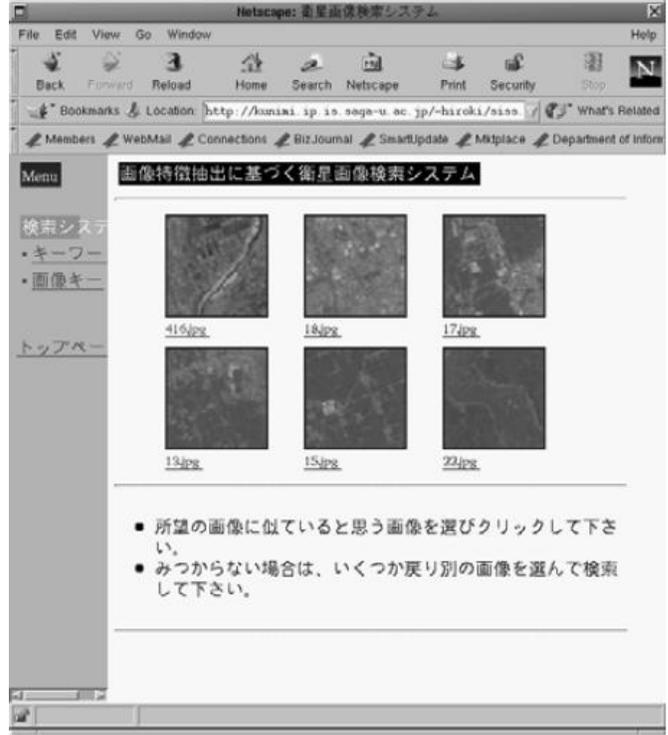


Fig. 8. Image feature based search is also done through the web page.

III. EXPERIMENTS

A. A. Data Used and Evaluation of Hit Ratio and the Number of Steps for Search

Landsat Thematic Mapper: TM data are used for experiments. From the TM data, 288 of 64 by 64 pixels of sample images with the different types of image features are selected.

12 students are nominated for performance tests of hit ratio and the required number of steps for image retrievals.

B. B. Image Retrieved Results

The examples of the first principal component images (the largest eigen vector which corresponds to the largest eigen value) are shown in Figure 11.



Fig. 9. Search results can be obtained

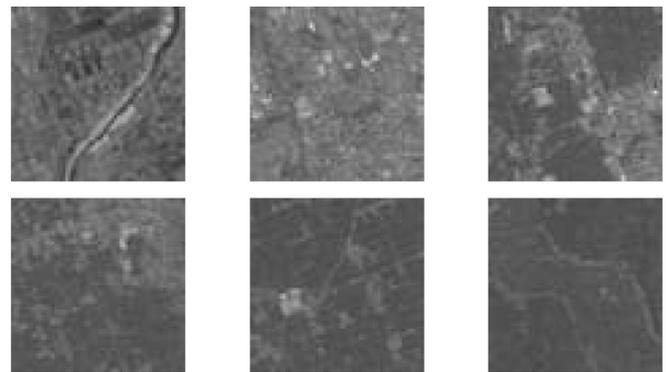


Fig. 11. Examples of the first principal component images (the largest eigen vector which corresponds to the largest eigen value)

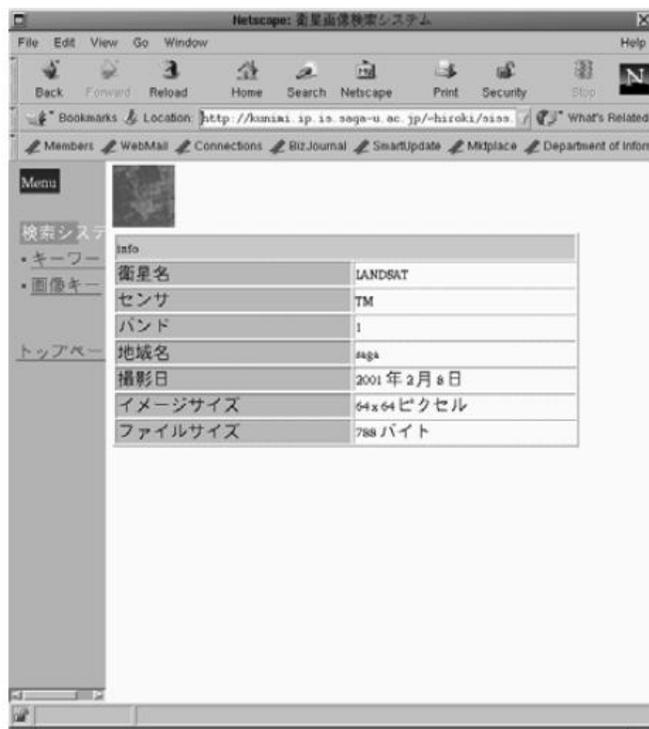


Fig. 10. Attribution information is obtained

Image feature based image retrieval is conducted. If user clicks one of the displayed image of Figure 11, then the retrieved image candidates are displayed in accordance with the distance between the query image and the candidate images as shown in Figure 12. Figure 12 shows the example images extracted from the first group of the first principal component images. It is found that image complexity can be expressed with the first principal component images.

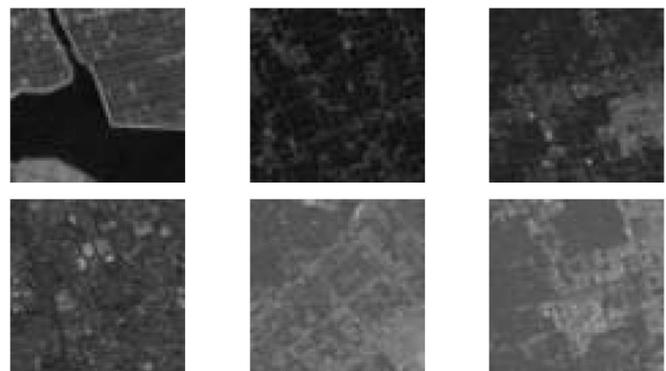


Fig. 12. Example images extracted from the first group of the first principal component images.

C. Evaluation of Hit Ratio and the Required Number of Steps for Search

By using the query images of Figure 13, image feature based image retrievals are conducted by 12 students who do not have any experience and knowledge about satellite imagery data at all. The evaluation results are shown in Table 1.

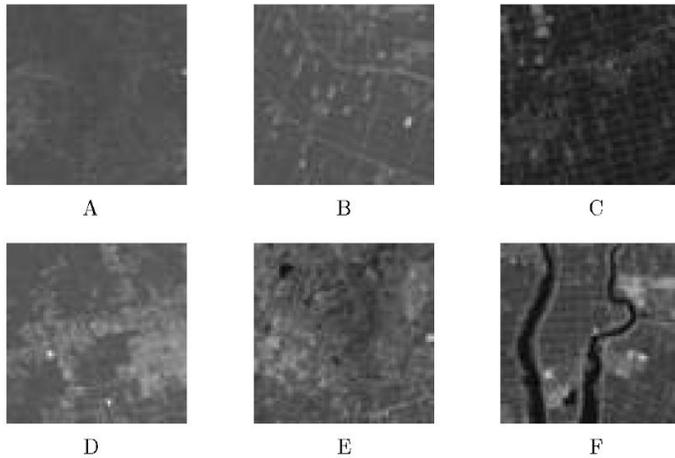


Fig. 13. Query images which are selected from the 288 of sample Landsat TM images in the image database.

TABLE I. HIT RATIO AND THE REQUIRED NUMBER OF STEPS FOR SEARCH WHICH ARE EVALUATED WITH LAND SAT SATELLITE TM IMAGERY DATA BY 12 STUDENTS.

	A	B	C	D	E	F
Hit Ratio (%)	54.5	36.4	18.2	72.7	9.1	100
No.of Steps	5.7	12	21	6.4	3	4.1

Hit ratio and the required number of steps for search are totally depending on the complexity of the query image. The highest hit ratio is shown for the query image “F” while the lowest hit ratio is shown for the query image of “E”. Therefore it is said that image feature rich image of “F” is easy to retrieve while image feature poor image of “E” is difficult for search. On the other hands, the number of steps required for image retrievals for the query image “C” is greatest and that for the query image “E” is smallest. This implies that it takes long time to search the query image “C”, at the same time, not so large hit ratio cannot be achieved for such query image. Meanwhile, user tried to search images for the query image “E”, and then user quit to search because it is difficult to

search. Therefore, not only hit ratio but also the number of steps required for search is poor.

IV. CONCLUSION

Remote Sensing satellite image database which allows image portion retrievals utilizing principal component which consists of spectral and spatial features extracted from the imagery data is proposed. Through the experiments with actual remote sensing satellite images, it is found that the proposed image retrieval does work so well.

Experimental results show that hit ratio depends on the image features which are containing in the satellite image in concern. The required number of steps does not depend on the image features which are contained in the satellite image in concern. Because user forgets about the image retrievals when they feel that it is difficult to search due to the fact that satellite image in concern does not have any specific features.

ACKNOWLEDGMENT

The author would like to thank Mr.Hiroki Fujikawa for his effort to conduct experiments.

REFERENCES

- [1] Henning M'uller, Antoine Geissbuhler, et al. "Benchmarking Image Retrieval Applications". San Francisco, CA: In The Seventh International Conference, September 2004. pp. 334-337.
- [2] Kherfi, M., Brahmi, D. and Ziou D. "Combining Visual Features with Semantics for a More Effective Image Retrieval". : in ICPR '04, vol. 2, 2004. pp. 961-964.
- [3] H. Yu, M. Li, H.-J. Zhang and Feng, J. "Color texture moments for content-based image retrieval". : In International Conference on Image Processing, 2002. pp. 24-28.
- [4] N. Sebe, Q. Tian, E. Loupias, M. Lew and T. Huang. "Evaluation of salient point techniques". s.l. : In International Conference, 2004.

AUTHORS PROFILE

Kohei Aarai He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission-A of ICSU/COSPAR since 2008. He wrote 30 books and published 332 journal papers.

Mashup Based Content Search Engine for Mobile Devices

Kohei Arai¹

Graduate School of Science and Engineering
Saga University
Saga City, Japan

Abstract—Mashup based content search engine for mobile devices is proposed. Example of the proposed search engine is implemented with Yahoo!JAPAN Web SearchAPI, Yahoo!JAPAN Image searchAPI, YouTube Data API, and Amazon Product Advertising API. The retrieved results are also merged and linked each other. Therefore, the different types of contents can be referred once an e-learning content is retrieved. The implemented search engine is evaluated with 20 students. The results show usefulness and effectiveness on e-learning content searches with a variety of content types, image, document, PDF files, moving picture.

Keywords—search engine; API; mashup; Andoride; e-learning content retrieval;

I. INTRODUCTION

Mashup technology is defined as search engine with plural different APIs. Mashup has not only the plural APIs, but also the following specific features.

- 1) it enables classifications of the contents in concern by using web 2.0,
- 2) it may use API from the different sites,
- 3) it allows information retrievals from both sides of client and server,
- 4) it may search contents as an arbitrary structured hybrid contents which is mixed contents formed with the individual contents from the different sites,
- 5) it enabling to utilize REST, RSS, Atom, etc. which are formed from XML conversions.

There are some Mashup tools, Yahoo! Pipes¹, Microsoft Popfly², etc. while there are some services using Mashup technology, ChaMap –Enjoy Geo Communication!³, Newsgraphy⁴, Flowser on Amazon⁵

Although Mashup allows content search which is same as portal, Mashup has the aforementioned different features from portal. Therefore, Mashup is possible to create more flexible search engine for any purposes of content retrievals. The search system which is proposed here is that make it possible to control the graph in the 3D space display with these

peculiarity on Android devices. Mashup technology utilized search engine for e-learning content retrievals is proposed. Example of the proposed search engine is implemented with Yahoo!JAPAN Web SearchAPI, Yahoo!JAPAN Image searchAPI, YouTube Data API, and Amazon Product Advertising API. The retrieved results are also merged and linked each other. Therefore, the different types of contents can be referred once an e-learning content is retrieved. The implemented search engine is evaluated with 20 students. The results show usefulness and effectiveness on e-learning content searches with a variety of content types, image, document, PDF files, moving picture.

The following section describes the proposed search engine followed by implementation results and experimental results with 20 students which show usefulness and effectiveness of the proposed content search engine for mobile devices. Then conclusion is described together with some discussions on future investigations.

II. PROPOSED METHOD AND SYSTEM

A. Background of the Proposed Search Engine

Figure 1 shows research background of the proposed search engine.

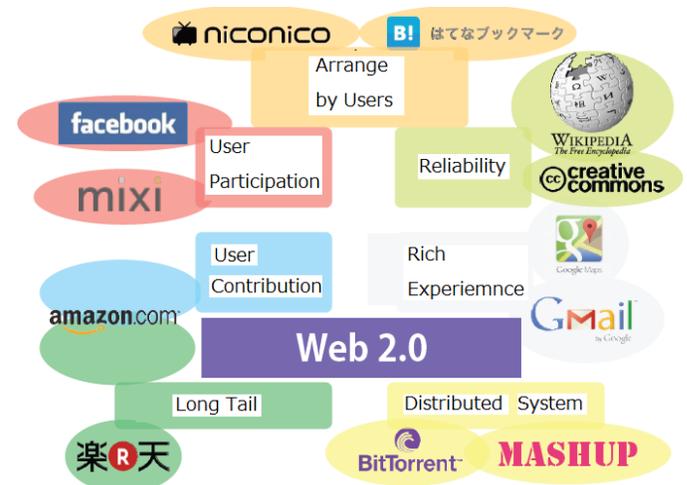


Fig.1. Research background of the proposed mashup based content search engine

There are seven key components, user involvement, user contribution, arranging by users, reliability, rich experiences,

¹ <http://pipes.yahoo.com/pipes/>

² <http://www.microsoft.com/ja-jp/dev/default.aspx>

³ <http://chamap.net/>

⁴ <http://newsgraphy.com/>

⁵ <http://www.flowser.com/>

long tail, and distributed system. Facebook, Mixi, etc. of social network system allows user involvement. Users may contribute to create huge database through amazon.com, for instance. Long tail becomes true with amazon.com, Rakuten, etc. Not only contents but also rearrangement of the contents can be done by users through Hatena Bookmark, Nikoniko, etc. Reliable Q&A is available through Wikipedia, Yahoo answer, etc. Also rich experiences are obtained through e-mail communications with Gmail, for instance.

These activities can be done based on Web 2.0 technology utilized distributed system. In particular, Mashup technology allows efficient and effective information retrievals for the distributed information providing systems.

B. Content Types and Representation Models for the Content Search

There are some contents types for search. Namely, document, moving picture, a variety of types of contents, images, etc. Also there are some representation models for the different search content types, helical, star, star-helical, and star slide models. Figure 2 shows an example of star slide model of representation of contents types (Blue star) and the search results (Smile marks). Namely, there are the different content types on the top and there are search results under the content types. With the touch panel function, these content types can be selected (swipe in horizontal direction) and also search results can be selected through swipe in vertical direction.



Fig.2. Content type and search result selections by star slide model of representation method

Figure 3 shows five content types as example. Namely, there are Youtube for moving picture, Amazon.com for product, Yahoo for document, Yahoo search for image, and Yahoo for Web searches as shown in Figure 3.

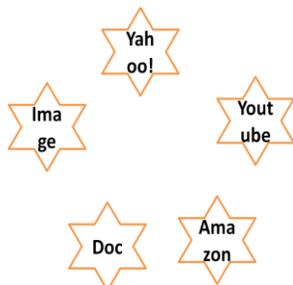


Fig.3. Five different content types are available for search

Under the icon of these five content types, candidate contents are aligned below the icons as search results. In order to show the candidate contents in a 3D representation, the sizes of icons and the candidate contents are changed depending on their locations. Namely, the icons which are located near forward are displayed with relatively large size while those which are located near backward are displayed with comparatively small size as shown in Figure 4.

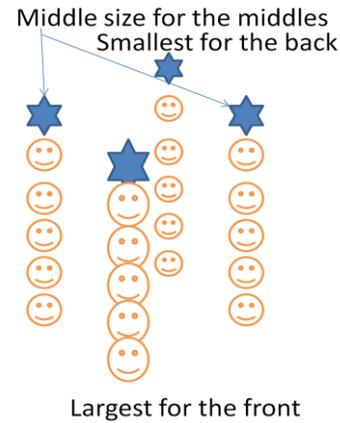


Fig.4. Icon sizes are changed depending on their locations

Also swipe for candidate URLs in vertical direction and flick for rotation of candidate URLs in horizontal direction are available as shown in Figure 5.

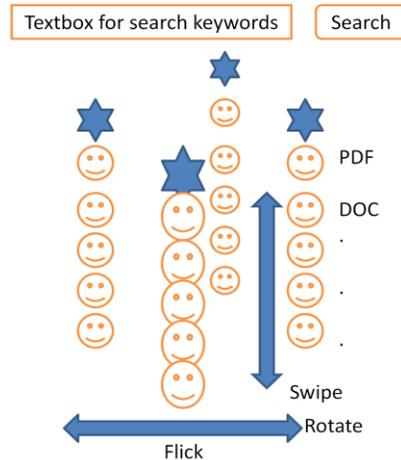


Fig.5. Swipe for candidate URLs in vertical direction and flick for rotation of candidate URLs in horizontal direction

Away3D on Andoroide is used for 3D representation of icons while the APIs which are shown in Figure 6 are used for Mashup. These are aligned in the pentagonal shape as shown in Figure 7. Meanwhile, example of icon movement through swipe for the case of Yahoo search is shown in Figure 8. Candidate of URL icon is selected with swipe (up and down operations).

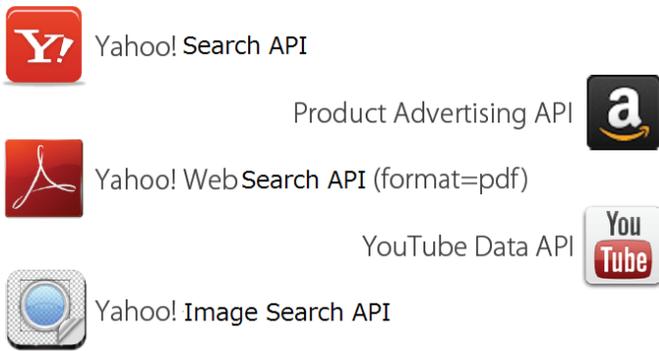


Fig.6. APIs used for mashup

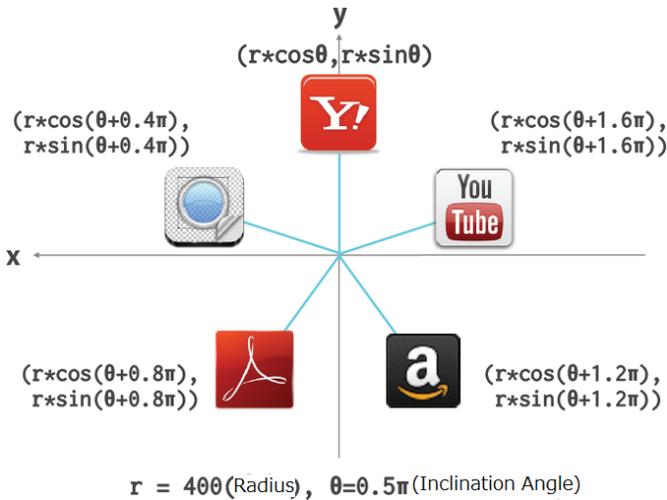


Fig.7. API icons alignment

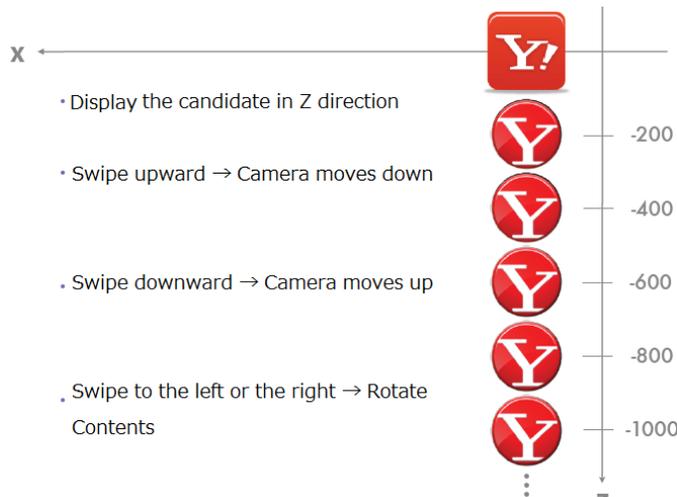


Fig.8. Example of icon movement through swipe for the case of Yahoo search

C. Examples on Content Search Operations

The icon of the application software, ELDOXEA: e-learning document search engine appears on the screen as shown in Figure 9. ELDOXEA is proposed for e-learning content search [1]-[5].

In this section, example of the search operation with the keyword of “java” is demonstrated as shown in Figure 10. First, users have to key-in their keyword through the dialog box. After the research results appear on the screen, then users have to select the candidate of the search results with swipe and flick operations.

When you click the icon, then Figure 11 (a) of dialog box with search start of radio button appears followed by keyboard screen as shown in Figure 11 (b). When you key-in the keyword, “java” for example, then search result of screen images appear as shown in Figure 12. Examples of swipe and flick operations are shown in Figure 13. Thus users may take a look at research results freely and easily. After that, when user click one icon of the candidate of the search results from the candidates of icons as the search results, the content of the candidate appears as shown in Figure 14.



Fig.9. Example of display image of smartphone of which the proposed search engine is implemented

github.com/legnoh/ledoxea



Fig.10. Example operations for search with keyword of “java”



Title
Discription

(a)Search startup screen image (b)Key-in screen image

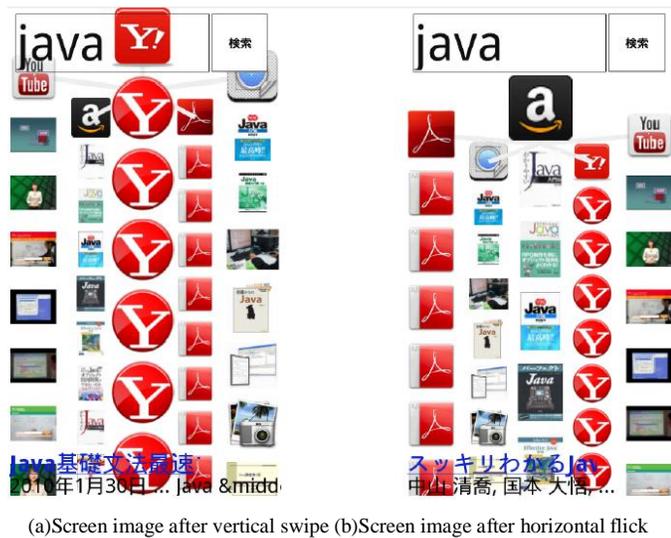
Fig.11. Search startup screen image



(a)Key-in screen image

(b)Search results screen image

Fig.12. Search results of candidates



(a)Screen image after vertical swipe (b)Screen image after horizontal flick

Fig.13. Example of search result screen images after swipe and flick operations



Fig.14. Examples of all the five content search results of screen images

D. Merged Saerch Results

At the bottom of the screen, there are four content types as shown in Figure 15. In the Figure 15, screen image is the search results from the Web search (Yahoo search). Therefore, the rest of four content types, Documents, Images, Moving picture, and PDF files are on the list. When you click one of those, then you get the closest related content because the contents have links to the closest different types of contents. There are some keywords in the Meta tag of the source code of URLs. By using the number of common keywords in the Meta tag keywords for both content types of URLs, the distance between URLs is defined. Thus the closest URL can be retrieved using the distance between URLs. This is also specific feature of the Mashup technology of search engine.



Fig.15. Example of linked retrieved results (Merged contents of Yahoo search results with the other retrieved contents)

III. EXPERIMENTS

20 students in the laboratory of department of information science, Saga University are participated to the experiments using the developed ELDOXEA. Then more than 30 of reports are sent as shown in Figure 16.

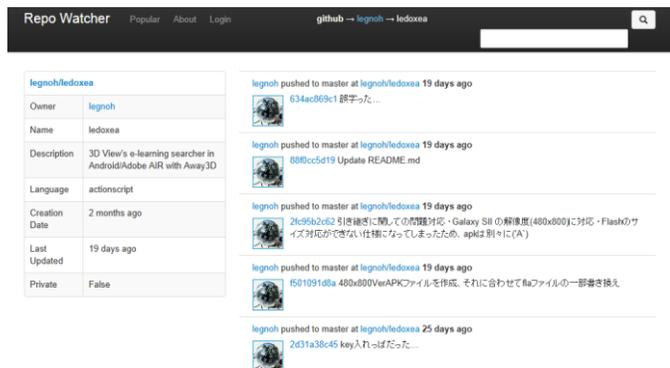


Fig.16. Figure 16 Small portion of the report from the users (20 of students) about ELDOXEA

Although most of those reports are positive, effective search and easy to use, a small number of negative report say that merged contents are not so easy to refer. Link structure has to be clearly shown in the screen.

IV. CONCLUSION

Mashup based content search engine for mobile devices is proposed. Example of the proposed search engine is implemented with Yahoo!JAPAN Web SearchAPI, Yahoo!JAPAN Image searchAPI, YouTube Data API, and Amazon Product Advertising API. The retrieved results are also merged and linked each other. Therefore, the different types of contents can be referred once an e-learning content is retrieved. The implemented search engine is evaluated with 20 students. The results show usefulness and effectiveness on e-learning content searches with a variety of content types, image, document, PDF files, moving picture.

Through experiments with 20 of students who use ELDOXEA for e-learning content search, it is found that the proposed search engine provides e-learning content retrievals comfortably and in a comprehensive manner. On the other hand, there is the report which says that merged contents which

are closely related to the retrieved content are not so easy to refer. It should be solved in the future.

ACKNOWLEDGMENT

The author would like to thank Mr. Ryoma Kai, for his effort to implement search engine application software and to conduct experiments.

REFERENCES

- [1] Kohei Arai, Herman Tolle, Model based content adaptation of composite e-learning content for delivering to mobile learners, International Journal of Computer Theory and Engineering, 3, 3, 382-387, 2011
- [2] Kohei Arai, Herman Tolle, Efficiency improvements of e-learning document search engine for mobile browser, International Journal of Research and Reviews on Computer Science, 2, 6, 1287-1291, 2011.
- [3] K.Arai, T.Herman, Efficiency improvement of e-learning document search engine for mobile browser, International Journal of Research and review on Computer Science, 2, 6, 1287-1291, 2012.
- [4] K.Arai, T.Herman, E-learning document search method with supplemental keywords derived from keywords in meta-tag and descriptions which are included in the header of the first search result, International Journal of Advanced Computer Science and Applications, 3, 4, 99-104, 2012.
- [5] K.Arai, T.Herman, Video searching optimization with supplemental semantic keyword for e-learning video searching, International Journal of Research and Review on Computer Science, 3, 3, 1640-1644, 2012.

AUTHORS PROFILE

Kohei Arai He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission-A of ICSU/COSPAR since 2008. He wrote 30 books and published 332 journal papers.

The preliminary results of a force feedback control for Sensorized Medical Robotics

Duck Hee Lee

Medical Engineering R&D Center
Asan Institute for Life Sciences, Asan Medical Center
Seoul, South Korea

Reza Fazel-Rezai

Department of Electrical Engineering
University of North Dakota
Grand Forks, North Dakota, USA

Seung Joon Song

Department of Convergence Biomedical Engineering
Daelim University College
Gyeonggi-do, South Korea

Jaesoon Choi

Medical Engineering R&D Center,
Asan Institute for Life Sciences, Asan Medical Center and
University of Ulsan, College of Medicine
Seoul, South Korea

Abstract—A laparoscopic surgery system by using a robot holds many problems. Among these, its inability in delivering touching sensation to a surgeon is raised as the biggest problem. The current paper attempted to find a force feedback controlling method at the time of performing movement by using one-degree of freedom (DOF) arm of slave and master system that is used in the programming based system. The study was two methods experience force feedback control; In the first place used force sensors and otherwise, conducted for the force feedback control by using a force sensor and for the case when the sensor could not be used due to the spatial and systematic limitation. The realization of force feedback system was successful, and the experiment results of force feedback control and current based force feedback control mode based on the force sensors of one-DOF system indicates that it could be directly applied to the another multi-DOF surgical robot system that is currently under the development.

Keywords—Laparoscopic surgery; force feedback control; degree of freedom (DOF).

I. INTRODUCTION

Nowadays, the robot has been used in various fields for the development of industrial technology. In particular, it is utilization high for automotive and specialty industrial fields. Using a precision operation of robot is very good matching for required medical fields, because it is available details and accuracy surgical operations. Hence, the application of robot in medical field has been rapidly progressed in last decade. Based on its application, a medical robot is classified into surgical robots, human robots to help physically handicapped and aged people, and Bio robots to study biomechanisms [1]. While using a surgical robot, the sensing information delivery by using hand provides much information by surpassing visual and aural information. All of sensations felt by hand are called as haptic, and the haptic interface related researches have been initiated from remote operation or from a devise for handicapped person, and developed into telesurgery, remote control, and application in space & aeronautic fields [2-3]. Especially, Minimally Invasive Surgery (MIS) in the field of medicine operates a surgery without opening a lesion, but a

surgeon operates a surgery by seeing a monitor after inserting surgical tools and camera through a hole in a lesion. Although the method's application scope is enlarged due to the advantages of shorter surgery time and shorter recuperation time of patients, but there is a three major disadvantage [4]; (a) As the surgeon does not have direct access to the operating field the tissue cannot be palpated any more. (b) Due to the friction in the trocar and due to the torque which are necessary to rotate the instrument around the entry point, the appearing contact forces between instrument and tissue can hardly be sensed. (c) As the instrument has to be moved around an invariant fulcrum point intuitive, direct hand-eye coordination is lost and due to the kinematic restrictions only four-DOF remains inside the body of the patient. Therefore, the surgeon could not feel the sensation when he directly operates a surgery [5-7]. Several robot surgical systems including Zeus and da-Vinci systems have been developed, but no report has been made yet for a haptic Interface providing system that could be applied onto an actual surgery. The current paper conducted a study for a force feedback control method that can be applied onto the currently developing surgical robot systems. The study was two type experiences for force feedback control; In the first place used force sensor and otherwise, conducted for the realization of force feedback control by using force/torque sensors and for the case when the sensors could not be used due to the spatial and systematic limitation.

II. MATERIALS AND METHODS

A. System Composition

In generally, surgical robot system is a consist of master and slave consoles. Therefore, many other researchers has been actively research of medical robot structure analysis and system developed the through after 1990s, but it is need to high cost and times. The development of haptic device that is used haptic interface for tactile transmission and control can be used to replace the master system. We used master system are PHANTOM device that it have 6-DOFs functions for SensAble™ [8]. This is characteristic as following; workspace of 381*276*191mm, nominal position resolution of 0.007mm,

maximum executable force of 37.5N and stiffness of 3.5N/mm. The slave system used for Adept 550 industry robot (adept technology, Inc) [9]. This Adept 550 robot is a four-axis SCARA (Selective Compliance Assembly Robot Arm) robot. Among then, we used attached to the surgical instrument dummy tool on the one-axis. Accordingly, the study was conducted by using translational 1-DOF among total 4-DOFs of slave. Figure 1 illustrates the overall composition of the electrical current and force sensor base force feedback control system block diagram. The Slave were equipped with DC motors (Maxon Motor AG, Switzerland), and general controllers were used. The general controller design is based on a nonlinear system control approach. The implemented of control software was generated from the C/C++ with Pentium processor 2.5GHz CPU (Intel Co., Ltd.) and 2Gbyte of RAM. For the communication method between the control software and general controller that is connected to the master/slave, it used a performance proven, a well recognized serial bus type CAN (Controller Area Network) communication method. Through the minimum 3 lines of simple and small volume of communication lines, the serial communication up to 1Mbps was possible. Since the collision prevention and error control processing mechanisms were already well established on a standard phase, reliable communication was possible. The bus type allowed multiple controllers to share communication lines efficiently, which fitted the proposal of the current system. The motor used for the control was controlled by using 1000Hz cycle.

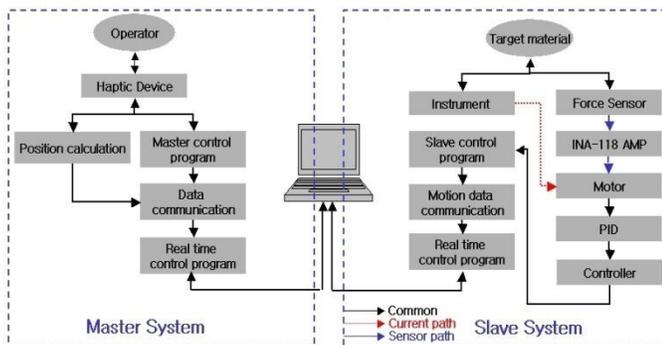


Fig. 1. Force feedback control block diagram

B. Control Method

The force feedback control scheme has an impedance control and admittance control about the haptic device [10]. Impedance controlled systems detect motion commanded by the operator and control the force applied by the haptic device. Admittance controlled systems detect the force commanded by the operator and control the velocity [11]. Also, The control methods for tele-manipulator systems used position to position control and position to force control. First case, controller torques applied to the master and slave are equal, but this is occur position and velocity error between the two device. Another case, It is similar to a position-position controller, except that the torque applied to the master system is based on a force sensor measurement on the slave system [12]. We system is a set the slave to follow the movement of the master haptic device and a controller was established based on the position of the master. From the movement position signals of

the master device, displacement is detected to transmit the value to the operation for the position control command value of the slave. The tracking of the master command value used PID control method for the slave system. To feedback the amount of load received in the slave to the master side, an electric current based control and a sensor using control methods were used. The electric current based control method recognizes the force applied into the slave by using an electric current value of the slave and transmits the value to the master side for the control. The calculation and programming based control method was used for the electric current value input control of the master in Figure 2. For the Figure 2 illustration as follows: The controller is described by the transfer functions PF_{MX} , PF_{MY} and PF_{MZ} which are each position and force of master system. The inverse Jacobian matrix JM^{-1} is computed by the estimated joint position θ . The delay time T_D dominates the delay of the communication network between master and slave that is a less than 1ms. For the force feedback control method by using sensors, the Load-Cell sensor was installed at the tool end of the slave as illustrated. The output of the Load-Cell sensor is approved as the input of control program through an amplification circuit. Such approved slave force could be felt at the master through the force and electric current control of the master haptic device.

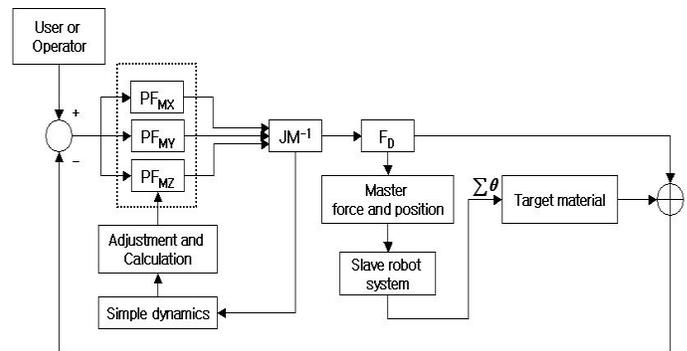


Fig. 2. Control system architecture for current and force sensor based force feedback

C. Current and force control function

In order to evaluate the interrelation of the electric current and force we compared current values according to the each condition after making the identical control condition to measure the current and force using no-load condition and soft-tissue.

For the no-load condition, the translation motor of the slave needs to be regular current to move the slave robot according to moving of master. Also when inflicting the load, it is received the position of translation movement and the force of reaching to the soft-tissue from the motor current. Therefore, in order to evaluate the interrelation of the current and force in this system, it is necessary to get information of the position as well as inflicting the force to the object. To examine the current and force, the following three equations are needed. For estimating the values of the force, three input variables, that is, position error (p_e) of the master and slave, the first variation ($\Delta p_e(n)$) of the error, and the second variation ($\Delta p'$) of the error as follows:

$$p_e = P_{master}(n) - P_{slave}(n)$$

$$\Delta p_e(n) = p_e(n) - p_e(n-1)$$

$$\Delta p'_e = \Delta p_e(n) - \Delta p_e(n-1)$$

The interrelation equation of the current according to the force is as follows:

$$i_{current} = f(\Delta a)$$

Where Δa denotes the first and second variation of the position.

We can figure out the current when reaching to the soft tissue using the current value from equation:

$$i_{noload} = C_{noload}(n) - MA_{ChangeVactor}(n)$$

$$i_{load} = p_e(n) - MA_{ChangeVactor}(n)$$

$$i_{contact} = i_{noload} - i_{load}$$

$$MA_{ChangeVactor}(n) = Ax(n) + Ax(n-1) + Ax(n-s+1) + Ax(n-s)$$

Where $MA_{ChangeVactor}(n)$ is the value of the moving average in the interval where the value changes positive to negatives.

III. EXPERIMENT

The force sensor used for the experiment was composed as bridge circuit, and the tension/compression available UMN-K5 (DACELL Co., Ltd.) was used. The force sensor operates at DC 5V with input and output impedance of 352Ω, rated output of 0.9102mV/V, small size of 27*34mm and maximum measurement up to 49.03N is possible.

A. Design of a force sensor amplification circuit

To amplify few mill-voltages (mV/V) of output signal values outputted from a force sensor, an amplification circuit (INA118U), which enables 1000 times of amplification was designed.

The amplifier used at the current paper is precise, which can be operated under the low voltage and amplification coefficient was adjusted by using variable resistance. This amplify circuit board was a size of 76.2*53.5mm and 4-layer PCB (printed circuit board). In addition, a notch filter was used to remove 60Hz of power noise (Figure 3).

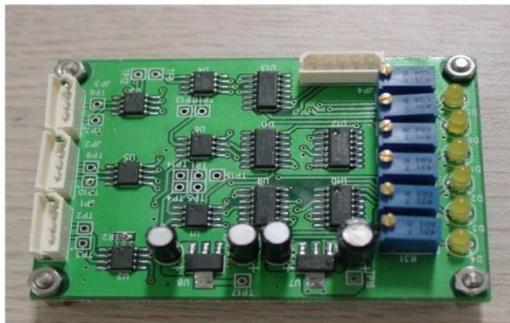


Fig. 3. Force sensor amplifier circuit board

B. Force sensor weight simulation

By using an experimental weight, the force sensor output at the 1-DOF translation movement was confirmed by the weight. The output experiment of the force sensor responding to 0.09N, 0.9N, 4.9N and 9.8N of weights was performed under the test condition of 10 seconds force application before 5-seconds of force removal (Figure 4).

C. Relationship of force with electric current on No-Load and test material

The electric current and force change was measured by actuating the master and slave system. As illustrated in Figure 5, it was possible to see the occurrence of motor current value under the No-Load state. It is also compared force sensor signals value. The motor current was developed by the interaction of equipment that holds a model tool in the slave robot arm at the time of Translation movement. To understand the relationship between electric current and force at the time of Translation movement on the test material, a laparoscopic surgery training use tissue suture pad was used to perform the test. The highest electric current was observed in test material model. However, the electric current under the No-Load state was observed in ± 500 mA range. It revealed that relatively large amount of electric current was flowing even under the No-Load state. It is a movement current following the master command to slave system.

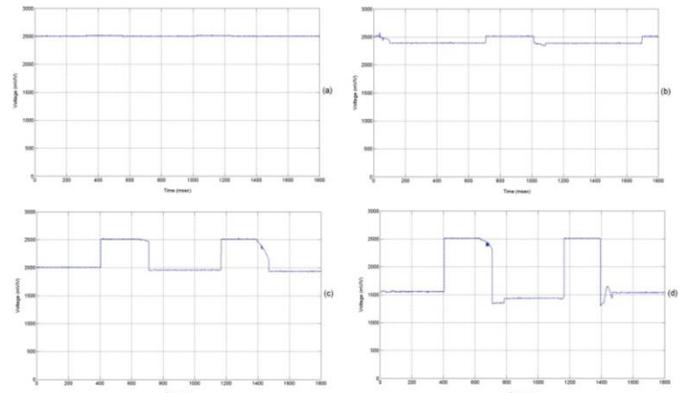


Fig. 4. Test of sensor weight simulation: (a) 0.09N; (b) 0.9N; (c) 4.9N; (d) 9.8N (X-axis is time(msec) and Y- axis is voltage(mV/V))

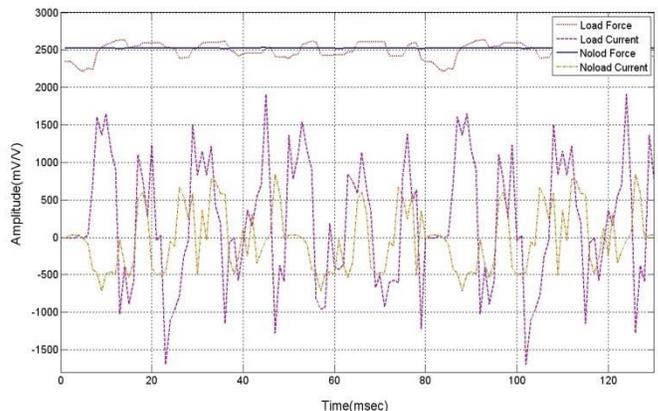


Fig. 5. Relationship of each No-Load and Load force with electric current

IV. RESULTS

To control the force feedback in this paper, we use electrical current values of the motor instead of sensor value, because of restriction of the space and structure in the laparoscopic surgery robot system that is in the process of the development.

Figure 6 shows the results of using the 5-point moving average algorithm for the current values that are obtained when no-load condition and pressing the test material object. That it is obtained from subtracting current values between no-load condition and pressing the test material. This means that the actual current values when pressing the soft-tissue object. Figure 7 shows position following performance of the slave. It shows good performance. The master and slave system of correction ($r=0.93$) for each position has an positive linear relation.

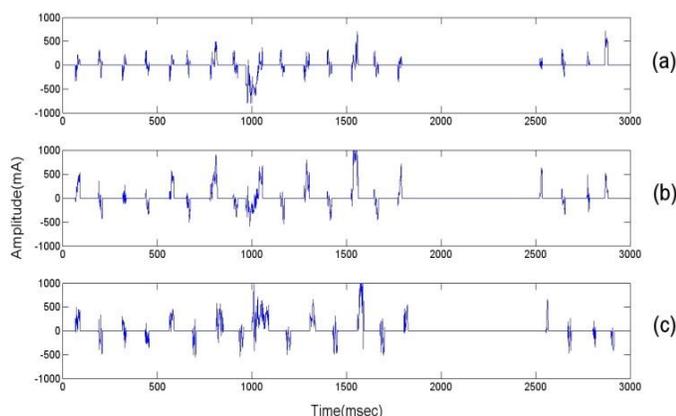


Fig. 6. Results of moving average feedback current value; (a) No-load current; (b) Load current; (c) Load and No-load current subtract

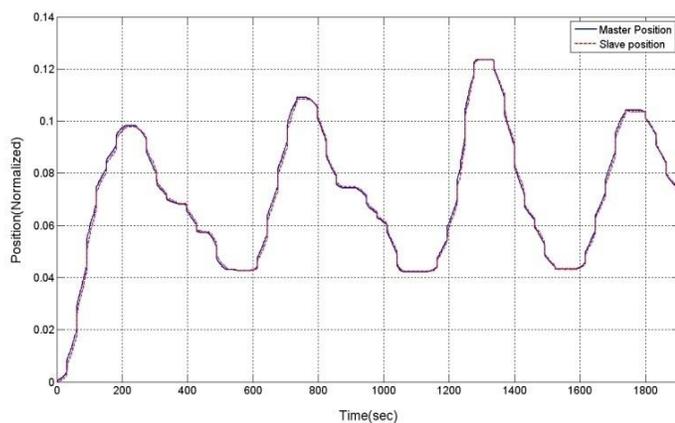


Fig. 7. Results of from master system to slave system position tracking

V. CONCLUSION

The surgical robot systems have been used on the surgery method for various diseases through the MIS. Therefore, the transit of haptic information to the surgeon's are such an environment as open surgery which it is appropriate microsurgery and bad location the affected part of surgery. The current paper describes a force feedback controlling method by using a force sensor or motor current in a laparoscopic surgery

robot system. As it has been tested in the current study, the application of the force feedback control method by using a force sensor onto a surgery robot tends to require use of more sensitive force sensor than the one currently used, and the problem of vibration at the slave robot arm has to be improved at the time of translation movement. In addition, the application of force sensor has to be applied onto the actuation, rotation, pitch, and yaw DOF other than the translation movement suggested in the current study, and reviews and studies have to be progressed on the force feedback control method by using motor current along with the force feedback control method on position.

Based on the above result, future studies have to be made on the method of using force/torque at the whole many other DOF of the slave and to realize force feedback system by using more various algorithms. We also should be study to development of component and small size sensor for medical robotics fields.

ACKNOWLEDGMENT

This work was supported by the Industrial Strategic technology development program (10041618) funded by the Ministry of Knowledge Economy (MKE, Korea).

REFERENCES

- [1] P. Dario et al., "Robotics for Medical Application", IEEE Robot and Automation Magazine, Vol. 3, No. 3, pp.44-56, 1996
- [2] G. C. Burdea, "The Synergy Between Virtual Reality and Robotics", IEEE Transactions on Robotics and Automation, Vol, 15, No.3, pp.400-410, 1999.
- [3] W. Barfield and T. A. Furness, "Virtual environments and advanced interface design", Oxford University Press. Inc, 1995, pp.415-436.
- [4] M.R. Treat, "Computer-Integrated Surgery, A Surgeon's Perspective on the Difficulties of Laparoscopic Surgery", MIT Press, 1995, pp. 559-560.
- [5] U. Seibold et al., "Prototype of Instrument for Minimally Invasive Surgery with 6-Axis Force Sensing Capability", International Conference on Robotics and Automation, Barcelona, Spain, 2005, pp.498-503.
- [6] B. Deml et al., "The Touch and Feel in Minimally Invasive Surgery", IEEE International Workshop on Haptic Audio Visual Environment and their Applications, Ontario, Canada, 2005.
- [7] M. Tavakoli et al., "A Force Reflective Master-Slave System for Minimally Invasive Surgery", IEEE/RSJ International Conference on Intelligent Robots and Systems, Las Vegas, USA, 2004, pp.3077-3082.
- [8] Available from: <http://www.sensable.com>
- [9] Available from: <http://www.adept.com>
- [10] R. Baumann and R. Clavel. "Haptic interface for virtual reality based minimally invasive surgery simulation," IEEE Interface Conference on Robotics and Automation, Leuven, Belgium, 1998, pp.381-386.
- [11] C. R. Carignan and K. R. Cleary. "Closed loop force control for haptic simulation of virtual environments," The Electronic Journal of Haptics Research, Vol. 1, No. 2, pp.1-14, 2000.
- [12] L. N. Verner and A. M. Okamura. "Sensor/Actuator asymmetries in telemanipulators: Implications of partial force feedback," Symposium on Haptic Interfaces for Virtual Environment and Teleoperator System, Virginia, USA, 2006, pp.309-314.
- [13] S. Wang et al., "A robotic system with force feedback for Micro-Surgery," IEEE International Conference on Robotics and Automation, Barcelona, Spain, 2005, pp.1999-204.
- [14] T. Hoshino et al., "A master-slave manipulation system with a force-feedback function for endoscopic surgery," International Conference of the IEEE Engineering in Medicine and Biology Society. Istanbul, Turkey, 2001, pp.3446-3449.

- [15] M. Wu et al., "Effects of velocity on human force control," EuroHaptics Conference and Symposium on Haptic Interface for Virtual Environment and Systems, Pisa, Italy, 2005, pp.73-79.
- [16] D. H. Lee et al., "An implementation of sensor based force feedback in a compact laparoscopic surgical robot," American Society of Artificial Internal Organs, Vol. 55, No. 1, pp.83-85, 2009.
- [17] D. J. Callaghan and M. M. McGrath. "A force measurement evaluation tool for telerobotic cutting applications: Development of an effective characterization platform," Genamics, International Journal of Mathematical, Physical and Engineering Sciences, Vol. 1, No. 3, pp.144-150, 2008.
- [18] J. Peirs et al., "A micro optical force sensor for force feedback during minimally invasive robotic surgery," Springer, Sensor and Actuators A: Physical, Vol. 115, No.2-3, pp.447-455, 2004.
- [19] N. Zemiti et al., "A new robot for force control in minimally invasive surgery," IEEE/RSJ International Conference on Intelligent Robots and Systems, Sendai, Japan, 2004, pp.3643-3648.
- [20] A. Krupa et al., "Achieving high precision laparoscopic manipulation through adaptive force control," IEEE International Conference on Robotics and Automation, Washington, DC, USA, 2002, pp.1864-1869.
- [21] M. Kitagawa et al., "Analysis of suture manipulation forces for teleoperation with force feedback," International Conference on Medical Image Computing and Computer Assisted Intervention, Tokyo, Japan, 2002, pp.155-162.
- [22] L. F. BAPTISTA et al., "Force control of robotic manipulators using a fuzzy predictive approach," Springer, Journal of Intelligent and Robotic Systems, Vol. 30, No. 4, pp.359-376, 2001.
- [23] Z. Zhang et al., "Master slave manipulators bilateral control system with force telepresence" IEEE International Conference on Robotics and Biomimetics, Sanya, China, 2007, pp.307-311.
- [24] D. Wang and N. H. McClamroch. "Position and force control for constrained manipulator motion: Direct method" IEEE Transactions on Robotics and Automation, Vol. 9, No. 3. pp.308-313, 1993.
- [25] V. Urban et al., "Robot assisted surgery system with kinesthetic feedback," Informa Healthcare, Computer Aided Surgery, Vol. 3, No. 4, pp.205-209, 1998.
- [26] M. J. H. Lum et al., "Optimization of a spherical mechanism for a minimally invasive surgical robot: Theoretical and experimental approaches," IEEE Transactions on Biomedical Engineering, Vol. 53, No. 7, pp.1440-1445, 2006.
- [27] K. Vachos and E. Papadopoulos. "Using force control for fidelity in low-force medical haptic simulators," IEEE International Conference on Control Applications, Munich, Germany, 2006, pp.181-186.
- [28] Y. Kobayashi et al., "Control method for surgical robot to prevent overload at vulnerable tissue," IEEE International Conference on Robotics and Automation, Roma, Italy, 2007, pp.1893-1899.
- [29] T. Chen and Z. Zhang. "Design and research of Tel-operation manipulator with force feedback," International Conference on Intelligent Computation Technology and Automation, Changsha, China, 2008, pp.979-983.
- [30] D. Kubus et al., "Improving force control performance by computational elimination of non contact Forces/Torque," IEEE International Conference on Robotics and Automations, Pasadena, USA, 2008, pp.2617-2622.
- [31] B. Finkemeyer et al., "Executing assembly tasks specified by manipulation primitive nets," Advanced Robotics, Vol. 19, No. 5, pp.591-611, 2005.

Performance Analysis and Evaluation of Different Data Mining Algorithms used for Cancer Classification

Gopala Krishna Murthy Nookala

Department of Computer Science & Engineering
Grandhi Varalakshmi Venkatarao Institute of Technology
Bhimavaram, Andhra Pradesh, India

Bharath Kumar Pottumuthu

Centre for Bioinformatics Research & S/w Development
Grandhi Varalakshmi Venkatarao Institute of Technology
Bhimavaram, Andhra Pradesh, India

Nagaraju Orsu

Department of Computer Science
Government Degree College
Macherla, Andhra Pradesh, India

Suresh B. Mudunuri*

Centre for Bioinformatics Research & S/w Development
Grandhi Varalakshmi Venkatarao Institute of Technology
Bhimavaram, Andhra Pradesh, India

* Corresponding Author

Abstract—Classification algorithms of data mining have been successfully applied in the recent years to predict cancer based on the gene expression data. Micro-array is a powerful diagnostic tool that can generate handful information of gene expression of all the human genes in a cell at once. Various classification algorithms can be applied on such micro-array data to devise methods that can predict the occurrence of tumor. However, the accuracy of such methods differ according to the classification algorithm used. Identifying the best classification algorithm among all available is a challenging task. In this study, we have made a comprehensive comparative analysis of 14 different classification algorithms and their performance has been evaluated by using 3 different cancer data sets. The results indicate that none of the classifiers outperformed all others in terms of the accuracy when applied on all the 3 data sets. Most of the algorithms performed better as the size of the data set is increased. We recommend the users not to stick to a particular classification method and should evaluate different classification algorithms and select the better algorithm.

Keywords—Weka; Cancer Classification; Micro-array; Data-mining; Classification Algorithms; Gene Expression Data;

I. INTRODUCTION

Advancement of Information Technology led to huge data accumulation in the recent years in several domains including banking, retail, telecommunications and medical diagnostics. The data from all such domains includes valuable information and knowledge which is often hidden. Processing the huge data and retrieving meaningful information from it is a difficult task. Data Mining is a wonderful tool for handling this task. The term Data Mining, also known as Knowledge Discovery in Databases (KDD) refers to the non trivial extraction of implicit, previously unknown and potentially useful information from data in databases [1]. Data mining in cancer research has been one of the important research topics in biomedical science during the recent years [2].

They are several different data mining techniques like Pattern Recognition, Clustering, Association and Classification

[3]. Classification has been identified as an important problem in the emerging field of data mining [4] as they try to find meaningful ways to interpret data sets. Classification of data is very typical task in data mining. There are large number of classifiers that are used to classify the data namely Bayes, Function, Rule's based, Tree based classification etc. The goal of classification is to correctly predict the value of a designated discrete class variable, given a vector of predictors or attributes [5]. In the age of bioinformatics, cancer data sets have been used for the cancer diagnosis and treatment that can improve human aging [6].

Cancer is a disease characterized by uncontrolled growth and spread of the abnormal cells and the capability to invade other tissues that can be caused by both external factors like radiation, chemicals, tobacco etc., and internal factors like inherited mutations, hormones, immune conditions, etc. There are more than 100 different types of cancers. Most of the cancers are named after the organ or type of cell in which they appear e.g., Melanoma, Colon Cancer, Breast Cancer etc.

All cancers begin in cells which are the structural and functional units of the body. These cells grow and divide in a controlled way to produce more cells as they are needed to keep the body healthy. When cells become old or damaged, they die and are replaced with new cells. However, sometimes life cycle of the cells fails or gets disturbed due to many reasons. When this happens, cells do not die as expected and new cells are formed even when the body does not need them. These extra cells may form a mass of tissue called a tumor. Tumors can be either benign or malignant. Some cancers do not form tumors. For example, leukemia is a cancer of the blood that does not form tumors.

Gene expression analysis of cancer is used to study regulatory gene defects and other devastating diseases, cellular responses to the environment, cell cycle variation, etc. When genes are expressed, the genetic information (base sequence) on DNA is first transcribed (copied) to a molecule named messenger RNA (mRNA). The mRNA molecules further

participate in protein synthesis by specifying the particular amino acids that make up individual proteins. Gene Expression Analysis is one of the major applications of the Micro-array. Microarray is a hybridization of a nucleic acid sample (target) to a very large set of oligo-nucleotide probes, which are attached to a solid support (chip), to determine sequence or to detect variations in a gene sequence or expression levels or for gene mapping.

In the recent years, tumor classification is frequently studied by applying various data mining classification algorithms on cancer gene expression micro-array data sets so as to predict the presence of cancer. However, the availability of several algorithms in data mining for classification often leads to confusion over the selection of the right algorithm. In this study, we have made a comparative analysis of the performances of various classification algorithms on different cancer micro-array data sets.

II. MATERIALS AND METHODS

We have used the popular, open-source data mining tool Weka (version 3.6.6) for this analysis. Three different data sets have been used and the performance of a comprehensive set of classification algorithms (classifiers) has been analyzed. The analysis has been performed on a HP Windows system with Intel® Core™ i3 CPU, 2.40 GHz Processor and 4.00 GB RAM. The data sets have been chosen such that they differ in size, mainly in terms of the number of attributes.

A. Data set 1:

The first data set is a small Breast Cancer Micro-array Gene Expression data used in an earlier study [7]. The data set contains 9 attributes apart from the class attribute with 286 instances.

B. Data set 2:

The second data set is a medium sized data set with Micro-array Gene Expression data of Lymphoma patients [8]. The data set has a total of 4,026 attributes and 45 instances.

C. Data set 3:

The large data set 3 is also a Micro-array Gene Expression data of Leukemia with 7,129 attributes and 34 instances [9].

D. Classifiers Used:

A total of 14 classification algorithms have been used in this comparative study. The classifiers in Weka have been categorized into different groups such as Bayes, Functions, Lazy, Rules, Tree based classifiers etc. A good mix of algorithms have been chosen from these groups that include Bayes Net & Naive Bayes (from Bayes), Multilayer Perceptron, Simple Logistics & SMO (from functions), IBk & KStar (from Lazy), NNge, PART & ZeroR (from Rules) and ADTree, J48, Random Forest & Simple Cart (from Trees). The following sections explain a brief about each of these algorithms.

1. Bayes Net

Bayes Nets or Bayesian networks are graphical representation for probabilistic relationships among a set of random variables. A Bayesian network is an annotated Directed

Acyclic Graph (DAG) that encodes a joint probability distribution [10].

2. Naive Bayesian

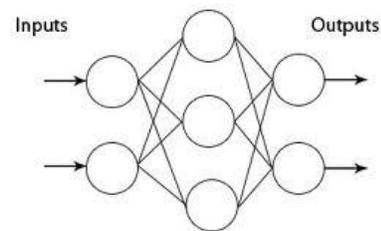
Naive Bayesian classifier is developed on bayes conditional probability rule used for performing classification tasks, assuming attributes as statistically independent; the word Naive means strong. All attributes of the data set are considered as independent and strong of each other [11].

3. Simple Logistics

It is a classifier used for building linear logistic regression models. LogitBoost with simple regression functions are base learners used for fitting the logistic models. The optimal number of LogitBoost iterations to perform is cross-validated, which leads to automatic attribute selection [12].

4. Multilayer Perceptron

Multilayer Perceptron is a nonlinear classifier based on the Perceptron. A Multilayer Perceptron (MLP) is a back propagation neural network with one or more layers between input and output layer. The following diagram illustrates a perceptron network with three layers [13].



5. SMO

Sequential Minimal Optimization (SMO) is used for training a support vector classifier using polynomial or RBF kernels. It replaces all missing the values and transforms nominal attributes into binary ones [14]. A single hidden layer neural network uses exactly the same form of model as an SVM.

6. IBk

IBk is a k -nearest-neighbor classifier that uses the same distance metric. k -NN is a type of instance based learning or lazy learning where the function is only approximated locally and all computation is deferred until classification. In this algorithm an object is classified by a majority vote of its neighbors [15].

7. KStar (K^*)

Aha, Kibler & Albert describe three instance-based learners of increasing sophistication. IB1 is an implementation of a nearest neighbor algorithm with a specific distance function. IB3 is a further extension to improve tolerance to noisy data. Instances that have a sufficiently bad classification history are forgotten and only instances that have a good classification history are used for classification. Aha [16] described IB4 and IB5, which handle irrelevant and novel attributes.

8. NNge

Instance-based learners are “lazy” in the sense that they perform little work when learning from the data set, but expend more effort classifying new examples. The simplest method, nearest neighbor, performs no work at all when learning. NNge does not attempt to out-perform all other machine learning classifiers. Rather, it examines generalized exemplars as a method of improving the classification performance of instance-based learners [17].

9. PART

PART uses the separate-and-conquer strategy, where it builds a rule in that manner and removes the instances it covers, and continues creating rules recursively for the remaining instances. Where C4.5 and RIPPER does global optimization to produce accurate rule sets, this added simplicity is the main advantage of PART [18].

10. ZeroR

ZeroR is the simplest classification method which depends on the target and ignores all predictors. ZeroR classifier simply predicts the majority category (class). Although there is no predictability power in ZeroR, it is useful for determining a baseline performance as a benchmark for other classification methods [19].

11. ADTree

Alternating Decision Tree is one of the classification method used in Machine learning which consists of decision nodes and prediction nodes. An instance is classified by an ADTree for which all decision nodes are true and summing any prediction nodes that are traversed. This makes it different from basic classification tree models that follow only one path through the tree [20].

12. J48

The J48 algorithm is WEKA’s implementation of the C4.5 decision tree learner. The algorithm uses a greedy technique to induce decision trees for classification and uses reduced-error pruning [21].

13. Random Forest

Random forest is an ensemble classifier which consists of many decision tree and gives class as outputs i.e., the mode of the class's output by individual trees. Random Forests gives many classification trees without pruning [22].

14. Simple Cart

CART is a recursive and gradual refinement algorithm of building a decision tree, to predict the classification situation of new samples of known input variable value. Breiman *et. al.*, 1984 provided this algorithm and is based on Classification and Regression Trees (CART) [23].

In our study, we have applied all the above classifiers on the 3 different cancer data sets and the results have been analyzed.

III. RESULTS AND DISCUSSION

The data sets have been submitted to a set of classification algorithms of Weka. We have used the 'Explorer' option of the Weka tool. Certain comparative studies conducted earlier [24][25][26][27][28] have shown that a particular algorithm has performed better on their data set and their conclusions however differ from each other. The studies either have used a very minimal set of classifiers or have used data sets that are not diverse resulting in an advantage or bias for a particular algorithm. Keeping that in mind, we have included a good number of classifiers in our analysis and used data sets that are diverse (in terms of size). The following sections describe the results obtained in our analysis.

A. Classification of Data set 1

The data set 1 is a small data set of micro-array gene expression data of Breast Cancer with 10 attributes and 286 instances. 5 out of the 14 algorithms got an accuracy of more than 95% where as the remaining algorithms reported the classification accuracy between 70% and 80%. Table 1 shows the results obtained in the analysis on data set 1.

The results in Table 1 indicate that the classifiers Multilayer Perceptron (ANN), IBk, KStar, NNge, and Random Forest performed better than the remaining algorithms. The Multilayer perceptron however took more time (11.68 secs) for classification whereas the remaining algorithms took almost less than 1 second. The kappa statistic for these 5 algorithms has been almost the same (~0.9). It should be noted that except IBk and KStar (Lazy classifiers), the classifiers among the better performers do not belong to the same group.

B. Classification of Data set 2

When a medium size data set (Lymphoma data set with 4,026 attributes and 45 instances) has been classified, the performance of the classifiers has significantly improved. All the classifiers (except ZeroR) reported more than 97% accuracy. Table 2 gives a summary report of the performances of all the classifiers when applied on Lymphoma data set.

10 out of 14 classifiers have got 100% accuracy as they correctly classified all the 45 instances. Though the number of instances decreased from 268 instances (from data set 1) to 45, the performance of the classifiers has been very good. The data set 2 has more number of attributes than data set 1 that resulted in better accuracy. The multilayer perceptron besides classifying all the instances correctly has however took a longer time (890.2 seconds) to get the results and hence, the accuracy of multi-layer perceptron can be ignored.

C. Classification of Data set 3

Finally, the large data set of Leukemia with 7,129 attributes and 34 instances has been used. The classifiers have achieved accuracies similar to the classification of medium size data set. However, the classifiers KStar and ZeroR underperformed. Rest of the classifiers achieved accuracies close to 100%. As expected, Multilayer perceptron took very long time to generate results. Table 3 gives a summary report of the performances of all the classifiers when applied on Leukemia data set.

TABLE I. Comparison of different classifiers using Breast Cancer Micro-array Gene Expression Data set with 10 attributes and 286 instances.

Classifier	Time Taken	Correctly Classified Instances	Incorrectly Classified Instances	Kappa statistic	Mean absolute error	Root mean squared error	Confusion Matrix
Bayes Net	0.02 Sec	217 (75.9%)	69 (24.1%)	0.3958	0.3018	0.4284	a b 173 28 41 44
Naive bayes	0.03 Sec	215 (75.2%)	71 (24.8%)	0.3693	0.3012	0.4278	a b 174 27 44 41
Multi layer Perceptron	11.7 Sec	276 (96.5%)	10 (3.5%)	0.9157	0.0482	0.1567	a b 197 4 6 79
Simple Logistics	0.87 Sec	218 (76.2%)	68 (23.8%)	0.32	0.3535	0.4183	a b 191 10 58 27
SMO	0.11 Sec	218 (76.2%)	68 (23.8%)	0.3615	0.2378	0.4876	a b 183 18 50 35
IBk	0 Sec	280 (97.9%)	6 (2.1%)	0.9491	0.0253	0.1053	a b 200 1 5 80
KStar	0 Sec	280 (97.9%)	6 (2.1%)	0.9494	0.0747	0.1399	a b 199 2 4 81
NNge	0.27 Sec	278 (97.2%)	8 (2.8%)	0.933	0.028	0.1672	a b 197 4 4 81
PART	0.21 Sec	229 (80.1%)	57 (19.9%)	0.4825	0.299	0.3866	a b 184 17 40 45
ZeroR	0 Sec	201 (70.3%)	85 (29.7%)	0	0.4183	0.457	a b 201 0 85 0
ADTree	0.08 Sec	223 (78.0%)	63 (22.0%)	0.4522	0.3659	0.4024	a b 175 26 37 48
J48	0.02 Sec	217 (75.9%)	69 (24.1%)	0.2899	0.3658	0.4269	a b 194 7 62 23
Random Forest	0.24 Sec	278 (97.2%)	8 (2.8%)	0.9326	0.1439	0.204	a b 193 8 5 80
Simple Cart	1.1 Sec	201 (70.3%)	85 (29.7%)	0	0.4177	0.457	a b 201 0 85 0

TABLE II. Comparison of different classifiers using Lymphoma Cancer Micro-array Gene Expression Data set with 4,026 attributes and 45 instances.

Classifier	Time Taken	Correctly Classified Instances	Incorrectly Classified Instances	Kappa statistic	Mean absolute error	Root mean squared error	Confusion Matrix
Bayes Net	0.27 Sec	45 (100%)	0 (0%)	1	0	0	a b 22 0 0 23
Naive bayes	0.24 Sec	45 (100%)	0 (0%)	1	0	0	a b 22 0 0 23
Multi layer Perceptron	890.2 Sec	45 (100%)	0 (0%)	1	0	0	a b 22 0 0 23
Simple Logistics	5.92 Sec	45 (100%)	0 (0%)	1	0.0641	0.0985	a b 22 0 0 23
SMO	0.18 Sec	45 (100%)	0 (0%)	1	0	0	a b 22 0 0 23

IBk	0 Sec	45 (100%)	0 (0%)	1	0.0213	0.0213	a b 22 0 0 23
KStar	0 Sec	45 (100%)	0 (0%)	1	0.0213	0.0213	a b 22 0 0 23
NNge	1.07 Sec	45 (100%)	0 (0%)	1	0	0	a b 22 0 0 23
PART	0.41 Sec	44 (97.8%)	1 (2.2%)	0.95	0.0425	0.1458	a b 22 0 1 22
ZeroR	0 Sec	23 (51.1%)	22 (48.9%)	0	0.4998	0.4999	a b 0 22 0 23
ADTree	0.82 Sec	45 (100%)	0 (0%)	1	0.0250	0.032	a b 22 0 0 23
J48	0.61 Sec	44 (97.8%)	1 (2.2%)	0.95	0.0423	0.1455	a b 22 0 1 22
Random Forest	0.17 Sec	45 (100%)	0 (0%)	1	0.1682	0.2078	a b 22 0 0 23
Simple Cart	1.73 Sec	44 (97.8%)	1 (2.2%)	0.95	0.0423	0.1455	a b 22 0 1 22

TABLE III. Comparison of different classifiers using Leukemia Cancer Micro-array Gene Expression Data set with 7,129 attributes and 34 instances.

Classifier	Time Taken	Correctly Classified Instances	Incorrectly Classified Instances	Kappa statistic	Mean absolute error	Root mean squared error	Confusion Matrix
Bayes Net	1.78 Sec	34 (100%)	0 (0%)	1	0	0	a b 20 0 0 14
Naive bayes	0.41 Sec	34 (100%)	0 (0%)	1	0	0	a b 20 0 0 14
Multi layer Perceptron	1313.87 Sec	33 (97.1%)	1 (2.9%)	0.9038	0.376	0.0267	a b 20 0 1 13
Simple Logistics	9.5 Sec	34 (100%)	0 (0%)	1	0	0	a b 20 0 0 14
SMO	0.19 Sec	34 (100%)	0 (0%)	1	0	0	a b 20 0 0 14
IBk	0.01 Sec	34 (100%)	0 (0%)	1	0.0278	0.0278	a b 20 0 0 14
KStar	0 Sec	20 (58.8%)	14 (41.2%)	0	0.5	0.5	a b 20 0 14 0
NNge	1.48 Sec	34 (100%)	0 (0%)	1	0	0	a b 20 0 0 14
PART	0.32 Sec	34 (100%)	0 (0%)	1	0	0	a b 20 0 0 14
ZeroR	0 Sec	20 (58.8%)	14 (41.2%)	0	0.4853	0.4922	a b 20 0 14 0
ADTree	1.5 Sec	34 (100%)	0 (0%)	1	0.0142	0.0145	a b 20 0 0 14
J48	0.52 Sec	34 (100%)	0 (0%)	1	0	0	a b 20 0 0 14

Random Forest	0.49 Sec	33 (97.1%)	1 (2.9%)	0.9386	0.1353	0.1955	a b 20 0 1 13
Simple Cart	2.0 Sec	34 (100%)	0 (0%)	1	0	0	a b 20 0 0 14

The results from the above 3 tables have been analyzed manually and they indicate that the classifiers work better when there is an increase in the number of attributes in the data set. But, none of the classifiers outperformed the others in terms of the accuracies. The classifiers Multilayer perceptron, IBk, NNge, and Random Forest have performed better on all 3 data sets. However, the performance of Multilayer Perceptron should not be considered because of the huge execution time taken by the classifier to generate results. The algorithm KStar reported around 58% accuracy for the large data set whereas the classifier ZeroR did not perform well on all 3 data sets. The remaining classifiers (except KStar and ZeroR) performed better on large data sets which are expected. The other statistics like kappa statistic and errors seem to be more or less same among all the classifiers in all three tests and are based on the accuracy of the prediction.

IV. CONCLUSION

This study focuses on finding the right algorithm for classification of data that works better on diverse data sets. However, it is observed that the accuracies of the tools vary depending on the data set used. It should also be noted that classifiers of a particular group also did not perform with similar accuracies. Overall, the results indicate that the performance of a classifier depends on the data set, especially on the number of attributes used in the data set and one should not rely completely on a particular algorithm for their study. So, we recommend that users should try their data set on a set of classifiers and choose the best one.

V. FUTURE WORK

We would like to develop web based software for performance evaluation of various classifiers where the users can just submit their data set and evaluate the results on the fly.

ACKNOWLEDGMENT

Bharath Kumar would like to thank Dr. DSVGK Kaladhar of GITAM University for his timely co-operation and guidance in completing the analysis. The authors greatly acknowledge the management and staff of GVVIT for their support and for providing lab facilities.

REFERENCES

[1] J. Han and M. Kamber, "Data Mining: Concepts and Techniques," Morgan Kaufmann, 2006.
[2] M.S. Chen, J. Han, and P.S. Yu. "Data mining: an overview from a database perspective," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 8, No.6, pp. 866 – 883, 2002.
[3] R. Chauhan, H. Kaur, and M.A. Alam, "Data Clustering Method for Discovering Clusters in Spatial Cancer Databases", *International Journal of Computer Applications* (0975 – 8887) Volume 10– No.6.
[4] R. Agrawal, T. Imielinski and A. Swami, "Data mining : A Performance Perspective". *IEEE Transactions on Knowledge and Data Engineering* (1993), 5(6):914-925.
[5] D. Grossman and P. Domingos. "Learning Bayesian Network Classifiers by Maximizing Conditional Likelihood". *Proceedings of the 21st*

International Conference on Machine Learning, Banff, Canada, 2004.
[6] J. C. Bailar, T.A. Louis, P.W. Lavori, and M. Polansky, "A Classification for Biomedical Research Reports," *N Engl J Med.*, Vol. 311, No. 23, pp. 1482-1487, 1984.
[7] R.S. Michalski, I. Mozetic, J. Hong, and N. Lavrac, "The Multi-Purpose Incremental Learning System AQ15 and its Testing Application to Three Medical Domains". In *Proceedings of the Fifth National Conference on Artificial Intelligence*, 1041-1045.
[8] A.A. Alizadeh, B. M.B. Eisen, R.E. Davis, C.M., et al "Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling." *Nature*, Vol 403, No. 3, pp. 503-511, 2000.
[9] T.R. Golub, D. K. Slonim, P. Tamayo, C. Huard et al, "Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring" *Science*, Vol. 286, pp. 531-537, 15 October 1999.
[10] Y. kumar and G. Sahoo, "Analysis of Bayes, Neural Network and Tree Classifier of Classification Technique in DataMining using WEKA", *International Journal of Information Technology in Computer Science*, vol. 4 (7), 359-369, 2012.
[11] D. Pedro and M. Pazzani "On the optimality of the simple Bayesian classifier under zero-one loss". *Machine Learning*, 29:103-137, 1997.
[12] N. Landwehr, M. Hall, and E. Frank. "Logistic model trees" *Machine Learning*, 59(1-2):161-205, 2005.
[13] Rosenblatt and X. Frank. Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. *Spartan Books*, Washington DC, 1961.
[14] S.S. Keerthi, S.K. Shevade, C. Bhattacharyya, K.R.K. Murthy "Improvements to Platt's SMO Algorithm for SVM Classifier Design". *Neural Computation*, 13(3), pp 637-649, 2001.
[15] D. Coomans, and D. L. Massart. "Alternative k-nearest neighbour rules in supervised pattern recognition: Part 1. k-Nearest neighbour classification by using alternative voting rules." *Analytica Chimica Acta*136 (1982): 15-27.
[16] A. Jemal, DVM, PhD; Freddie Bray, PhD *CANCER J CLIN* 2011;61:69-90
[17] B. Martin. Instance - Based Learning: Nearest Neighbour with generalisation, Department of Computer Science, University of Waikato, Hamilton, New Zealand
[18] E. Frank and I. H. Witten. "Generating accurate rule sets without global optimization". In *Proc 15th International Conference on Machine Learning*.
[19] I.H.Witten and E. Frank, "Data mining: practical machine learning tools and techniques" ISBN: 0-12-088407-0.
[20] Y. Freund and L. Mason. "The Alternating Decision Tree Algorithm." *Proceedings of the 16th International Conference on Machine Learning*, pages 124-133 (1999)
[21] J. R. Quinlan "C4.5: programs for machine learning" *Morgan Kaufmann Publishers Inc.*, San Francisco, CA, USA, 1993.
[22] L. Breiman, "Random Forests" *Machine Learning*. 45(1):5-32, 2001.
[23] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone. Classification and Regression Trees. *Wadsworth International Group*, Belmont, California, 1984.
[24] D.S.V.G.K. Kaladhar and B. Chandana "Data mining, inference and prediction of Cancer datasets using learning algorithms, *International Journal of Science and Advanced Technology*, Volume 1 No 3 May 2011
[25] B. Othman, Md. Fauzi, and T. M. S. Yau. "Comparison of different classification techniques using WEKA for breast cancer." *3rd Kuala Lumpur International Conference on Biomedical Engineering 2006*. Springer Berlin Heidelberg, 2007.
[26] A.Rohit. "Comparative Analysis of Classification Algorithms on Different Datasets using WEKA." *Int. Journal of Computer Applications*

54.13, 2012.

- [27] D. Delen, G Walker, and A. Kadam. "Predicting breast cancer survivability: a comparison of three data mining methods." *Artificial intelligence in medicine*, 34.2: 113-128, 2005.
- [28] H. Mark, et al. "The WEKA data mining software: an update." *ACM SIGKDD Explorations Newsletter*, 11.1: 10-18, 2009.

AUTHORS PROFILE

Mr. Gopala Krishna Murthy Nookala obtained his B.Sc degree from Andhra University, Visakhapatnam, India. He obtained his Master's degree in Computer Applications from Bharathiar University, M.Tech in Information Technology from Andhra University and also obtained Master of Philosophy in Computer Science from Alagappa University. He is pursuing his PhD in Computer Science & Engineering in Acharya Nagarjuna University since 2008. Currently, he is working as a Professor in the Department of Computer Science and Engineering, G.V.V.R. Institute of Technology, Bhimavaram, Andhra Pradesh. His specializations include Bioinformatics, Data mining, Networking, and Software Engineering. His current research interests are complexity reduced tumor classification using micro-array data-sets, usage of bioinformatics tools and optimization techniques for ant based networks using swarm intelligence.

Mr. Bharath Kumar Pottumutthu obtained his M.Sc in Bioinformatics from GITAM University, Visakhapatnam, India and B.Sc from Andhra University, Visakhapatnam. He is currently working as a Research Assistant in Centre for Bioinformatics and Software Development (CBRSD) at GVVR Institute of Technology, Bhimavaram. His current research interests include Comparative Genomics, Bioinformatics and Machine Learning.

Dr. NagaRaju Orsu received Ph.D. degree in Computer Science in the faculty of Engineering in 2011 from Acharya Nagarjuna University, Guntur, India and Masters Degree in Computer Science & Engineering from Acharya Nagarjuna University. He has a satisfactory and consistent academic track record. He is currently working as Head of the Department of Computer Science, Government Degree College, Macherla. He has occupied various academic positions like practical examiner, project adjudicator for various universities and colleges in and around in Andhra Pradesh. His current research is focused on Bioinformatics, Software Engineering, Image Processing & Database Management System. He has published several papers in National & International Journals.

Dr. Suresh B. Mudunuri received his PhD in Computer Science & Systems Engineering from Andhra University College of Engineering, Visakhapatnam, India in the year 2012. He received his BCA and MS degrees from Andhra University, and International Institute of Information Technology (IIIT), Hyderabad in 2002 and 2005 respectively. During 2005 and 2007, he worked on a bioinformatics research project in the Lab of Computational Biology (LCB) at Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad. Later, he has also worked as an Associate Professor in CSE departments at Aditya Engineering College, Surampalem and Vishnu Institute of Technology, Bhimavaram. He is currently working as Head of Centre for Bioinformatics Research & Software Development (CBRSD) at GVVR Institute of Technology, Bhimavaram. He is passionate to do research in the areas of Bioinformatics, Computational Biology and Data mining. He has published good number of quality research papers in international journals and conferences.

Path Based Mapping Technique for Robots

Amiraj Dhawan¹, Parag Oak, Rahul Mishra, George Puthanpurackal²
Department of Computer Engineering,
Fr. C. Rodrigues Institute of Technology, Vashi,
Navi Mumbai, Maharashtra, India

Abstract—The purpose of this paper is to explore a new way of autonomous mapping. Current systems using perception techniques like LAZER or SONAR use probabilistic methods and have a drawback of allowing considerable uncertainty in the mapping process. Our approach is to break down the environment, specifically indoor, into reachable areas and objects, separated by boundaries, and identifying their shape, to render various navigable paths around them. This is a novel method to do away with uncertainties, as far as possible, at the cost of temporal efficiency. Also this system demands only minimum and cheap hardware, as it relies on only Infra-Red sensors to do the job.

Keywords—Robots; map generation; navigation; AI planning; path planning;

I. INTRODUCTION

As the research in robotics continues and has become more involved in the past decade, the need for an intelligent navigation system has been realized. Many students and researchers around the world have devoted great amounts of time in developing new ways for an artificial agent to navigate both locally and globally. It has become essential for an agent to be able to navigate unknown environments autonomously. Hence when we talk about navigation we imply three things

- Mapping (the new environment)
- Path Finding
- Actuation

Several methods have been devised but most of them require expensive perception techniques using laser (LIDAR), SONAR or expensive cameras. Such methods are justified when time is crucial or when the environment is highly stochastic. But for simple indoor navigation something simpler can be sufficient.

The maps generated by existing mapping methods are usually discrete and hence can be made more informative and less stochastic either by increasing the resolution of the sensor or by increasing the input sampling rate. Unfortunately, this either increases the cost or increases the processing time and the storage space required.

A desired system would be the one that renders the required map resolution without unreasonable rise in resource consumption. We propose a method that will help an artificial agent map its immediate surroundings in an indoor setting. This method can work with cheap sensors (IR) and will generate a map which is technically discrete but logically continuous. It will allow motion planning with continuous paths.

The proposed system is such that irrespective of the environment complexity or the map resolution, the memory space requirement does not vary drastically and the sensor used will still be the same. The only dimension affected is time and that too varies only with environment complexity while it remains independent of the map resolution. The main steps involved are:

- **Generation of Boundaries:** The unknown environment is mapped and the boundaries are generated. As a result reachable and unreachable areas are demarcated. Also the detected objects are classified.
- **Shape identification:** The generated boundaries are analyzed. It is determined whether they are concave or convex in shape. This will supplement the next step.
- **Preparing path plan:** Finally, continuous paths are generated such that they occupy only the reachable areas of the environment. Path generation depends on the boundary shape.

The layout of the paper is as follows: Section II reviews few existing systems, followed by the proposed system in Section III. The proposed system is broken down into several sections explaining the environment, hardware, boundary generation, shape identification of the boundary, path planning etc. Section IV discusses several merits of the proposed system. Finally we conclude with Section V.

II. EXISTING SYSTEM

Robotic Mapping is a branch in Robotics, dealing with the application and study of map or floor plan construction by an autonomous robot. There are two types of internal representation:

•Metric: The metric framework is the most common for humans. It considers a 2D space in which it places the objects at their known coordinates. This model is very useful, but is sensitive to noise. Calculation of precise distance is also quite difficult.

•Topological: The topological framework only considers places and their relation. Usually, the metric stored is the distance separating the places. Finally a graph is created in which the nodes correspond to places and the arcs correspond to the paths. Some of the existing systems related to robot mapping are as follows:

A. Robot Localization and Mapping Using Sonar Data

Without any prior knowledge of the environment the robot generates a global map dynamically and computes the robot location for localization by correlating it with the local map. To create the local map, the robot uses range measurements in different directions from the sonar sensor. A servo holds the sonar sensor, so that 180 degree sweeps are possible [1]. This system uses 2-dimensional grid to provide a map of the robot's environment. Each grid stores the occupancy and certainty value obtained from the robot mapping algorithm which is later used for its localization.

B. Occupancy Algorithm

The system uses occupancy algorithm to create the map. As the robot explores its environment over time, it uses its range of range sweeping sensor values and current location to determine the occupancy of each grid. It classifies every grid as occupied, empty or unexplored. The occupancy of each cell in the pie is finally updated based on the previous value and the one inferred from the range readings which contributes to the generation of the grip map.

C. Autonomous topological modeling of an environment

This system uses the method of autonomous topological modeling and localization in a home environment using SONAR. The topological model is extracted from a grid map using cell decomposition and normalized graph cut. Autonomous topological modeling involves the incremental extraction of a sub-region without predefining the number of sub-regions [2]. The following are the important steps involved:

- Cell decomposition can systematically extract empty regions in the grid map and produce a roughly modeled graph structure for an empty environment.
- Normalized graph cut produces an effective clustering result by maximizing the similarity within clusters; this has low computational burden because of the cell decomposition process.
- Finally the topological map is constructed with an unknown number of sub-regions.

D. Robot Map Creation Algorithm using sensor data

This system describes an algorithm by which a robot can construct a map on the fly, and localize itself to the self-constructed map. In the given system, the robot begins by taking sonar readings, to generate a polar distance map of the robot's immediate neighborhood. These initial soundings are taken to be the robot's initial map [3].

Then the robot starts to move in some direction, stops at a particular location, and takes another sounding. The assumption is taken that there is no major changes in robot's environment, which contributes the best fit sounding map. The best fit returns a most likely location of the robot relative to the origin. The soundings are then shifted with respect to the robot's current location and used to modify the map. Several iterations of this cycle are performed until the robot has finished exploring.

III. PROPOSED SYSTEM

We plan to implement a system for construction of a navigation map and mapping the position and shape of the objects present in an area using a robot. The system would be limited to finite indoor geographical locations. The robot moves around in the enclosed area and at the end gives a map of the area reachable by the robot and the position of the objects present in the area with their estimated shape and size.

A. Environment

The environment needs to be a finite indoor geographical location like room or courtyards which are surrounded by some form of boundary (like walls). There can be two types of objects:

- Objects which are not in contact with the boundary of the area i.e. Extrinsic Objects
- Objects which are in contact with the boundary of the area i.e. Intrinsic Objects

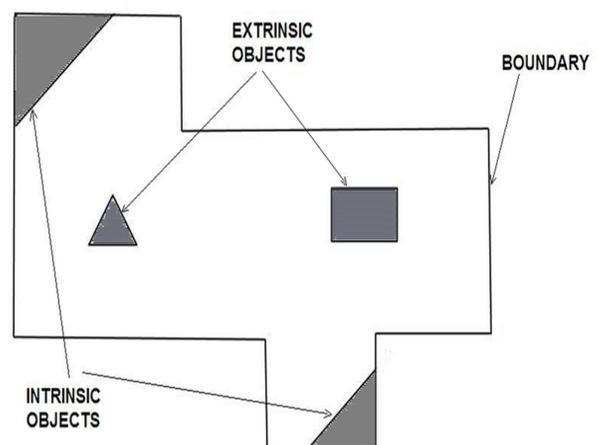


Fig.1. Boundary generated for Intrinsic and Extrinsic objects.

Intrinsic objects are considered as part of the boundary. Hence the system maps only the reachable area, the boundary of which can be the outline of an Intrinsic Object. For extrinsic objects, the system can report their positions and shapes accurately. Figure 1 shows these two types of objects.

B. Robot Hardware

The robot will consist of a four wheel or two wheel drive chassis with geared DC motors. The detection of the walls or boundaries can be done by using Digital Infrared Proximity Sensor's on all four sides of the robot. The sensors only detect the presence of the wall at a fixed distance d . More such sensors can be used on the diagonals to increase the mapping quality of the curved boundaries of the room or courtyards. In addition to infrared proximity sensors, the robot would also use wheel encoders to measure the distance traveled in a given time frame.

A Gyroscope sensor would be used to maintain orientation of the robot. The robot will be wireless capable for communication with a host system. The host system would collect data acquired by the robot's sensors and execute the resource heavy algorithm. The host system controls the actions of the robot using a command set. The host system can also be built on the robot itself if the mobile system can provide sufficient computing power for the algorithm to be executed in real time.

C. Generation of Boundaries

The first task is to map the reachable area by the robot which gives the boundaries of the area and the intrinsic objects. The intrinsic objects and the boundaries are indistinguishable.

The robot starts inside the closed area and randomly moves in any one direction until its proximity sensors give a high signal indicating an obstacle. This obstacle can be:

- An Intrinsic object or the boundary of the area: In which case the robot just moves along the detected obstacle and tries to alter its path whenever the sensor stops detecting the obstacle until the obstacle is again detected by the sensor. For example if the left side sensor detects an obstacle, and on moving along it, the sensor loses the obstacle, the robot would keep moving left, until the sensor again picks up the lost obstacle. The wheel encoders' & proximity sensors' data are sent to the host which checks if the robot has reached the same point from which it first picked up the obstacle signal. At this point the host sends a STOP command to the robot and generation of the boundaries of the area is done. Figure 2 below illustrates this scenario.
- An Extrinsic object: In which case the robot again moves along the detected obstacles and follows the same procedure as in Intrinsic or boundary scenario. But after the boundary generation, the robot finds itself outside the last boundary formed. Since it is a closed finite environment and the robot starts from inside the closed area, this is an impossible case and thus the obstacle found has to be an extrinsic object. Figure 3 shows the mapped extrinsic object.

Now the robot starts again by selecting a random direction and repeats the process but this time ignoring all the previous extrinsic objects mapped.

The boundary generation phase repeats every time when the extrinsic object scenario is detected and it ends successfully as soon as the system detects that the mapped obstacle is an intrinsic object or a boundary scenario.

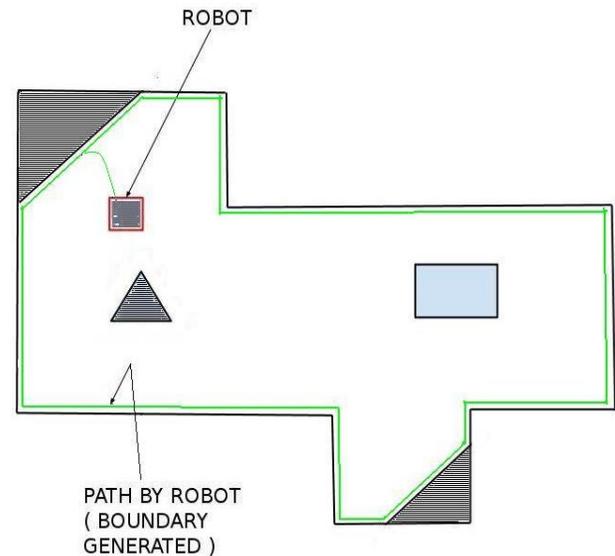


Fig.2. Generation of Boundary of an Intrinsic Object

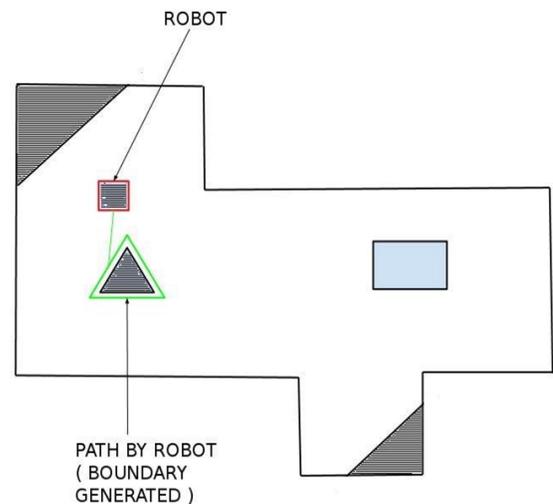


Fig.3. Generation of Boundary phase for Extrinsic Object Scenario

D. Shape Identification of the Generated Boundary

After the generation of the boundary, we try to identify the shape of the boundary. This phase is important since different strategies are applied to boundaries depending on its shape.

The types of shapes that are to be identified are Concave Polygon, Convex Polygon, Circular and Complex shapes

The complex shape is basically a mixture of Convex/Concave Polygon and Circular arcs.

The first step is to detect all the points where the robot took a turn of more than 1 degree in any direction. If no such points are detected then the shape of the generated boundary is Circular. Using these points, the shape can be identified as Concave or Convex Polygon easily using one of various Convex Hull algorithms with slight variations [4]. The preferred algorithm is Jarvis's March also known as Gift Wrapping algorithm due to its running time of $O(n)$ where n is the number of points [5]. According to Jarvis's March algorithm, if at any time the succeeding point vector goes in an anti-clockwise direction then the boundary is classified as a concave polygon and convex polygon otherwise. A complex shape also is passed as either a concave or a convex polygon. For this, the generated boundary is passed through Hough Transform for Circle. If a circle or curve is detected by the algorithm then it is a complex shape. In case a complex shape is detected, the shape is divided into convex/concave polygon and a circular part. These parts are then treated as individual reachable areas by the rest of the algorithm and the objects are mapped independently in these parts.

E. Preparing The Path Plan

To map the extrinsic objects, we first create a Path Plan using which the robot decides the movement in the reachable area. The path plan is basically a set of straight lines starting from one of the vertices of the reachable area to the edge opposite to the vertex. It also contains the angle at which the robot should start moving from the starting point of the line. We define a parameter which defines how accurately extrinsic objects are detected. This accuracy factor is denoted by the variable α . The accuracy factor α can help us deduce the approximate dimensions of the extrinsic objects which might not be mapped by the robot. Thus by customizing the accuracy factor, it would be possible to alter the completeness in the mapping of extrinsic objects. Also, increasing the accuracy would increase the time required for the mapping. We use different approaches for the different types of shapes of the reachable area mapped by the robot as follows:

- *Convex Shape*

The host system first finds all the vertices of the convex shape. Then from the first vertex a straight line is formed with any vertex of the opposite edge. This straight line is added into the path plan for the robot. Figure 4 shows the vertex P of a convex reachable area.

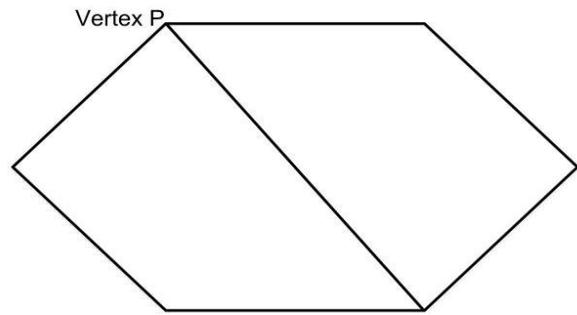


Fig.4. First path formed from Vertex P

Now we recursively form such straight lines from the same vertex P towards the opposite edges until we reach the other end vertex say R of the opposite edge. The endpoint of each straight line is at a distance α apart from the previous straight line. If the opposite edge length is not a multiple of α , then the last straight line is at a distance less than α from the straight line between P and R maintaining the accuracy of α .

From the reachable area we can deduce the vertical distance between the vertex P and the opposite edge. Using this vertical distance and α , we calculate the angle that each straight line makes with the opposite edge and store them in the path plan.

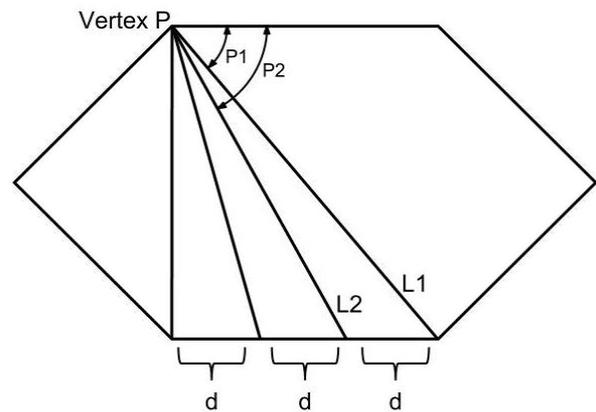


Fig.5. Straight lines from vertex P to the current opposite edge

We now take another edge opposite to the vertex P, if any, and redo the process considering this new edge. Figure 6 shows all the straight lines from the vertex P.

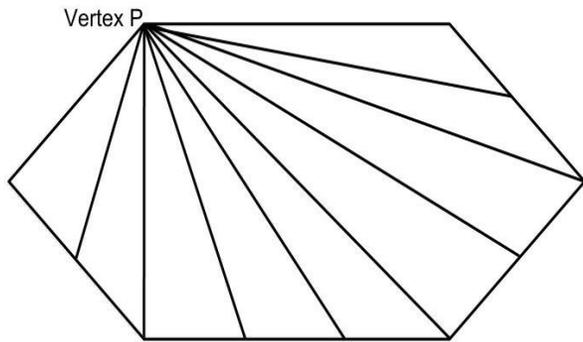


Fig.6. All the straight lines from vertex P towards the opposite edge

A sample entry in the path plan would consist of the starting vertex, angle made by the line to the boundary in the anti-clockwise direction and a line identifier. For example [P, P1, L1] and [P, P2, L2]. The path plan is complete when the above procedure is repeated for all the vertices. Straight lines in the completed path plan that coincides with any of the reachable area boundary are excluded from the path plan since the path was already used by the robot for mapping the reachable boundary. Figure 7 shows the complete path plan of the example convex shape.

• Concave Shape

For concave shape, we do the same procedure as we did in the convex shape with one addition. Every time a straight line is added into the path plan, it is first checked if the entire line falls with the boundary or not. If any part of the line is outside the reachable area i.e. the concave shape, it is not added into the path plan. Figure 8 shows the accepted and rejected straight lines from the Vertex P.

After every vertex is done, the final Path Plan for the sample Concave Shape is shown in Figure 9.

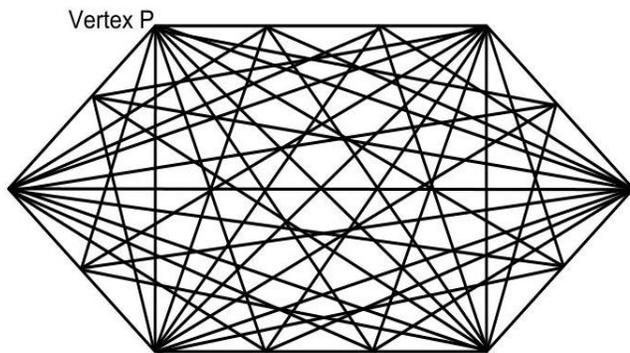


Fig.7. Path Plan for the sample convex shape

• Circular Shape

For Circular shape reachable area, the host system starts from a random point on the circular shape and makes a straight line passing from the center. The next point on the circular shape is selected at an arc length of α from the last point in a clockwise direction. The process ends when the new point has already been processed. If the circular shape is a semicircle, then the straight lines start from the boundary and ends at the center. But in a full circle, the straight lines start the boundary, passes through the center and ends on the other side of the boundary. Figure 10 shows the complete path plan for a circle shape [6].

F. Mapping an Extrinsic Object

After preparing the path plan, the host system commands the robot to move according to this plan. At all times, the proximity sensors are active and sends a stop signal to the robot if an obstacle is detected. The vertex, say C, closest to the robot's current position after mapping the boundary and getting the reachable area, is first selected, and all the paths stored in the path plan with starting vertex C are traversed. The path with the lowest angle is selected first and after reaching the opposite edge, the next path, which is closest to the first path, is traversed from the opposite edge to the vertex. During traversing on any path, if any of the proximity sensors send a Stop signal, the robot sends back the information to the host system indicating that it has detected an extrinsic object. The host system, who knows the current position of the robot, the position of the boundary and the direction of the proximity sensor, act differently as follows:

- The robot is sufficiently close to the boundary. In this case the robot is commanded to continue its path without any deviation.
- An extrinsic object, which the host system is already aware of, is detected. In this case, the robot is commanded to ignore the object and alter the path to move along the boundary of the object until it reaches its original path and then continues.
- Detects an extrinsic object, which the host system is unaware of. In this case the host system creates an entry of the extrinsic object. The robot is commanded to alter its path to move along the boundary of the object. It moves in the anti-clockwise direction, along the object if the right or the center proximity sensors detected the object else it moves clockwise. After moving along the object if the robot comes back on its actual path which it was following, the robot stops altering its path according to the object and continues along the original path, reaching the opposite edge of the starting vertex. Then it again takes the exact same path from the path plan but this time starts from the opposite edge and towards the vertex.

Again it picks up the same object and it alters its path along the boundary of this object going along the same side as it did in the previous traversal. The movement data from the wheel encoders and gyroscope data are used by the host system to map the extrinsic object detected, into the reachable area. After reaching the starting vertex, the entire object would be mapped since the robot went around the object for the entire 360 degrees during the 2 traverses of the same path.

When all of the paths in the path plan are traversed at least once, the mapping system ends. The Host system gives an image of the reachable area boundary, and the mapped Extrinsic objects in the environment.

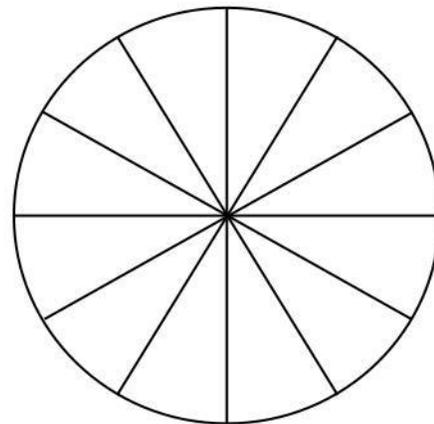


Fig.10. Sample path plan for a circle

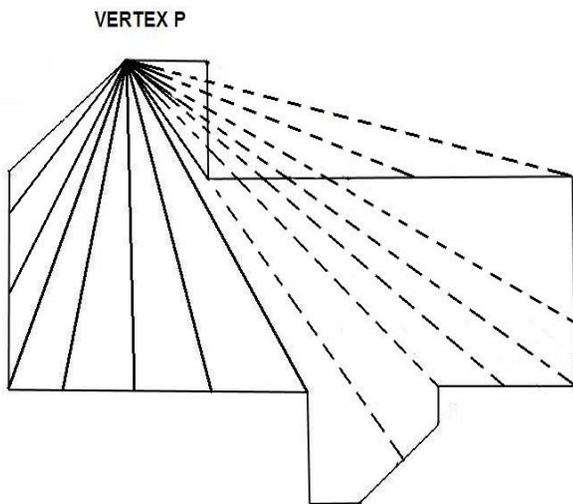


Fig.8. Accepted and rejected straight lines from Vertex P for a Concave shaped boundary

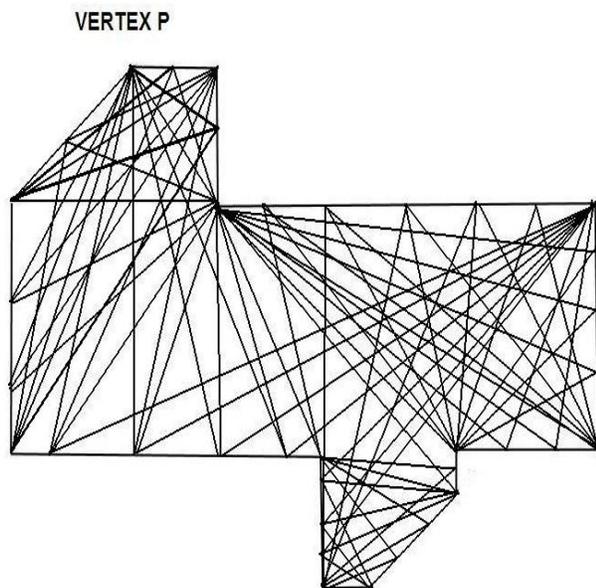


Fig.9. Shows the final path plan for the sample Concave boundary

IV. COMPARISON WITH EXISTING SYSTEM

One major advantage of our proposed system over existing ones is our system uses IR technology unlike SONAR, LASER technique as used by others. Besides this some points worth noting are:

1) *The storage space required by the map does not vary too much with change in the accuracy factor unlike current systems using LASER or SONAR. On the other hand, increasing accuracy of any of the existing systems, demands better sampling rates or higher resolution, hence requiring more storage and processing resources. The complexity in processing or using the map generated by proposed system remains independent of the accuracy.*

2) *Since the proposed system does not use any SONAR sensor, that use audio frequencies, to detect objects, the drawbacks posed by the existing SONAR technology such as with textured walls are not an issue. Also in case of mapping of an auditorium kind of a structure where sound gets absorbed and reflected in a particular way, the accuracy of the SONAR technology worsens, whereas in our proposed system the accuracy would still remain intact.*

3) *Mapping in areas like courtyards without any roofs is not possible using the SONAR technology since the sound wave loss is far too much. But our proposed system can be used in such conditions and areas since IR technology is unidirectional and the loss is minimal.*

4) *The object detection by the proposed system is robust. That is the user can be sure that an object of particular dimensions would be recognized by the robot, while in case of the current sound based systems, object detection depends on the extent of its reflective nature and probabilistic models used.*

5) *Extending the above point, the accuracy factor α described in the proposed system gives a measure to the dimension of the objects that would be detected by the robot. For example, if the robot is mapping in an area where a lot of small objects are expected to be encountered, the accuracy factor can be set to suit the environment i.e. detect the smaller*

objects. Thus changing the accuracy factor allows the robot to behave optimally in different surroundings.

6) The proposed system as compared to the existing system also gives more focus on detecting objects present at corners, this is due to the fact that one of the endpoint of each path line is a corner, so they have a denser network of path lines thus they provide more precision for mapping objects. Hence smaller objects present near the corners have better chances of detection, as compared to one in the interiors. It has applications in areas like autonomous cleaning robots where corners need to be emphasized.

V. CONCLUSION

The gist of the entire technique is as follows:

A. Generation of the boundaries:

The outermost boundary is mapped and objects within it are also detected. They are classified as follows:

- Extrinsic - these are isolated objects. Only some or none of the objects may be detected at this stage. The remaining ones are mapped after the path generation
- Intrinsic - these are objects attached to the outermost boundary and not differentiated from it.

B. Shape Identification of the Generated Boundary:

It is determined whether the boundaries are concave, convex, circular or complex. For this the turns made by the robot if greater than 1 degree are recorded as points.

- If there are no such points on a boundary, then it is a circle
- Else Jarvis March Algorithm is used to determine whether it is concave or convex boundary
- Complex boundaries are broken into their simpler components i.e. concave polygon or convex polygon and a circle (detected using Hough Transform)

C. Preparing the Path Plan:

Depending on the boundary shape the following steps are applied:

- Convex shape - For this the vertices and the edges (at distances determined by accuracy factor α) are connected to form paths.
- Concave shape - The approach is same as that for convex but requires removing the paths that do not lie entirely within the boundary.

- Circular shape - Multiple diameters act as paths. Each adjacent diameter is spaced apart by an angle determined by accuracy factor α .

D. Mapping extrinsic objects:

Here we are only mapping the objects that were not covered in the first step. The paths that have been generated are traversed. Whenever objects are encountered for the first time they are mapped.

We must keep in mind that the effectiveness of this technique will vary with the type of environment. For example more number of corners will require more paths to be traversed, in turn increasing time required for mapping.

The variable accuracy factor α is the essence of this technique, but it needs to be used carefully. With increase in the accuracy, i.e. with decrease in value of α , the number of paths traversed will increase. This means the system will require more time to complete the mapping process. The time required is directly proportional to the number of paths to be traversed.

We are trying to make major improvements by finding out ways to eliminate paths during planning, mainly redundant paths. As we eliminate more paths, the time required to map will also improve. This will make the technique useful in a wider range of environments.

ACKNOWLEDGMENT

We would like to express our sincere gratitude to Prof. H.K. Kaura, the Head of Department, Computer Engineering, our guide Ms. M. Kiruthika and Ms. Smita Dange for their support and guidance. Also we thank our friend Rohit Jha for his invaluable views.

REFERENCES

- [1] Vassilis Varveropoulos. *Robot Localization and Map Construction Using Sonar Data* [Online]. Available: <http://rossum.sourceforge.net>.
- [2] Jinwoo Choi *et al.*, "Autonomous topological modeling of a home environment and topological localization using a sonar gridmap" in *Springer Science+Business Media*, vol. 30, Issue. 4, pp. 351-368, May, 2011.
- [3] Jon Howell. *A Robot Map-Creation Algorithm* [Online]. Available: <http://www.jonh.net/~jonh/robots/mapping/submitted-paper.html>.
- [4] Eric Eilberg. *Convex Hull Algorithms* [Online]. Available: <http://www.denison.edu/academics/departments/mathcs/eilberg.pdf>.
- [5] Thomas H. Cormen *et al.*, "Finding the convex hull", in *Introduction to Algorithms*, 2nd ed. MIT Press and McGraw Hill, 2001, pp. 955-956.
- [6] Gary Bradsky and Adrian Kaehler, "Image Transforms", in *Learning OpenCV*, 1st ed. Mumbai, India: SPD, 2008, ch. 6, pp. 158-16

A New Optimization Algorithm For Combinatorial Problems

Azmi Alazzam and Harold W. Lewis III
Systems Science and Industrial Engineering Department
State University of New York at Binghamton
Binghamton, NY 13902, USA

Abstract—Combinatorial optimization problems are those problems that have a finite set of possible solutions. The best way to solve a combinatorial optimization problem is to check all the feasible solutions in the search space. However, checking all the feasible solutions is not always possible, especially when the search space is large. Thus, many meta-heuristic algorithms have been devised and modified to solve these problems. The meta-heuristic approaches are not guaranteed to find the optimal solution since they evaluate only a subset of the feasible solutions, but they try to explore different areas in the search space in a smart way to get a near-optimal solution in less cost and time. In this paper, we propose a new meta-heuristic algorithm that can be used for solving combinatorial optimization problems. The method introduced in this paper is named the Global Neighborhood Algorithm (GNA). The algorithm is principally based on a balance between both the global and local search. A set of random solutions are first generated from the global search space, and then the best solution will give the optimal value. After that, the algorithm will iterate, and in each iteration there will be two sets of generated solutions; one from the global search space and the other set of solutions will be generated from the neighborhood of the best solution. Throughout the paper, the algorithm will be delineated with examples. In the final phase of the research, the results of GNA will be discussed and compared with the results of Genetic Algorithm (GA) as an example of another optimization method.

Keywords—*meta-heuristic; optimization; combinatorial problems*

I. INTRODUCTION

Many optimization problems have been encountered in different domains of manufacturing and industry. Usually the optimization problem that needs to be solved is first formulated and all the constraints are given. The optimization problem mainly consists of an objective function and a set of constraints. The objective function can be in mathematical form or combinatorial form. Once the objective function of the optimization problem is formulated and all the constraints are defined, then the main issue is to solve this problem.

The solution is usually the best values of the variables or the best scenarios which can also be called the optimal solution. This optimal solution should give us the best performance or best fitness in terms of the objective function.

In most optimization problems there is more than one local solution. Therefore, it becomes very important to choose a good optimization method that will not be greedy and look only in the neighborhood of the best solution; because this will mislead the search process and leave it stuck at a local solution. However, the optimization algorithm should have a mechanism to balance between local and global search. An example of a two-dimensional function that has more than one local and global solution is shown in Fig.1 [1].

There are multiple methods used to solve optimization problems of both the mathematical and combinatorial types. In fact, if the optimization problem is simple or if the search space is small, then the optimization problem can be solved using conventional analytical or numerical procedures. However, if the optimization problem is difficult or if the search space is large, it will become difficult to solve the optimization problem by using conventional mathematics or using numerical induction techniques. For this reason, many meta-heuristic optimization methods have been developed to solve such difficult optimization problems. These include Genetic algorithm (GA), simulated annealing (SA), ant colony algorithm (ACA), and particle swarm (PS). Most of these meta-heuristic optimization problems are inspired by nature, biology, or environment.

The term meta-heuristic refers to a specific class of heuristic methods. Fred Glover first used this term and defined it as follows, “A meta-heuristic refers to a master strategy that guides and modifies other heuristics to produce solutions beyond those that are normally generated in a quest for local optimality.

The heuristics guided by such a meta-strategy may be high level procedures or may embody nothing more than a description of available moves for transforming one solution into another, together with an associated evaluation rule.” [2].

The meta-heuristic algorithms do not always guarantee an optimal solution. However, in most cases a near optimal solution can be obtained in much less time than the computational methods [3-4].

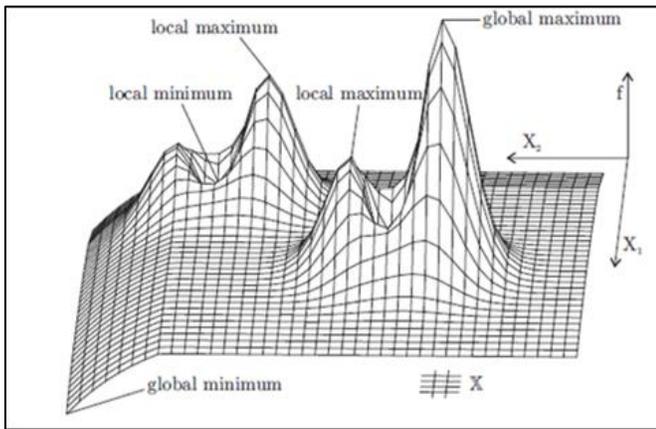


Fig. 1. Global and local optima of two-dimensional function [1].

The meta-heuristic algorithms can be classified based into different categories[5-6]:

1) Constructive and Improvement:

A constructive heuristic (also known as a greedy approach) usually constructs a solution from scratch based on a certain criteria. Some of the common constructive heuristics are nearest neighbor, multiple fragment and insertion heuristics [7]. An improvement or neighborhood search, which is usually known as a local search, attempts to improve the solution by exploring the neighborhood of the current solution [8]. The neighborhood of a solution is the set of solutions that are close to the current solution. The local optimal solution is the best solution in each neighborhood and the global optimum is the best solution with respect to the whole search space. An improvement or local search begins the search from a given solution, and then iteratively attempts to improve the solution quality by using move operators, the move operator is usually determined based on the neighborhood structure, and it aims to change (move) the solution to a newer solution in the same neighborhood but with a better fitness.

2) Single Solution and Population based approaches:

In the single based solution, a unique solution is first generated and then based on a certain move criteria, other solutions are generated. Some of the meta-heuristic methods that start with a single solution are: Tabu Search (TS) and Simulated Annealing (SA). Population based algorithms on the other hand start by generating a set of multiple initial solutions. Examples of those methods would be Genetic Algorithm (GA) and Ant Colony Algorithm (ACA).

The computational drawbacks of mathematical techniques and methods (i.e., complex derivatives, sensitivity to initial values, and the large amount of enumeration memory required) have forced researchers to rely on meta-heuristic algorithms based on simulations and some degree of randomness to solve optimization problems [9]. Although, these meta-heuristic approaches are not very accurate and they do not always give the optimal solution, in most cases they give a near optimal solution with less effort and time than the mathematical methods [10].

The meta-heuristic algorithms are general purpose stochastic search methods simulating natural selection and

biological or natural evolution [11]. Different meta-heuristic algorithms have been developed in the last few decades simulating and emulating different processes. Some of these meta-heuristic algorithms were inspired by the biological evolutionary processes; such as the evolutionary strategy (ES) [12], evolutionary programming [13-15], and the genetic algorithm (GA) proposed by Holland [16-17].

Some meta-heuristic algorithms emulate different animal behaviors; like the tabu search (TS) proposed by Glover [18], the ant colony algorithm (ACA) by Dorigo et al [19], Particle Swarm Optimization (PSO) [20], Harmony Search (HS) algorithm [21], Bee Colony Optimization (BCO) [22]. Other meta-heuristic algorithms were inspired by different physical and natural phenomena like the simulated annealing (SA) [23], and the Gravitational Search Algorithm (GSA) [24].

The distribution of publications which applied the meta-heuristics methods to solve the optimization problem in the past decade is shown in Fig.2. [25].

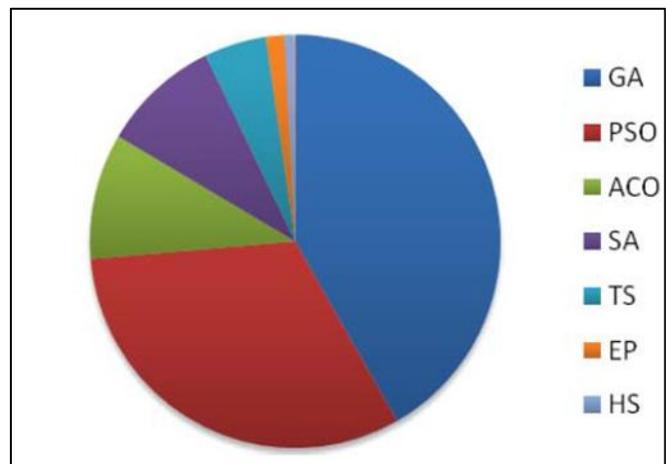


Fig. 2. Pie chart showing the publication distribution of the meta-heuristics algorithms [25].

In this paper we introduce a new optimization algorithm that can be applied to combinatorial problems. The new optimization problem is named Global Neighborhood Algorithm (GNA), and it is a population based and derivative free algorithm like other evolutionary optimization algorithms including Genetic Algorithm (GA), Ant Colony (ACA) and Evolutionary Strategy (ES). A set of randomly generated solutions from the entire search space are first generated and then the best of these solutions is chosen. After that, the algorithm will iterate, and in each iteration there will be two sets of generated solutions; one from the global search space and the other set of solutions will be generated from the neighborhood of the best solution. This paper starts with a background about optimization problems, then the methodology of the GNA algorithm is explained, and after that results for using this algorithm to solve the well-known Traveling Salesman(TSP) problem are also discussed.

II. METHOLODOLGY

The algorithm proposed in this paper is used to optimize combinatorial problems. The combinatorial problems could have more than one local and global optimal value within the

search space values. The proposed methodology will work to find the optimal value among these local optima by switching between exploration and exploitation. Exploration allows for exploring the whole search space. Exploitation allows focusing the search in the neighborhood of the best solution of generated solutions.

In order to explain the methodology of the GNA algorithm, assume we have discrete function that we need to optimize and let us say that we need to minimize this function (without loss of generality).

So the objective function we have is:

$$\min \quad g = f(x_1, x_2, \dots, x_n) \quad (1)$$

Where:

x_1, x_2, \dots, x_n are the different combinations of the solution sequence; we can think of these combinations as the city sequence in the TSP problem.

We need to find the optimal combination or solution (x_1, x_2, \dots, x_n) that will give the optimal (minimum) value for the objective function (g). In general, if each of the variables (x_1, x_2, \dots, x_n) can be chosen in (n_1, n_2, \dots, n_n) ways respectively, then if we want to enumerate all the possible solutions this will yield ($n_1 \times n_2 \times \dots \times n_n$) solutions. However, this process could take several hours or days depending on the size of the problem. Thus, using a meta-heuristic approach is better even if does not always give the optimal solution, but in most cases it will give a solution that is close to the optimal solution with less computational power.

According to the GNA algorithm, a set of (m) random solutions are first randomly generated from the set of all possible solution, where: (x_1, x_2, \dots, x_n) can be chosen in (n_1, n_2, \dots, n_n) ways.

The generated solutions will then look like:

$$(x_1^q, x_2^q, \dots, x_n^q) \text{ where } q = 1, 2, \dots, m$$

The fitness for the above solution will be evaluated and this can be done by substituting them in the objective function (g).

The solutions are then sorted according to their fitness obtained from the objective function:

$$f(s_1) < f(s_2) < f(s_3) < \dots < f(s_m)$$

$s_1 = (x'_1, x'_2, \dots, x'_n)$ is the solution sequence with best fitness.

The best combination (s_1) is then used as a good measure for the local optimal solution and it is also initially set as the best known solution.

In the next iteration, 50% of the (m) generated solutions will be generated near the best solution neighborhood by using a suitable move operator.

The other 50% of the (m) generated solutions will be still generated from the whole search space, and the reason for that is to allow for the exploration of the search space, because if we just choose the solutions close to the best solution we will only be able to find the local solution around this point, and since the function that need to be optimized could have more than one local optima, which might lead us to get stuck at one of these local optima.

Next, the best solutions from the above (m) solutions (50%, 50%) is calculated. The new value for the best solution is compared to best known solution and if it was found to be better it will replace it.

The procedure is then repeated until a certain stop criterion is met. This stop criterion can be a pre-specified number of iterations (t), or when there is no further improvement on the final value of the optimal solution we obtained.

The pseudo code for the GNA algorithm is shown in Fig.3.

```
Define objective function (g)
Initialize the values for all parameters: m,t
Generate (m) feasible solutions from the search space
Evaluate the fitness from the objective function (g)
Optimal solution= the best solution.
i=1
Do while i<t,++
    Generate 50% × m solutions from the
    neighborhood of the best solution

    Generate 50% × m solutions from the
    search space

    Find the best solution from the (m)
    generated solution

    If best solution is less (better) than
    optimal solution

        Optimal solution=best solution

    End If
End DO
```

Fig. 3. Psuedo Code For GNA Algorithm

The flowchart for the GNA algorithm is shown in Fig.4.

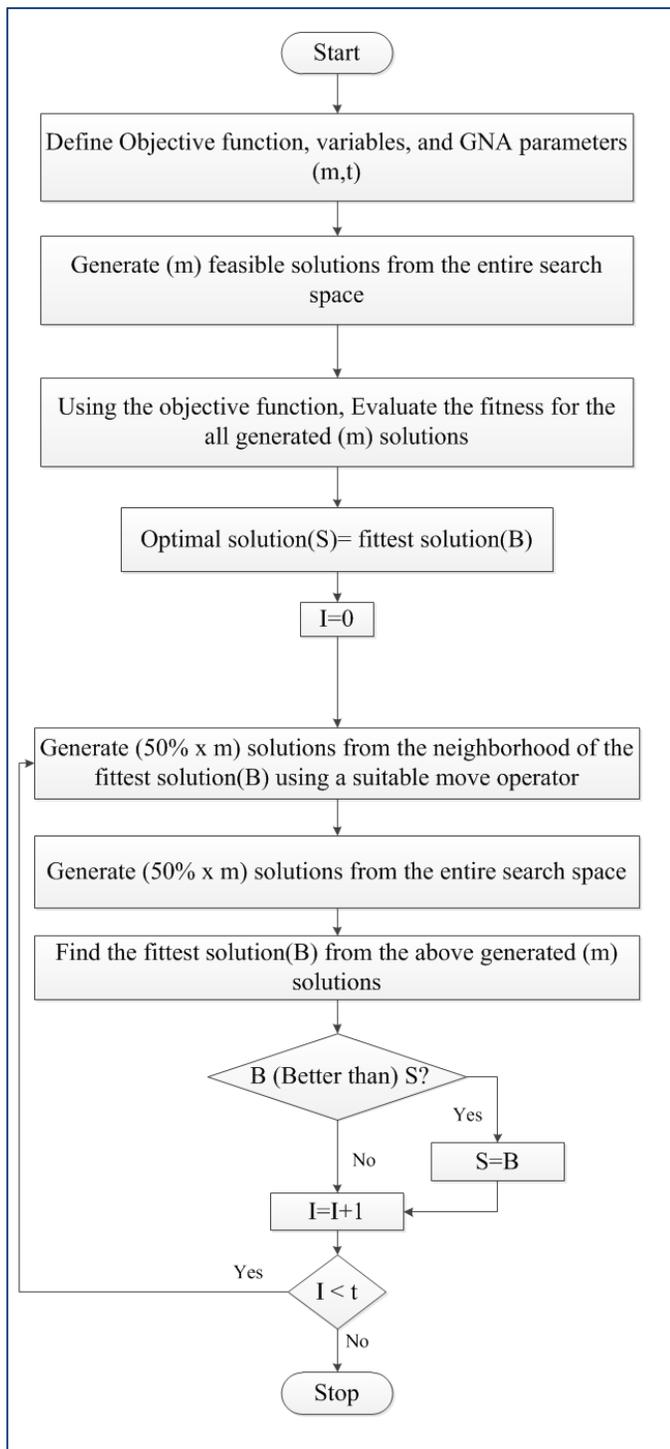


Fig. 4. Flow Chart For GNA Algorithm.

III. EXPERIMENTAL RESULTS AND ANALYSIS

The GNA algorithm was used to solve the Traveling Sales man Problem (TSP). The TSP problem consists of a number of cities that need to be visited one time for each, starting from one city and ending at the same city. In order to optimize the TSP problem, the optimal sequence of the different cities that gives the minimum cost (distance) of the tour length has to be

found. Thus, the objective (distance) function for the TSP is given by:

$$D(S) = d(C_N, C_1) + \sum_{n=1}^{N-1} d(C_n, C_{n+1}) \quad (2)$$

Where:

$D(S)$: The total distance for a sequence of N cities.

$d(C_n, C_{n+1})$: The Euclidean distance between the current city and the next city to be visited.

$d(C_N, C_1)$: The Euclidean distance between the last visited city and the first visited city.

To solve the TSP problem, we have to find the optimal sequence (S) that will give the shortest distance. If all the possible solutions are to be checked, then a total number of the combinations will be (N!) for asymmetric TSP or (N!/2) for the symmetric TSP. Obviously, if the number of the cities (N) is small then all the combination can be tried and a deterministic optimal solution can be found. However, if the number of the cities is large, then checking all the possible solution will take very long time and the complexity of the TSP problem will grow exponentially. For this reason, different meta-heuristic algorithms have been widely used to solve TSP problems.

In this paper, the GNA algorithm is used to solve a 29 cities TSP problem. The data were obtained from a real world problem that contained 29 cities in Bavaria, Germany; the source of these data is Zuse Institute Berlin [26]. The optimal solution for this problem is known and documented (2020). The GNA algorithm was implemented using MATLAB software, and the total number of solutions (m) generated at each iteration was 50.

At each iteration 25 feasible solutions were generated from the whole search space and the other 25 solutions were generated from the neighborhood of the fittest solution. The neighborhood move operator that was used in our case is the two-opt swap; where two cities were randomly chosen and swapped. The code was run for different times and at each time the obtained optimal solution and the run time were recorded. The stopping criteria used was 10000 iterations. The results for the GNA algorithm from the MATLAB code are shown in TABLE.1.

As it can be seen from TABLE.1, in the 10 different run times, we obtained a near-optimal solution that is very close to the known optimal solution.

TABLE I. MATLAB CODE OUTPUT FOR USING GNA TO SOLVE THE TSP PROBLEM.

RUN	Fitness of Optimal Solution	Number of iterations	Run Time
1	2026	10000	28.64
2	2022	10000	28.31
3	2026	10000	28.39

4	2033	10000	27.94
5	2046	10000	28.11
6	2033	10000	27.77
7	2022	10000	28.24
8	2020	10000	28.09
9	2047	10000	27.68
10	2026	10000	28.05

The results of the GNA algorithm were also compared to the Genetic algorithm (GA).

The parameters for the genetic algorithm were as the following:

- Generation size: 50
- Crossover probability: 90%
- Mutation probability: 10%
- Number of iterations: 10000

TABLE II. MATLAB CODE OUTPUT FOR USING GA TO SOLVE THE TSP PROBLEM.

RUN	Fitness of Optimal Solution	Number of iterations	Run Time
1	2132	10000	66.90
2	2295	10000	72.48
3	2066	10000	71.23
4	2191	10000	69.35
5	2084	10000	74.12
6	2097	10000	67.97
7	2178	10000	70.59
8	2226	10000	68.24
9	2320	10000	67.38
10	2218	10000	73.82

TABLE. II shows that the run time for the GA is more than twice the run time for the GNA, and the solution obtained by the GA is not always close to the known optimal solution. MINTAB software was used to conduct a statistical analysis between the means of the two optimal solutions obtained by both GA and GNA.

Statistical Analysis was conducted to test if there is a statistical difference between the average for each algorithm. A two- Sample T-Test was used for this purpose. The output from MINITAB is shown in Fig.5.

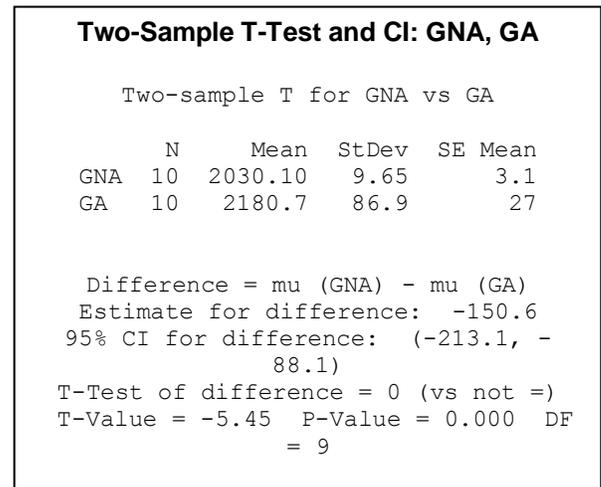


Fig. 5. MINITAB output for GNA and GA statistical analysis.

A 2-Sample t test showed that there is a significant difference between the optimal solution obtained from both GNA and GA , P-value= 0.000. The 95% CI for difference was (-213.1, -88.1).

Since the difference is always negative as indicated by the confidence interval, This shows us that on average the optimal solution is always higher for the GNA.

The difference in the means between the solutions obtained by the two algorithms is also clear in the Box plot, as shown in Fig.6. It can be seen the GNA outperformed the GA in terms of obtaining a near optimal solutions, and the run time to get this solution was also less. The reason for that is the selection process in GA is more complicated, and it requires sorting the solutions in each generation, whereas in our GNA, the best solution is always selected. Also, the method by which the solutions evolve in each iteration is much simpler in the GNA, unlike the GA that uses Crossover and mutation at each iteration; which makes it take longer time.

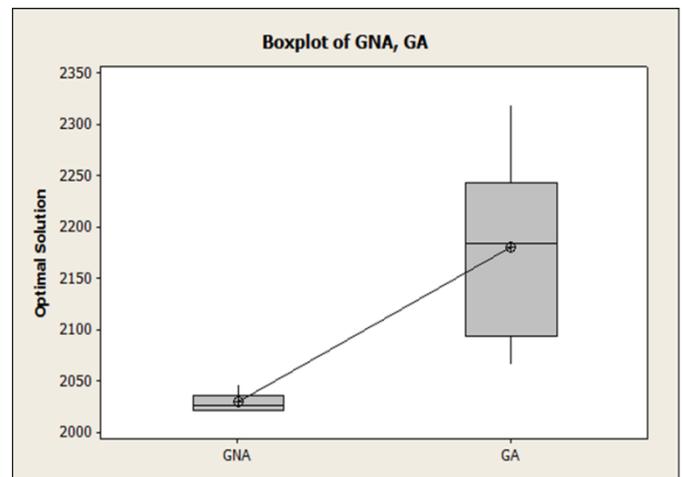


Fig. 6. Boxplot for GNA and GA output.

IV. CONCLUSION

In this paper, a new meta-heuristic optimization method was introduced and named Global Neighborhood algorithm (GNA). This optimization method is a population based algorithm; since it starts with generating a set of random solutions from the search space for the optimization problem. The proposed algorithm can be used to solve combinatorial optimization problems. These combinatorial problems are usually more difficult to solve than other continuous optimization problems. The methodology of this algorithm was elaborated and 29-cities TSP optimization problem was solved using the GNA. The TSP optimization problem was also solved using genetic algorithm (GA) and the results were compared to the GNA. Statistical analysis was conducted using MINITAB software, and it was found that the GNA showed better performance, and the results obtained were very close to the known optimal solution. Future studies can include different variants for the basic GNA algorithm to enhance the search power.

REFERENCES

- [1] T. Weise, Global Optimization Algorithms – Theory and Application, Germany: it-weise.de (self-published), [Online]. Available: <http://www.it-weise.de/>, 2009
- [2] F. Glover and M. Laguna. Tabu search. Kluwer Academic Publishers, 1997.
- [3] C. R. Reeves and J. E. Beasley, Modern heuristic techniques for combinatorial problems, McGraw-Hill, 1995
- [4] W. Wang, P. C. Nelson, and T. M. Tirpak, "Optimization of high-speed multistation SMT placement machines using evolutionary algorithms," IEEE Transactions on Electronics Packaging Manufacturing, 22(2), 137-146, 1995.
- [5] E. Silver, R. V. Vidal, and D. de Werra, "A tutorial on heuristic methods," European Journal of Operational Research, 5, 153-162, 1980.
- [6] S. H. Zanakis, J. R. Evans, and A. A. Vazacopoulos, "Heuristic methods and applications: a categorized survey," European Journal of Operational Research, 43, 88-110, 1989.
- [7] D. S. Johnson, "Local optimization and the traveling salesman problem," In Goos, G. and Hartmanis, J. (eds) Automata, Languages and Programming, Lecture Notes in Computer Science, 442, Springer, Heidelberg, 446-461, 1990.
- [8] E. Aarts and J. K Lenstra, "Local search in combinatorial optimization," Wiley, 1997.
- [9] K. S. Lee, and Z. W. Geem, "A new structural optimization method based on the harmony search algorithm," Computers and Structures 82(2004) 781–798, 2004.
- [10] K. S. Lee, Z. W. Geem, "A new meta-heuristic algorithm for continuous engineering Optimization," Comput. Methods Appl. Mech. Engrg. 194 (2005) 3902–3933, 2005.
- [11] M. G. Omran and M. Mahdavi, "Global-best harmony search," Applied Mathematics and Computation 198, 643–656, 2008.
- [12] I. Rechenberg, Cybernetic solution path of an experimental problem, Royal Aircraft Establishment, Library Translation no. 1122, 1965.
- [13] L. J. Fogel, A. J. Owens and M. J. Walsh, Artificial intelligence through simulated evolution, Chichester, UK: John Wiley, 1966.
- [14] K. De Jong, "Analysis of the behavior of a class of genetic adaptive systems", Ph.D. Thesis, Ann Arbor, MI: University of Michigan, 1975.
- [15] J. R. Koza, "Genetic programming: A paradigm for genetically breeding populations of computer programs to solve Problems," Report No. STAN-CS-90-1314, Stanford, CA: Stanford University, 1990.
- [16] D. E. Goldberg, Genetic algorithms in search optimization and machine learning, Boston, MA: Addison-Wesley, 1989.
- [17] J. H. Holland, Adaptation in natural and artificial systems, Ann Arbor, MI: University of Michigan Press, 1975.
- [18] F. Glover, "Heuristic for integer programming using surrogate constraints," Decision Science, 8(1):156–66, 1977.
- [19] M. Dorigo, V. Maniezzo and A. Coloni, "The ant system: Optimization by a colony of cooperating agents," IEEE Trans Systems Man Cybernet, 26(1), 29–41, 1996.
- [20] J. Kennedy and R. Eberhart, "Particle swarm optimization," IEEE International Conference on Neural Networks Perth, Australia, pp: 1942-1948, 1995.
- [21] Z. W. Geem, J. H. Kim and G. Loganathan, "A new heuristic optimization algorithm: Harmony search," Simulation, 76(2), 60, 2001.
- [22] S. Nakrani and C. Tovey, "On honey bees and dynamic server allocation in internet hosting centers," Adapt. Behav., 12(3-4), 223, 2004.
- [23] S. Kirkpatrick, C. Gelatt and M. Vecchi, "Optimization by simulated annealing," Science 1983, 220(4598), 671–80, 1983.
- [24] E. Rashedi, H. Nezamabadi-Pour and S. Saryazdi, "GSA: A gravitational search algorithm," Inform. Sci., 179(13): 2232-2248, 2009.
- [25] M. Khajehzadeh, M. R. Taha, A. El-Shafie and M. Eslami "A Survey on Meta-Heuristic Global Optimization Algorithms," Research Journal of Applied Sciences, Engineering and Technology 3(6), 569-578, 2011.
- [26] Z. I. Berlin, "MP-TESTDATA- the TSPLIB Symmetric Traveling Salesman Problem Instances," vol. 2010: Zuse Institute Berlin, 2010.

Gesture Controlled Robot using Image Processing

Harish Kumar Kaura¹, Vipul Honrao², Sayali Patil³, Pravish Shetty⁴,
Department of Computer Engineering
Fr. C. Rodrigues Institute of Technology, Vashi
Navi Mumbai, India

Abstract—Service robots directly interact with people, so finding a more natural and easy user interface is of fundamental importance. While earlier works have focused primarily on issues such as manipulation and navigation in the environment, few robotic systems are used with user friendly interfaces that possess the ability to control the robot by natural means. To facilitate a feasible solution to this requirement, we have implemented a system through which the user can give commands to a wireless robot using gestures. Through this method, the user can control or navigate the robot by using gestures of his/her palm, thereby interacting with the robotic system. The command signals are generated from these gestures using image processing. These signals are then passed to the robot to navigate it in the specified directions.

Keywords— Gestures; OpenCV; Arduino; WiFly; L293D.

I. INTRODUCTION

In today's age, the robotic industry has been developing many new trends to increase the efficiency, accessibility and accuracy of the systems. Basic tasks could be jobs that are harmful to the human, repetitive jobs that are boring, stressful etc. Though robots can be a replacement to humans, they still need to be controlled by humans itself. Robots can be wired or wireless, both having a controller device. Both have pros and cons associated with them. Beyond controlling the robotic system through physical devices, recent method of gesture control has become very popular. The main purpose of using gestures is that it provides a more natural way of controlling and provides a rich and intuitive form of interaction with the robotic system. This mainly involves Image Processing and Machine Learning for the system or application development. Beyond this, it also requires some kind of hardware for interfacing with the system for gesture control. There are some systems that have been developed in the same field using various techniques.

A. Existing Systems

Many systems exist that are used for controlling the robot through gestures. Some gesture recognition systems involve, adaptive colour segmentation [1], hand finding and labelling with blocking, morphological filtering, and then gesture actions are found by template matching and skeletonising. This does not provide dynamicity for the gesture inputs due to template matching. Another system uses machine interface device to provide real-time gestures to the robot [2]. Analog flex sensors are used on the hand glove to measure the finger bending [3], also hand position and orientation are measured by ultrasonics for gesture recognition [4]. And in another approach, gestures are recognized using Microsoft Xbox 360 Kinect(C) [5]. Kinect gathers the colour and depth information

using an RGB and Infra-Red camera respectively. This system though is not very cost effective.

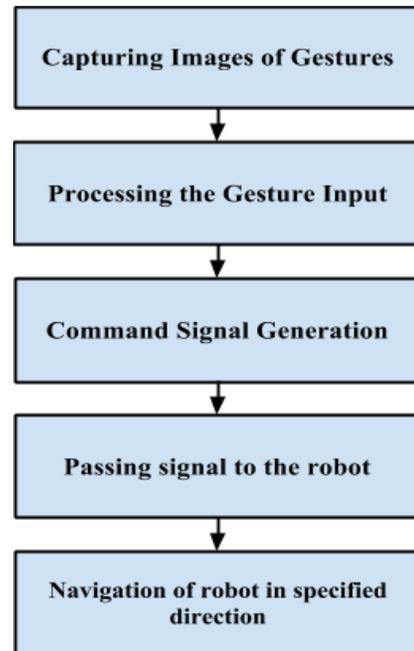


Fig. 1. Basic flow of the System

B. Proposed System

We propose a system, using which the user can navigate the wireless robot in the environment using various gestures commands. The main objective is to provide reliable and a more natural technique for the user to navigate a wireless robot in the environment using gestures.

In this system, user operates the robot from a control station that can be a laptop or a PC with a good quality in-built webcam or external webcam. This webcam is used to capture real time video stream of hand gestures to generate commands for the robot. Gesture commands are given using hand palm. Mainly two kinds of gestures are used which are explained further. Irrespective of the gesture technique used, robot is moved in all possible directions in the environment using four possible types of commands which are Forward, Backward, Right and Left. Image frame is taken as an input and processed using Image Processing. Processed image is then used to extract the gesture command. This gesture command can have one of the four possible commands as specified.

From this generated gesture command, signal is generated to pass the given command to the robot.

Generated signal is stored in the file at the control station. Wi-Fi shield on the robot accesses this file to transmit the signals from the control station to the robot. As soon as the Wi-Fi shield gets command from the control station, it is passed to the Arduino microcontroller. Arduino takes this signal as input from the Wi-Fi shield and generates some output signals that are passed to the motor driver. This output signal generation depends on the gesture input, for every four possible gesture input, different output signal is generated. The motor driver is used to drive the DC motors of the robot. It takes digital signals as the input from the Arduino and gives these signals as an output to the DC motors. Once a command signal is given to the robot, it continues to move in that direction till the next command is given or any obstacle comes in the path. Figure 1 shows basic flow of the system.

II. TECHNOLOGIES USED

A. C++ with OpenCV:

OpenCV (Open Source Computer Vision Library) is a library of programming functions mainly aimed at real-time computer vision, developed by Intel. The library is platform independent. It focuses mainly on real time image processing and computer vision. OpenCV is written in C and its primary interface is C with wrapper classes for use in C++. [6] Also

there are full interfaces available for Python, Java, MATLAB and Octave. It is used for recognition of gesture commands given by the user for the robot.

B. Arduino: Duemilanove

Arduino is an open-source electronics prototyping platform based on flexible, easy-to-use hardware and software. Arduino can sense the environment by receiving input from a variety of sensors and can affect its surrounding by controlling lights, motors and other actuators [7]. The microcontroller on the board is programmed using the Arduino Programming Language (based on Wiring) and the Arduino Integrated Development Environment (IDE) (based on processing). Arduino B. projects are stand-alone or they can communicate with software running on a computer (e.g. Flash, Processing, MaxMSP).

C. Wi-Fi Shield: WiFly GSX:

The Wi-Fi Shield connects Arduino to a Wi-Fi connection. In the proposed system we used a Wi-Fi Shield called WiFly. It takes power from the Vin pin of Arduino which is regulated at 3.3V and is provided to main Wi-Fi chip called RN-131C.

Communication between Arduino and WiFly shield is over SPI using Arduino's digital pins 10-13 (CS, MOSI, MISO, and SCLK respectively). [8]

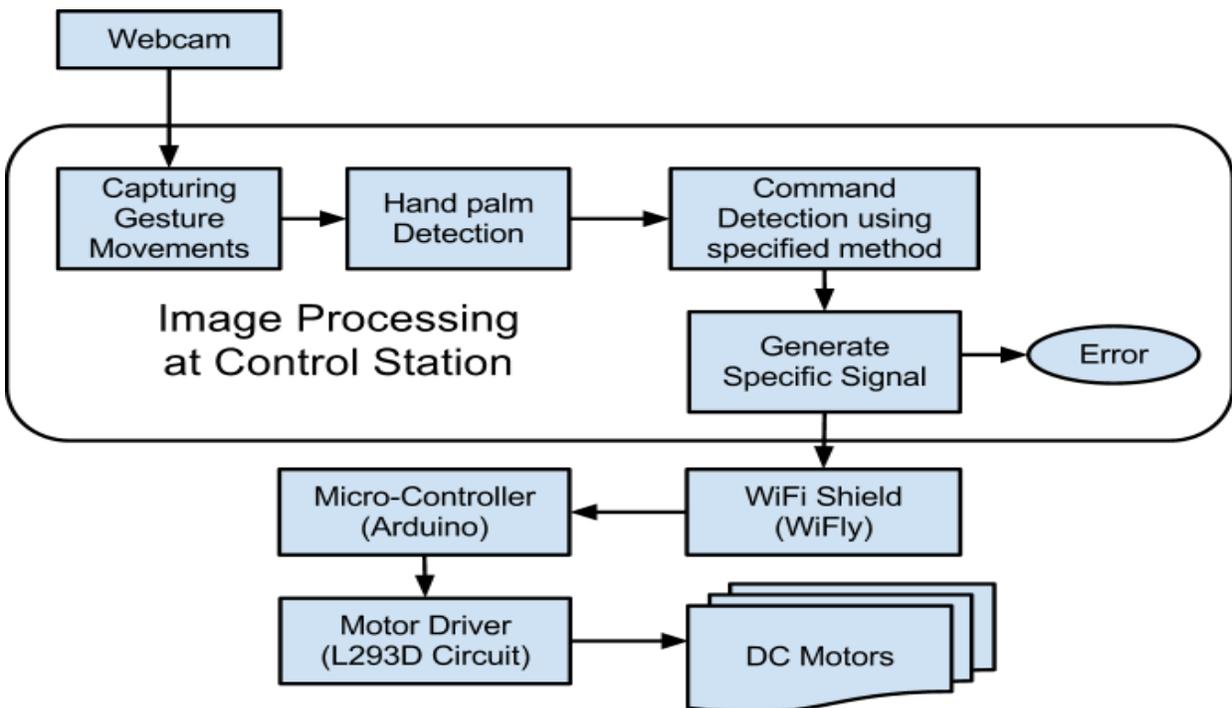


Fig. 2. Design of System



Fig. 3. Wi-Fi Shield (WiFly GSX)

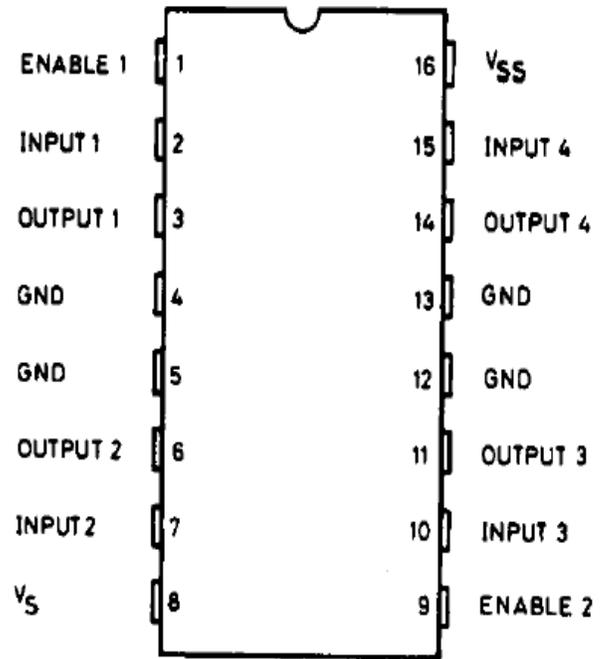


Fig. 5. Motor Driver (L293D)

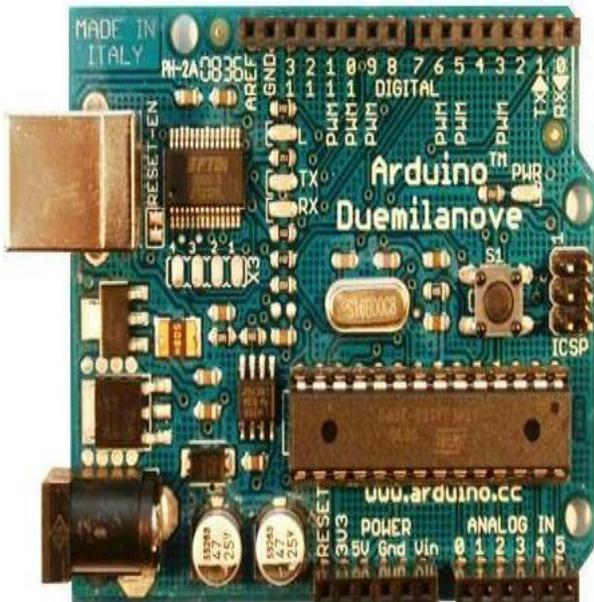


Fig. 4. Micro-Controller (Arduino-Duemilanove)



Fig. 6. DC Motor

D. L293D: Motor Driver

The L293D is a monolithic, integrated, high voltage, high current and has 4 channel drivers. Basically this means using this chip one could use at most four DC motors and provide power supplies of up to 36V. The L293D chip uses H-Bridge. The H-Bridge is typically an electrical circuit that enables a voltage to be applied across a load in either direction to an output [9], e.g. motor. This means that just reversing the direction of current leads to reversing of the direction of motor.

E. PHP:

PHP is an open-source server-side scripting language designed for Web development to produce dynamic Web pages. It is one of the first developed server-side scripting languages to be embedded into an HTML source document rather than calling an external file to process data. The code is printed by a Web server with PHP processor module which generates the resulting Web page.

III. DESIGN

Our design primarily focuses on gesture command Recognition. Command is generated at the control station and sent to the robot through Wi-Fi within the Wi-Fi range. The robot moves in the specified direction according to the specified command. The following section illustrates the steps carried out in the design.

A. Capturing Gesture Movements:

The end result of gesture recognition system is to generate a command and that is given to the robot. There are mainly four possible gesture commands that can be given to the robot (Forward, Backward, Right and Left). These gesture commands are given by the user in two ways. First method is based on the Finger Count and second is based on the direction given by the hand palm. Both methods involve recognition of the palm and process it further accordingly.

B. Hand Palm Detection:

This involves two steps to detect Hand Palm, which are as follows.

1) Thresholding of an Image Frame:

An image frame is taken as input through webcam. Binary Thresholding is then done on this image frame for the recognition of hand palm. Initially minimum threshold value is set to a certain constant. This value can be used to threshold an image and thus to increment the value till the system detects only one single blob of white area without any noise.

2) Drawing Contour and Convex Hull:

After thresholding an image frame, contour is determined and drawn on the thresholded white blob. Number of contours generated will be many due to noise. So threshold value is incremented and same procedure is applied till one contour is generated. A convex hull is drawn around the generated contour points. Convexity defect is used to generate the list of defects in the convex hull. This is in the form of vectors that provides all the defect parameters which includes defect point in the contour, depth of the defect, start point and end point of the line in the convex hull for which defect has occurred.

C. Command Detection using Specified Method

These 2 basic tasks are performed on the input frame. Depending on the kind of gestures further processing on an image frame is done. As stated earlier following are the two ways of controlling the robot through gestures:

1) Finger Count based Gesture Control:

The convexity defect's list provides depth parameters for every defect. Defects with the highest depth are easily extracted from the list. These defects are basically gaps between two consecutive fingers. Finger count is determined as one more than the number of defects. For example, for one defect, finger count is two. Thus each count from two to five can be specified as a command for robot to move in the particular direction.

2) Direction of Hand Palm Gesture Control:

In this method instead of giving a finger count, direction of the palm is used to determine the gesture command. Convexity defect specifies the depth point, start point and end point for the line of convex hull where the defect has occurred. Start and end points of the line specify fingertip. A depth point in the defect is the point in the palm where fingers meet. By comparing these point coordinate values, command is predicted. For this comparison the midpoint on the defect line is taken into consideration for calculations. This midpoint is then compared with the depth point. For a small difference in x coordinates and large difference in y coordinates, gesture command is predicted as forward and backward respectively. If the difference between midpoint and depth point is negative then gesture command is forward or else backward. The same thing is applied for a small difference in y coordinates and large difference in x coordinates. Instead of using all fingers of palm few fingers can also be used to specify the orientation.

D. Generate Specific Signal:

A text file forms an interface between Control Station and the Robot. Gesture command generated at control station is written in the file with a tagged word. This is done using C++ fstream library. A particular value is written into the file for given command. To make robot move in forward, backward, right and left direction simple values are written as 1, 2, 3 and 4 respectively. For example, Forward command is represented as 1. Value written in a file for 1 is sig1 where sig is a tagged word in the file. As soon as there is a real time change in gesture command, file is updated. Similarly for the next command to turn right, value is written as sig3 in the file where sig3 is a tagged word in the file.

E. Wi-Fi Shield: WiFly

WiFly works on HTTP port 80 for provided library. This library is used for communication of WiFly with the Control station through the Wi-Fi Hotspot. There is a simple web page present on the control station having PHP script. This is used for accessing the file and reading a command written in it.

WiFly on the robot connects itself with the Wi-Fi Hotspot. Then it continuously pings this web page on the control station. This is also called as Polling. From this, WiFly gets a command signal with a tagged word. This command signal is then transmitted to the Arduino without any modifications. Figure 3 shows WiFly GSX by Sparkfun.

F. Micro-Controller: Arduino (Duemilanove)

When Arduino gets command signal from the WiFly, it is having HTTP headers sent by web page with a tagged signal. Signal is read character by character and appended in the string. Every time after appending the character, Arduino checks for the tagged word in the string. For every iteration it checks the substring of tagged word at the end. If 'sig' is the tagged word in the signal, then program check for substring sig at the end of a string in each iteration loop. As it gets a tagged word at the end of the string, it terminates the loop for reading the signal character by character. Then it reads only the next character which is an actual command signal generated by the gestures. Depending on this command, signal is sent to L293D motor driver through the digital pins of the Arduino. Four digital pins of the Arduino is set as input to the L293D PIC, two pins on both sides. It has four possible methods as forward(), backward(), right(), left(). Depending on the command signal, a particular method is called for every iteration. Each method is defined with a specified command to make each digital pin HIGH or LOW. Figure 4 shows Arduino of Duemilanove.

G. Motor Driver: L293D Circuit

It has four input pins two on each side of the PIC. All these four pins are connected to the digital pins of an Arduino and four output pins are connected to DC motors of the Robot. Enable pins are used to enable input/output pins on both the sides of PIC and Vcc is used for supplying external power to the DC motors. Both the motors on the same side of the robot move in the same direction at a time. So positive ends of both motors are put in output pin 1 of PIC and negative ends of same motors are put into output pin 2, same thing is done for other side of PIC too. Vcc pin is used to provide external power supply to the DC motors. Figure 5 shows L293D PIC.

H. DC Motors:

A DC motor is mechanically commutated electric motor powered from direct motor (DC). The stator is stationary in space by definition and therefore so is its current. The current in the rotor is switched by the commutator. DC motors better suited for equipment ranging from 12V DC systems in automobiles to conveyor motors, both which require fine speed control for a range of speeds above and below the rated speeds. The speed of a DC motor can be controlled by changing the field current [10]. Fig 6 shows most widely used DC motor.

IV. IMPLEMENTATION

A. Capturing Gesture Movements:

Image frame is taken as input from the webcam on the control station and further processing is done on each input frame to detect hand palm. Figure 7 is an example of input frame. This involves some background constraints to identify the hand palm correctly with minimum noise in the image.

B. Hand Palm Detection

After capturing the image frame from the webcam, some basic operations are performed on this frame to prepare it for further processing of command detection. These operations are necessary for implementing both the techniques of gesture

control, following two main processes are done to detect hand palm.

1) Thresholded Image:

Image frame taken as input from webcam is thresholded starting from minimum thresh value till single contour is formed in an image, same is in the case of intensity based thresholding. In the Figure 7 two fingers are shown by the user as a gesture command having dark background. That image is thresholded so that only a single contour can be formed on it. This thresholding is done on the basis of intensity in the image, which neglects the dark background and thresholds the fingers. Thresholded image is shown in Figure 8.

2) Drawing Contour and Convex Hull:

As shown in Figure 9, after obtaining thresholded image two main things are done, drawing the Contour on the thresholded part and fitting this contour in the Convex Hull. Contour is drawn in the thresholded image by using function drawContour() in the library OpenCV. This is done on the intermediate image, this image is then passed for drawing the convex hull. This covers the whole contour by joining the minimal points to form Convex Hull. These two basic operations are performed on every image frame taken from the webcam, and then depending on the kind of gesture technique chosen by the user, further processing on the images is done. These two techniques are Finger Count based gesture control and Direction of Hand Palm Gesture Control.

C. Command Detection using Specific Method:

After completion of pre-processing of an input frame, further processing is done on the extracted image according to specified technique. These two methods of giving gesture commands are as follows.

1) Finger Count based Gesture Control:

In this technique of giving gesture commands, first defects in the convex hull are found out using function convexityDefects(). Convex hull is formed using minimal set of points, so it does follow the contour path always, this causes the formation of defects in the convex hull. convexityDefects() function gives information about these defects stored in the form of vector.

This vector has values of each defect point in the coordinate format, end points of the line of defect, and depth of the defect point. The non-linear structure of the contour causes hundreds of defects for the convex hull. But the defect formed due to gap between two fingers has largest depth value as compared to other depth values of defects. Minimum threshold value is considered for the comparison with all depth values thus the depth values more than this threshold value would be the depth point of a defect, which would be formed due to gaps between two fingers. This is well applicable for the count two to five. As for the 1 such depth, count of finger will be 2 i.e. depth plus one and so on. In the Figure 10 two fingers are shown having only one depth point. So for that one depth point count is two. Also in Figure 11 same techniques are used to show a different command. After processing the image two defects are found out that specifies count as 3. This technique has some disadvantages so another technique can be used.

2) Direction of Hand Palm:

In the previous technique of Finger Count Gesture Control, image having no large depth fails. For example, for finger count one, there is no such large depth so it is difficult to recognise, as it is count one or there is no such count. So counts from two to five are used as command signals. In this technique of gesture command, orientation of hand palm is taken into the consideration for recognition of the command signal. Thus an orientation of hand palm gives direction in which robot is to be moved. This command can be given with minimum two fingers or with the whole palm.

In the Figure 12, gesture command of this technique is given using two fingers. Orientation of these two fingers is towards right side, so the command signal for right is generated and passed to the robot. For deciding orientation of the palm two things are used, depth point and end points of the line of defect. These parameters have been found out by initial processing of the gesture command recognition. Midpoint of line of defect is calculated by taking average of two end points. Then this midpoint's coordinate position is compared with depth point's position.

Each time two main conditions are checked for each frame. For small difference between y-coordinate and large difference between x-coordinate, possible commands are Right or Left. In Figure 12, depth point and midpoint has small y-coordinate difference and large positive difference between them, so orientation of fingers is towards right is predicted correctly.

For the difference between y-coordinate of the point, which is negative, orientation is in the left direction specifying command as left. Similarly in the above Figure 13, the orientation of two fingers is down, that is a command given to the robot is backward. In this image frame, there is a small change in x-coordinate and large positive change in y-coordinate of the depth point and midpoint of the line of defect. So the orientation of the two fingers is downwards, specifying command as backward. Similarly for the difference between y-coordinate of the point as negative, orientation is in the upward direction specifying command as forward.

D. Generate Specific Signal:

After detecting gesture command specific signal value is generated, unique for every gesture command. This signal value is written in the file using C++ file reading/writing functions.

E. Wi-Fi Shield: WiFly

As soon as the command is generated on the control station, it is written in a file with tagged word with it. This file is read by WiFly after regular interval. As it is a wireless communication, so WiFly communicates with the control station using a hotspot where control station is also in the same network. This hotspot can be a wireless router or any kind of Smartphone having tethering facility. Both control station and WiFly is provided with an IP address by using this IP address WiFly accesses the information from the control station using provided IP address of control station.

F. Micro-Controller: Arduino- Duemilanove

WiFly is connected to the Arduino through stackable pins shown in Figure 14. When the process of communication starts, WiFly tries to connect itself with the hotspot. For that it requires ssid and passphrase of the hotspot. These are provided in the burned code of the Arduino. After forming a successful connection with the hotspot, WiFly tries to get the access of the control station, with the provided IP address of the control station and port number of HTTP port which is by default 80.

As soon as WiFly gets connected to the control station, it continuously pings a PHP webpage on the control station which has a small PHP script, which returns the value of signal command written in the file with a tagged word. These received signal values are then passed to Arduino, which extracts command and calls which are specified for that command. Arduino sends four digital signals as an input to the L293D motor driver.

G. Motor Driver: L293D

L293D motor driver circuit is shown in the Figure 15. It takes digital signal as an input from the Arduino and gives digital output to the DC motors of the robot. Power supply to the circuit is given by rechargeable batteries. In this system some rechargeable mobile batteries are used as power supply each of 3.7V. To provide more voltage to drive the motors, 2-3 such batteries are connected in series.

H. DC Motors in Robot:

This is the final end product of robot consisting of all the hardware WiFly, Arduino and L293D motor Driver circuit on the robot chassis having power supply provided by the rechargeable batteries. Four DC motors are connected to this robot chassis as shown in Figure 16. This is controlled through gestures by the user at control station.

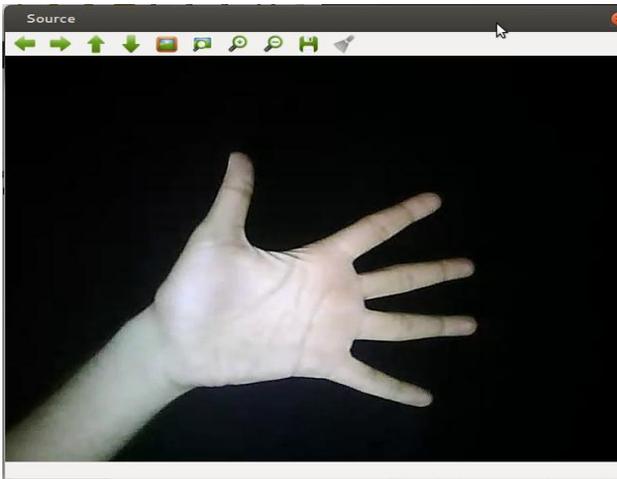


Fig. 7. Input Image Frame

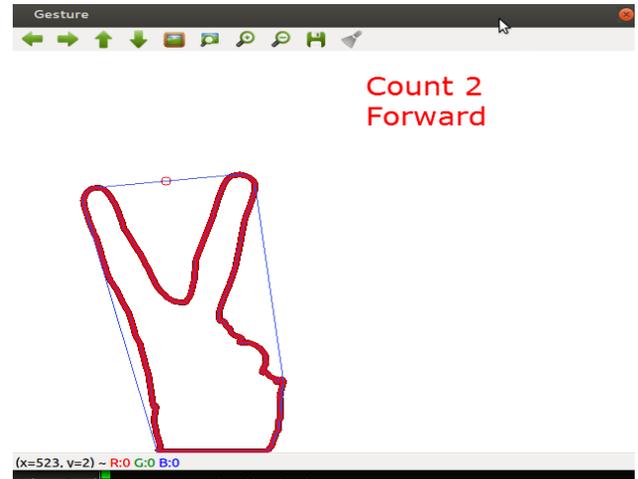


Fig. 10. Finger Count based Gesture Control I

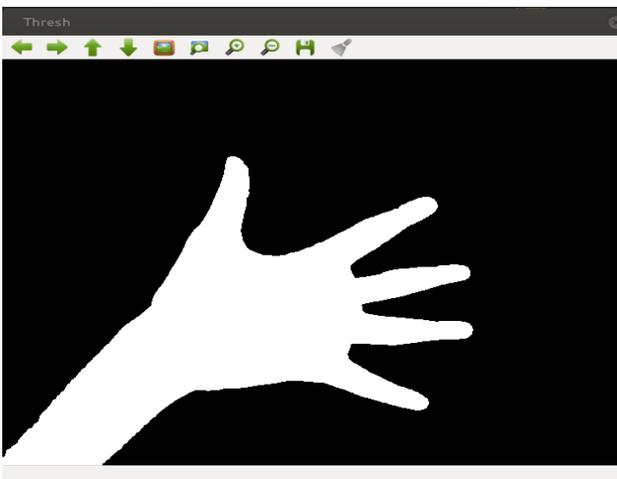


Fig. 8. Image Thresholding

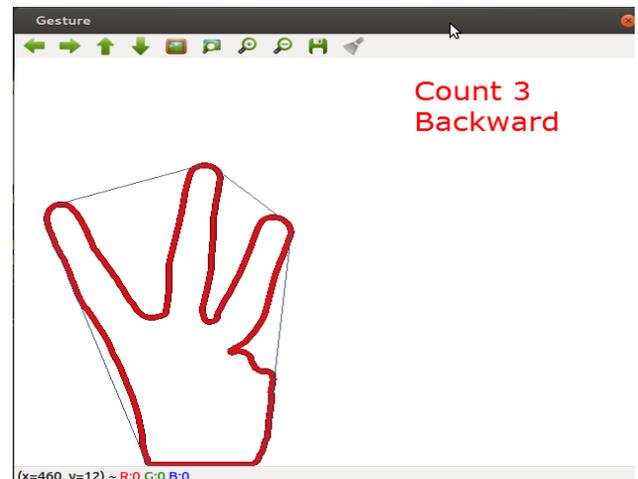


Fig. 11. Finger Count based Gesture Control II

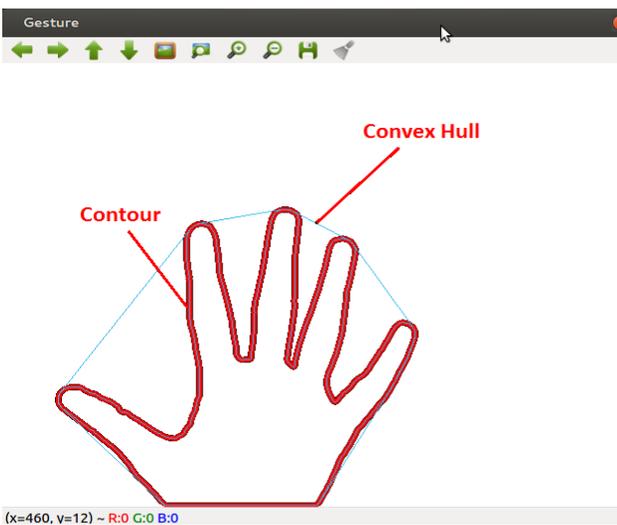


Fig. 9. Contour and Convex Hull

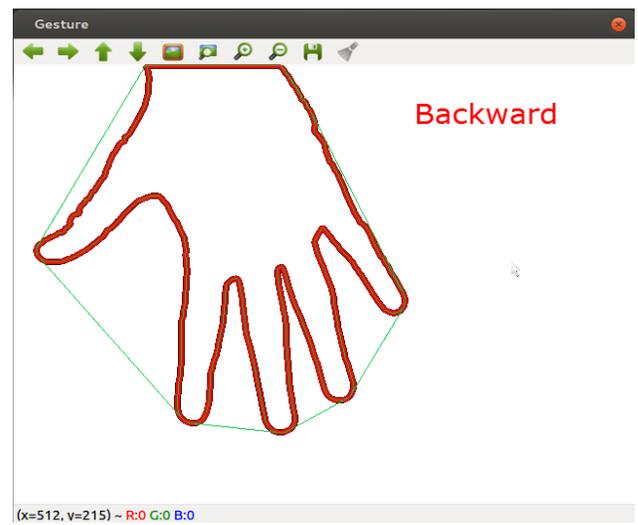


Fig. 12. Direction of Hand Palm I

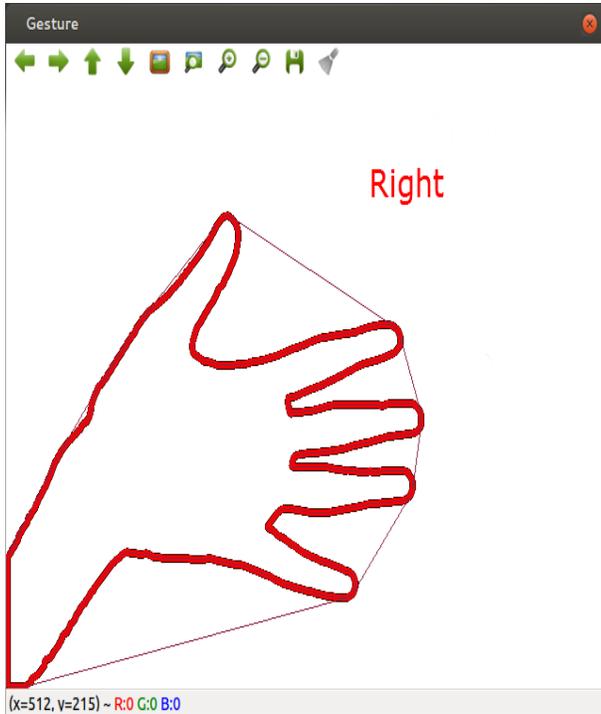


Fig. 13. Direction of Hand Palm II

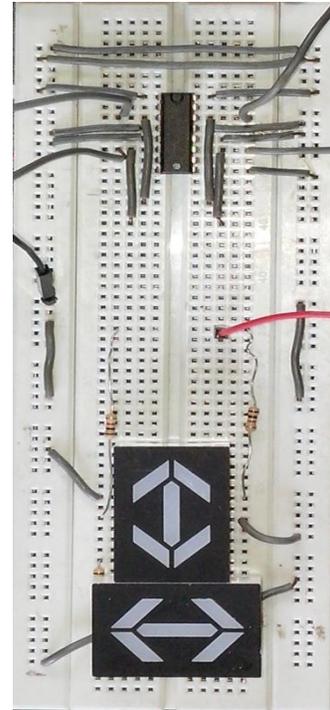


Fig. 15. Motor Driver- L293D Circuit

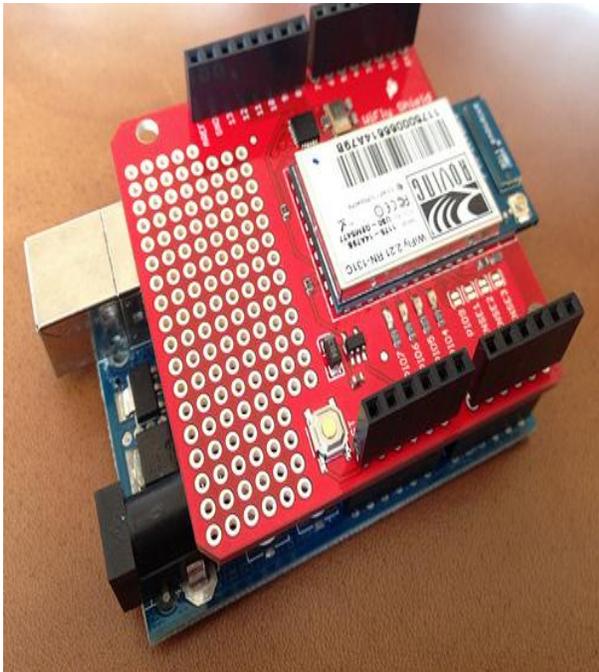


Fig. 14. WiFly interfaced with Arduino

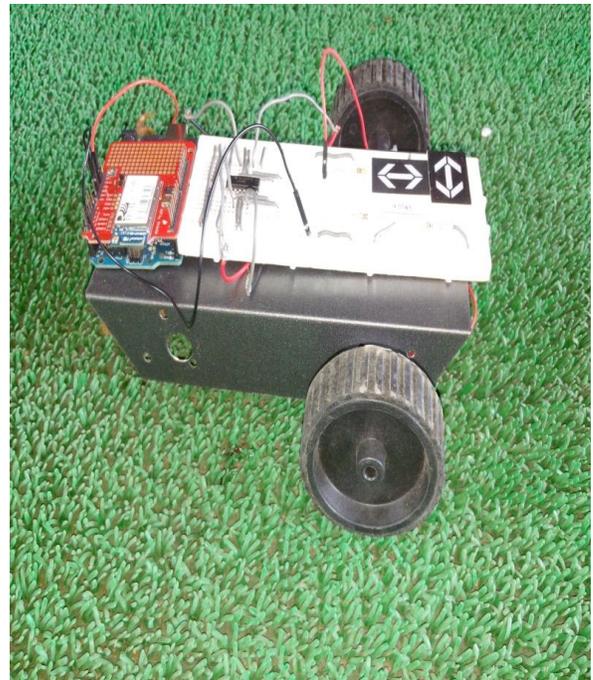


Fig. 16. Robot

V. COMPARISON WITH EXISTING SYSTEM

The major advantage of our system over other systems is that it provides real time palm gesture recognition, leading to an effective and natural way for controlling robots. Additional advantages are:

- This implemented system is much more cost effective than the existing systems. As it does not involve any hardware requirement or configuration, there is little or no cost for the system's implementation. Moreover, ordinary webcams on PCs or laptops can be used for capturing gesture inputs.
- As mentioned earlier, it does not involve any specific hardware for gesture inputs - a normal webcam on laptop or PC can be used for gesture recognition. This system can be installed on any of these usable devices for gesture recognition. This provides flexibility to the user and the system is portable.
- The implemented system takes real-time gesture inputs from the user, processes these gesture inputs to generate command signals. For both methods of gesture input, processing is done by a method provided by the system, and it does not involve template matching to identify the finger count or direction of palm. Each image frame is processed and a command is generated in real-time. This provides higher accuracy for gesture recognition.

VI. CONCLUSION

The Gesture Controlled Robot System gives an alternative way of controlling robots. Gesture control being a more natural way of controlling devices makes control of robots more efficient and easy. We have provided two techniques for giving gesture input, finger count based gesture control and direction of hand palm based gesture control.

In which each finger count specifies the command for the robot to navigate in specific direction in the environment and direction based technique directly gives the direction in which robot is to be moved.

At a time any one of the method can be used according to user's reliability, without using any external hardware support for gesture input unlike specified existing system. After gesture recognition command signal is generated and passed to the robot and it moves in specified direction.

VII. FUTURE SCOPE

Presently in the system a minimum threshold value is set up and using that value input image frame is thresholded to binary. This approach put some constraints on the background, so a dark background is used. To avoid this, colour based thresholding can be done. According to user's hand palm colour, image can be thresholded within a tight bound limit. So hand palm can be easily detected irrespective of the background.

Also the current implementation makes use of periodic polling from WiFly to the web server to access the command signal in real time. This method of periodic polling may overload the server. Thus, instead of using periodic polling, a persistent connection between server and WiFly can be set up using HTML5's Web Socket API. Through this connection, the web server can push a command signal to WiFly asynchronously thereby reducing the load on the server.

ACKNOWLEDGMENT

We would like to thank Ms. M. Kiruthika for supervising all activities and for lending her full cooperation, despite her busy schedule, in successful implementation of this system. We are also very grateful to Mr. Amiraj Dhawan and Mr. Rohit Jha for their guidance, help and support in the implementation.

REFERENCES

- [1] Chao Hy Xiang Wang, Mrinal K. Mandal, Max Meng, and Donglin Li, "Efficient Face and Gesture Recognition Techniques for Robot Control", CCECE, 1757-1762, 2003.
- [2] Asanterabi Malima, Erol Ozgur, and Mujdat Cetin, "A Fast Algorithm for Vision-Based Hand Gesture Recognition for Robot Control", IEEE International Conference on Computer Vision, 2006.
- [3] Thomas G. Zimmerman, Jaron Lanier, Chuck Blanchard, Steve Bryson and Young Harvill, "A Hand Gesture Interface Device", 189-192, 1987.
- [4] Jagdish Lal Raheja, Radhey Shyam, Umesh Kumar and P Bhanu Prasad, "Real-Time Robotic Hand Control using Hand Gestures", Second International Conference on Machine Learning and Computing, 2010.
- [5] Gesture Controlled Robot using Kinect <http://www.e-yantra.org/home/projects-wiki/item/180-gesture-controlled-robot-using-firebirdv-and-kinect>
- [6] OpenCV Library <http://docs.opencv.org/>
- [7] Arduino <http://arduino.cc/en/Guide/HomePage>
- [8] WiFly Library <https://github.com/sparkfun/WiFly-Shield>
- [9] L293D Motor Driver <http://luckylarry.co.uk/arduino-projects/control-a-dc-motor-with-arduino-and-l293d-chip/>
- [10] DC Motors www.globalspec.com/learnmore/motion_controls/motors/dc_motors