ISSN : 2165-4069(Online)
ISSN : 2165-4050(Print)

IJARAI

International Journal of Advanced Research in Artificial Intelligence

Volume 3 Issue 2

www.ijarai.thesai.org

A Publication of
The Science and Information Organization

IJARAI

INTERNATIONAL JOURNAL OF
ADVANCED RESEARCH IN ARTIFICIAL INTELLIGENCE

SAI

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!

Editor-in-Chief
IJARAI
Volume 3 Issue 2 February 2014
ISSN: 2165-4069(Online)
ISSN: 2165-4050(Print)

Editorial Board

Reviewer Board Members

CONTENTS

Paper 1: Static Gesture Recognition Combining Graph and Appearance Features
Authors: Marimpis Avraam
PAGE 1 – 4

Paper 2: Comparative Study of Feature Extraction Components from Several Wavelet Transformations for Ornamental Plants
Authors: Kohei Arai, Indra Nugraha Abdullah, Hiroshi Okumura
PAGE 5 – 11

Paper 3: Human Gait Gender Classification using 3D Discrete Wavelet Transform Feature Extraction
Authors: Kohei Arai, Rosa Andrie Asmara
PAGE 12 – 17

Paper 4: Method for Traffic Flow Estimation using On-dashboard Camera Image
Authors: Kohei Arai, Steven Ray Sentinuwo
PAGE 18 – 22

Paper 5: Predicting Quality of Answer in Collaborative Question Answer Learning
Authors: Kohei Arai, ANIK Nur Handayani
PAGE 23 – 26

Paper 6: Preliminary Study on Phytoplankton Distribution Changes Monitoring for the Intensive Study Area of the Ariake Sea, Japan Based on Remote Sensing Satellite Data
Authors: Kohei Arai, Toshiya Katano
PAGE 27 – 32

Paper 7: Using Unlabeled Data to Improve Inductive Models by Incorporating Transductive Models
Authors: ShengJun Cheng, Jiafeng Liu, XiangLong Tang
PAGE 33 – 38

Paper 8: Migration Dynamics in Artificial Agent Societies
Authors: Harjot Kaur, Karanjeet Singh Kahlon, Rajinder Singh Virk
PAGE 39 – 47

Static Gesture Recognition Combining Graph and Appearance Features

Marimpis Avraam
Computer Science Department
Aristotle University of Thessaloniki, AUTH
Thessaloniki, Greece

***Abstract*—In this paper we propose the combination of graph-based characteristics and appearance-based descriptors such as detected edges for modeling static gestures. Initially we convolve the original image with a Gaussian kernel and blur the image. Canny edges are then extracted. The blurring is performed in order to enhance some characteristics in the image that are crucial for the topology of the gesture (especially when the fingers are overlapping). There are a large number of properties that can describe a graph, one of which is the adjacency matrix that describes the topology of the graph itself. We approximate the topology of the hand utilizing Neural Gas with Competitive Hebbian Learning, generating a graph. From the graph we extract the Laplacian matrix and calculate its spectrum. Both canny edges and Laplacian spectrum are used as features. As a classifier we employ Linear Discriminant Analysis with Bayes' Rule. We apply our method on a published American Sign Language dataset (ten classes) and the results are very promising and further study of this approach is imminent from the authors.**

Keywords— Gesture Recognition; Neural Gas; Linear Discriminant Analysis; Bayes Rule; Laplacian Matrix

I. Introduction

The purpose of a gesture recognition process is to interpret gestures performed by humans [24]. The domain of such processes can vary significantly, from remote robot control, virtual reality worlds to smart home systems [25]. Most events that are recognized as gestures originate from the body's motions or state, but with most commonly origin the face or the hands.

A topic highly interwoven with gesture recognition and it is equally significant is the sign language recognition (SLR). A sign language is composed of three types of features. Manual features such as, hand shapes, their orientation and movement. Non-manual features are related to arms, the body and the facial expressions. Finally, finger-spelling that corresponds to letters and words in natural languages [19]. Specifically the non manual features are used to both individually form part of a sign or support other signs and modify their bearing meaning [26].

The ultimate purpose of sign language recognition systems is to build automated systems that are able recognize signs and convert them into text or even speech. This could ideally lead to a translation system to communicate with the deaf people. Finally, another challenge in this field is that there is no international sign language, thus every region defines its own and there is no imposed constraint that is should be based on the spoken language of the region [27].

In the literature have been proposed methods that utilize features extracted in real time and identify the gestures as they are performed. We, on the other, will focus on a series of static images. Our purpose is to demonstrate a novel combination of features, classic appearance-based characteristics and graph descriptors (e.g. adjacency matrix). Through experimentation, we concluded that these features are sufficient to describe in a discriminative way a gesture.

II. Related Work

Soft computing techniques include a plethora of well-established algorithms such as Self-Organizing Neural Gas (SGONG). In [17], the authors employ SGONG by generating an approximation of the hand topology. Based on a number of presumptions such as the hand should be placed vertical, etc they produce a statistical model to recognize the gestures. Moreover, they are able to identify the fingers individually by the number of neighbors that a neuron has. By combining Self-Organizing Maps (SOM) and Markov chains in [18] the authors propose a novel probabilistic classification model. SOMs are used to model the position of the gesture while via a quantization process they describe the hand direction. These features are cast in discrete symbols and then used to construct a Markov model for each hand.

In relation to image processing methods and edge detection methods specifically, in [23] Canny edges in order to produce thinned skeletons of the gestures. While in [19] the Canny edges are further processed by Hough transformation in order to conclude into the gesture features. Simpler but very effective, in [21] the authors utilize the difference between the detected (strong & weak) edges as features. An extensive work, based solely on edges as presented in [20], by which the edges are processed and clip in order to describe the fingers individually. Edges have also been proposed in conjunction with other low level features such as salient points and lines for sparse feature representation in [22].

Each of these proposals manages achieve very high recognition rates utilizing either graph or appearance-based features but not an explicit combination of the two. To our humble knowledge our proposal is one of the few, if not the only one, gesture recognition system (even testing on a simple dataset) that combines appearance-based features and graph properties simultaneously.

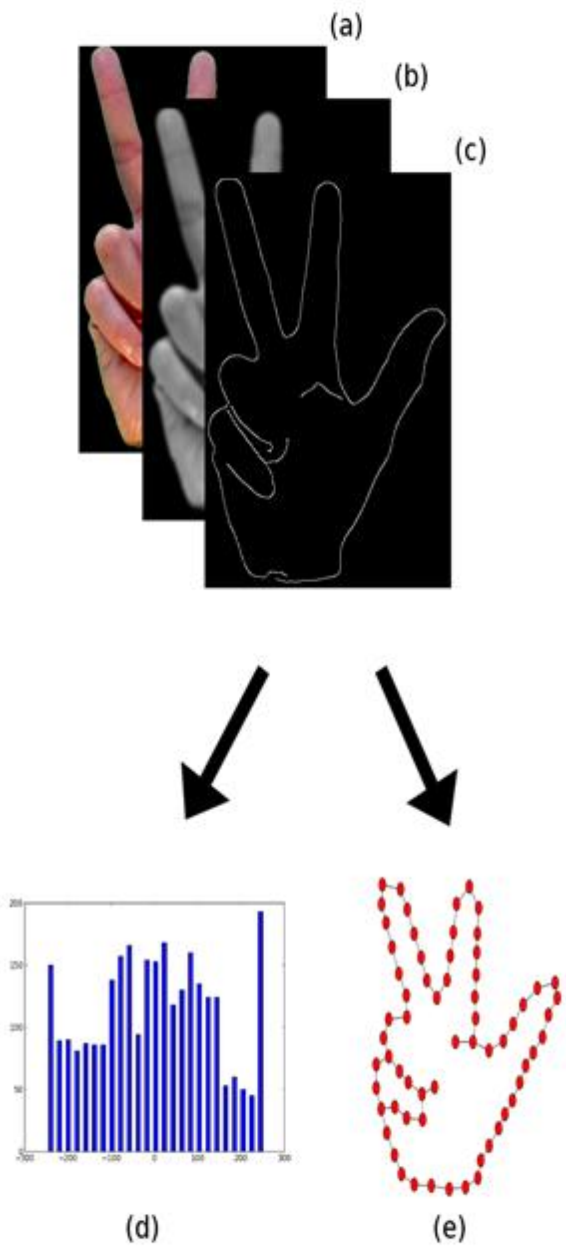

Fig. 1. (a) Orignal gesture image. (b) Grayscaled and passed with a Gaussian Blur filter. (c) Canny edges detected. (e) Histogram of 1 dimensional PCA projection. (f) Graph extracted by Neural Gas with CHL.

III. Proposed Method

As stated earlier our method utilizes appearance-based features as well as properties that are induced from graphs. On one hand the Canny edges provide a more general and rich description, without any specificity. On the other, the graph-related features are sufficient to describe the topology of a gesture, but fail to discriminate gestures with similar topographies. We concluded to this combination because these features complement each other.

In the following paragraphs we will provide a detailed description of our system's overview.

A. Image Preprocessing

Before anything else, we have to segment the hand from the image. Because of the nature of the specific dataset, a simple thresholding is sufficient; the background is just the black color.

The dimensions of the images in the dataset [4] are variable, because a number of transform operations have been applied. Some gestures have bee scaled along the X or Y axis while others have been rotated. The only property that we modify is the dimension of each image. We resize all images to $256x256$ pixels in order to restrict the possible variances of the generated graph topologies.

Finally, in order to restrict the Canny edge detector that we will user later on, we convolve a Gaussian blur kernel with the image as suggested in [16]. This way, we reduce the qualitative resolution of image and the edge detector will highlight only what we have issued to be, the most significant edges on the hand. The edges on the fingers will be greatly enhanced.

B. Appearance-based features

In Canny edge detector algorithm [3], initially the raw image is convolved with a Gaussian filter in order to reduce its susceptibility to noise. Then, utilizing two filters for each of the directions as in horizontal and vertical it computes the gradient [13]. Finally, it calculates the magnitude and the orientation of the gradient and with the usage a simple threshold technique it suppresses edges with low values.

For our purposes, after the edges are extracted we use PCA [14] and project them into one dimension and calculate the histogram.

C. Graph-induced features

1) Neural Gas with Competitive Hebbian Learning (NGCHL)

Martinetz and Schulten introduced the Neural Gas manifold learning algorithm in [5] and were inspired by Kohonen's Self Organizing Maps. The most notable difference between these two is that the first does not define explicitly a neighborhood. To overcome this issue an enhanced version of Neural Gas was proposed, integrating Competitive Hebbian Learning (CHL) [1, 2].

This new alternative approach follows closely the description of Neural Gas, that of a feature manifold learning method but through the generation of topology preserving maps. The learning function that is incorporated into the algorithm is described by a gradient descent as in the original algorithm. As in the original Neural Gas, the projection of a feature vector marks the adaptation of all neurons with a step size from the closest to the furthest. The key difference is that the Neural Gas with CHL also creates or destroys the synaptic links between the neurons. Each such connection is described by a synaptic strength. This indicator is used to decide whether or not a connection should be refreshed and kept or removed from the graph.

Below is a short description of the Neural Gas with Competitive Hebbian Learning algorithm. We must declare and set the following parameters: $T_{\max}$ (maximum number of iterations), α_T (final edge age), a_0 (initial edge age), e_T (final learning rate), e_0 (initial learning rate), λ_T (initial neighborhood reference), λ_0 (final neighborhood reference).

Create a number of neurons and assign them random positions and with no connections between them (the adjacency matrix is empty).

Step 1: Project $\vec{x}$ (drawn from R^n space) to our network.

Step 2: Sort the nodes based on their distances from $\vec{x}$ (keep the relative index position to k).

Step 3: Adapt the nodes' position vectors ($\vec{w}$) using:

$$N_{\vec{w}} \leftarrow N_{\vec{w}} + [N_{\vec{w}} * e(t) * h(k) * (\vec{x} - N_{\vec{w}})] \quad , \forall N \in GraphNodes$$

where: $h(t) = \exp\frac{-k}{\sigma^2}(t)$, $\sigma^2 = \lambda_0(\frac{\lambda_T}{\lambda_0})^{\frac{t}{T_{\max}}}$, $e(t) = e_0(\frac{e_T}{e_0})^{\frac{t}{T_{\max}}}$

Step 4: Connect the two closest neurons (according to their distance from $\vec{x}$). If already connected, refresh the age of the connection to 0.

Step 5: Increase the age by 1, of all edges connected with the best matching unit (the neuron closest to $\vec{x}$).

Step 6: Check for old edges in the graph and remove them if their age is exceeding the threshold given by the following formula:

$$\alpha(t) = \alpha_0(\frac{\alpha_T}{\alpha_0})^{\frac{t}{T_{\max}}}$$

The above function is a gradient descent. It uses the ranking of nodes in the aggregation of step 3. Increase the iteration counter and repeat from step 2 until the required criteria are met, usually $t = T_{\max}$.

2) Graph Spectra

Having obtained an undirected graph from NGCHL we are now able to apply specific linear algebra concepts. One of the most interesting notions in graph theory is the Laplacian matrix [6, 7].

In simple terms, given an adjacency matrix A and a degree matrix D, the Laplacian matrix is given by: $L = D - A$. These matrices have a numerous variations, properties and applications but these topics are out of the scope of this paper.

We would like to use the Laplacian matrix in order to extract its spectrum [8]. The computed Eigenvalues are very important and widely researched (especially the λ_2) [9] in the field of graph analysis and theory and can be employed when comparing graphs [12]. In our case we employ all of the Eigenvalues as a strong discriminative feature.

D. Linear Discriminant Analysis (LDA) and Bayes Optimal Classification

In our proposed method we use LDA to estimate the within class density (based on the assumptions that the distributions are Gaussians and the classes share a common covariance matrix), resulting into a linear classifier model [10, 11].

Based on Bayes' rule, we will assign a gesture to the class with the maximum posterior probability given the feature vector $\vec{X}$. So the conditional probability of class G given X is:

$$\Pr(G = k \mid X = x) = \frac{fk(x)\pi k}{\sum_{l=1}^{k} fl(x)\pi l}$$

Based on the earlier assumption of the normal distributions we can derive the following linear discriminant model:

$$g_i(x) = -\frac{1}{2}(x - \mu_i)^t \Sigma^{-1}(x - \mu_i) + \ln P(\omega_i)$$

This assigns a feature vector x into class i with the maximum g_i. Notice that the above equation contains a term that it is a Mahalanobis distance between the feature vector to and each the c mean vectors, and assign x to the category of the nearest mean [10, 15].

E. American Sign Language Dataset

We applied our proposed in the dataset proposed in [4]. It is a dataset containing 36 American Signs (ASL) each one of which is performed by five different persons. Each gesture is captured using a neutral-colored background (a green background), making the segmentation easier. This enables us to focus our research on the feature extraction and the classification methodology. In order to extend the dataset and provide more realistic (as "in the wild") poses, the original authors of the dataset, rotated each image in different angles and scales resulting into 2425 images.

In this paper, we focus on the ten classes corresponding to the numbering gestures from one to ten. Each class is composed of 65 samples.

IV. Experimental Results

We applied our proposal in the aforementioned dataset. Ten different gestures (classes), each one a gesture sign counting from zero to nine. The topographical network graph was set in 128 nodes. Finally as far the classification is concerned we followed a 5-fold cross validation scheme. For training we used 50 samples, while the rest 15 composed the testing set. Below we tabulate the results acquired.

TABLE I. CLASSIFICATION ACCURACY

Bayes Classification	
Gesture Class	***Accuracy Score***
0	93%
1	86%
2	73%
3	65%
4	100%
5	93%
6	86%
7	80%
8	86%
9	69%

[a.] Mean accuracy of a 5-cross validation procedure.

The results look very promising. Two gestures, namely the "3" and the "9" seem to perform poorly compared to others. The reason for this, is that both gestures are similarly signed (performed) resulting into also similar topographic graphs.

V. FUTURE WORK

We could possibly improve the performance by considering a method that respects the geodesic distances of the manifold, meaning that the edges would be regulated. This would greatly adjust the generated graphs and ultimately it would refine the quality of the extracted spectrums. We believe that this would result into greater results. In a more extensive work we will integrate a skin detection and segmentation algorithm possible based on a Gaussian Mixture Model that have been widely utilized in the literature. This will help build a more complete system. Another field that enjoys a lot of attention is the usage of wavelets (discrete) in order to extract low approximations of an image. This could result into a new system, in which the edges from the vertical and/or horizontal approximations are used.

VI. CONCLUSION

In this paper we presented the combination of the some properties and descriptors that can be extracted from a topology graph with Canny edges. A small topology results to a highly connected (or even fully) graph where all nodes are close to each other on the other hand, in a large topology all nodes spread to cover the topology and the connections are sparse, usually limited to their nearest neighbors. Based on the adjacency and degrees matrices we extract the Laplacian spectrum. This fusion yield very satisfactory results based on a five-fold cross validation procedure in a ten-class dataset.

REFERENCES

[1] Martinetz, T., & Schulten, K. (1994). Topology representing networks. Neural Networks, 7(3), 507-522.

[2] Martinetz, T. (1993). Competitive Hebbian learning rule forms perfectly topology preserving maps. In ICANN'93 (pp. 427-434). Springer London.

[3] Canny, J. (1986). A computational approach to edge detection. Pattern Analysis and Machine Intelligence, IEEE Transactions on, (6), 679-698.

[4] Barczak, A. L. C., Reyes, N. H., Abastillas, M., Piccio, A., & Susnjak, T. (2011). A new 2D static hand gesture colour image dataset for asl gestures.

[5] Martinetz, T., & Schulten, K. (1991). A" neural-gas" network learns topologies (pp. 397-402). University of Illinois at Urbana-Champaign.

[6] Newman, M. E. (2006). Modularity and community structure in networks. Proceedings of the National Academy of Sciences, 103(23), 8577-8582.

[7] Newman, M., Barabási, A. L., & Watts, D. J. (Eds.). (2006). The structure and dynamics of networks. Princeton University Press.

[8] Newman, M. (2009). Networks: an introduction. Oxford University Press.

[9] Mohar, B., & Alavi, Y. (1991). The Laplacian spectrum of graphs. Graph theory, combinatorics, and applications, 2, 871-898.

[10] Duda, R. O., Hart, P. E., & Stork, D. G. (2012). Pattern classification. John Wiley & Sons.

[11] Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. The Mathematical Intelligencer, 27(2), 83-85.

[12] Koutra, D., Parikh, A., Ramdas, A., & Xiang, J. (2011). Algorithms for Graph Similarity and Subgraph Matching.

[13] Semmlow, J. L. (2004). Biosignal and biomedical image processing: MATLAB-based applications (Vol. 22). CRC press.

[14] Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. Chemometrics and intelligent laboratory systems, 2(1), 37-52.

[15] Teknomo, Kardi. Discriminant Analysis Tutorial. http://people.revoledu.com/kardi/ tutorial/LDA/

[16] Hai, W. (2002). Gesture recognition using principal component analysis, multi-scale theory, and hidden Markov models (Doctoral dissertation, Dublin City University).

[17] Stergiopoulou, E., & Papamarkos, N. (2009). Hand gesture recognition using a neural network shape fitting technique. Engineering Applications of Artificial Intelligence, 22(8), 1141-1158.

[18] Caridakis, G., Karpouzis, K., Drosopoulos, A., & Kollias, S. (2010). SOMM: Self organizing Markov map for gesture recognition. Pattern Recognition Letters, 31(1), 52-59.

[19] Munib, Q., Habeeb, M., Takruri, B., & Al-Malik, H. A. (2007). American sign language (ASL) recognition based on Hough transform and neural networks. Expert Systems with Applications, 32(1), 24-37.

[20] Ravikiran, J., Mahesh, K., Mahishi, S., Dheeraj, R., Sudheender, S., & Pujari, N. V. (2009, March). Finger detection for sign language recognition. In Proceedings of the International MultiConference of Engineers and Computer Scientists (Vol. 1, pp. 18-20).

[21] MANISHA, L., SHETE, V., & SOMANI, S. (2013). SOFT COMPUTING APPROACHES FOR HAND GESTURE RECOGNITION. International Journal of Computer Science.

[22] Georgiana, S., & Caleanu, C. D. (2013, July). Sparse feature for hand gesture recognition: A comparative study. In Telecommunications and Signal Processing (TSP), 2013 36th International Conference on (pp. 858-861). IEEE.

[23] Barkoky, A., & Charkari, N. M. (2011, July). Static hand gesture recognition of Persian sign numbers using thinning method. In Multimedia Technology (ICMT), 2011 International Conference on (pp. 6548-6551). IEEE.

[24] Chaudhary, A., Raheja, J. L., Das, K., & Raheja, S. (2011). A survey on hand gesture recognition in context of soft computing. In Advanced Computing (pp. 46-55). Springer Berlin Heidelberg.

[25] Chaudhary, A., Raheja, J. L., Das, K., & Raheja, S. (2013). Intelligent approaches to interact with machines using hand gesture recognition in natural way: a survey. arXiv preprint arXiv:1303.2292.

[26] Cooper, H., Holt, B., & Bowden, R. (2011). Sign language recognition. In Visual Analysis of Humans (pp. 539-562). Springer London.

[27] Dreuw, P., Forster, J., Gweth, Y., Stein, D., Ney, H., Martinez, G., ... & Wheatley, M. (2010, May). Signspeak–understanding, recognition, and translation of sign languages. In Proceedings of 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (pp. 22-23).

Comparative Study of Feature Extraction Components from Several Wavelet Transformations for Ornamental Plants

Kohei Arai, Indra Nugraha Abdullah, Hiroshi Okumura
Graduate School of Science and Engineering
Saga University
Saga City, Japan

***Abstract*— Human has a duty to preserve the nature, preserving the plant is one of the examples. This research emphasis on ornamental plant that has functionality not only as ornament plant but also as a medicinal plant. Purpose of this research is to find the best of the particular feature extraction components from several wavelet transformations. It consists of Daubechies, Dyadic, and Dual-tree complex wavelet transformation. Dyadic and Dual-tree complex wavelet transformations have shift invariant property. While Daubechies is a standard wavelet transform that widely used for many applications. This comparison is utilizing leaf image datasets from ornamental plants. From the experiments, obtained that best classification performance attained by Dual-tree complex wavelet transformation with 96.66% of overall performance result.**

Keywords—wavelet transformation; shift invariant; rotation invariant; feature extraction; leaf identification.

I. INTRODUCTION

Oxygen is an essential part for all living things in the world. Plant plays an important role to produce oxygen and supply it for their sustainable life. The cycle between human and the plant is the interesting one. Carbon dioxide as the output of human respiratory is needed by plant for photosynthesis activity. Then, this activity is resulting oxygen which vital for human.

According to this cycle, human supposed to preserve the plant to maintain availability of oxygen. Based on International Union for Conservation of Nature and Natural Resources, the number of identified plant species which consist of Mosses, Ferns and Allies, Gymnosperms, Flowering Plants, Green Algae, Red Algae is about 307.674 species. [1].

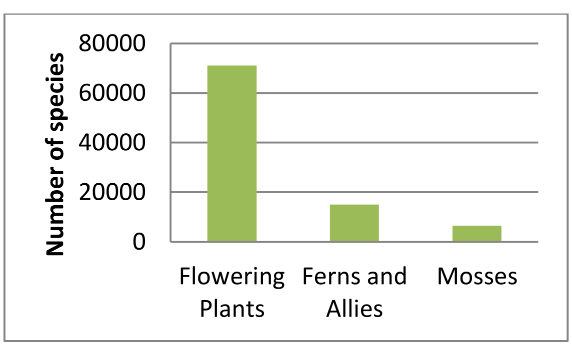

Fig. 1. Column charts number of unidentified plant species

On the other side, the approximate number of unidentified species is 80.500 species. It consists of Flowering Plants with 71.000 species, Ferns and Ales with 15.000 species, Mosses with 6.500 species [2]. Considering the highest number possessed by Flowering Plants, identification of the plants, which include also ornamental plant, has become a challenge for us.

As a recognition step for unidentified species, this research emphasis on ornamental plant, which has medicinal function. However, only few people know about its function as a treatment of the disease. As in Indonesia, this ornamental plant is mostly cultivated in front of the house. If this easiness and its medicinal function are taken into consideration, this plant should be an initial treatment or option towards full chemical-based medicines. In 2005, Gu et al. proposed leaf recognition based on the combination of Haar discrete wavelet transformation and Gaussian interpolation [3]. They used the wavelet to make a leaf skeleton and the extract the feature using run-length feature. Followed by Casanova et al. in 2009, conducted a research with title plant leaf identification using Gabour wavelets [4]. They were used 20 Brazilian plant species, then identified using Gabour wavelets and compared it with Fourier descriptor and co-occurrences matrices.

This research is a sub-research of the main research. The main research is identification ornamental plant based on its color, shape, and texture information. In this study, we want to investigate the best of the particular feature extraction components, which had previously been conducted [5]. For this measurement process, we utilize several wavelet transformations. Start from Daubechies wavelet transformation, and Dyadic wavelet transformation which are the previous researches [6], and then Dual-tree wavelet transformation which is current main focus. Because of there are small numbers of papers dealing with utilization of wavelet transformation for leaf identification. We propose to do comparative study utilizing aforementioned wavelet transformations.

This paper is organized as follows. Section 2 will describe ornamental plant that used in this research. Follow by the next section that explains about the proposed method. Experiments results in section 3. In the end of this paper, there are conclusion and future work.

II. Ornamental Plants

Ornamental plant in this research is not general the ornamental plant, because besides has function as an ornamental plant, also can be used as herbal medicine to cure many diseases. The main focus of this research is to recognize this kind of plant from its leaf.

Image data set of ornamental leaf in this research was obtained from direct acquisition using digital camera. This data set is taken based on tropical ornamental plant that usually cultivated in front of the house in Indonesia. This data set contains 8 classes with 15 images for each class. The classes are Bay (*syzygium polyanthum*), Cananga (*canagium odoratum, lamk*), Mangkokan (*nothopanax scutellarium merr.*), Jasmine (*jasminum sambac [soland]*), Cocor bebek (*kalanchoe pinnuta*), Vinca (*catharanthus roseus*), Kestuba (*euphorbia pulcherrima, willd*), Gardenia (*gardenia augusta, merr*). In order to avoid expensiveness of computation, size of the image is 256 x 256 pixels. The sample images of each class are presented as follow:

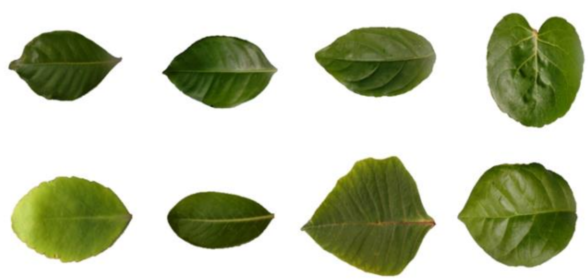

Fig. 2. Sample Images from every class inside data set

The Table 1 shows specific medicinal functions of the corresponding ornamental leaf. In Indonesia, there are two mostly used serving ways of these ornamental leaves. Firstly, boils the leaf together with water and apply as drinks. Secondly, put the leaf to the skin surface [7].

TABLE I. MEDICINAL FUNCTION OF ORNAMENTAL LEAF

Name	Medicinal Function
Bay	Diarrhea, scabies and itching
Cananga	Asthma
Mangkokan	Mastitis, skin injury, hair loss
Jasmine	Fever, head ache, sore eyes
Cocor Bebek	Ulcer, diarrhea, gastritis
Vinca	Diabetes, fever, burn
Kestuba	Bruise, irregular menstrual
Gardenia	Sprue, fever, constipation

III. Proposed Method

A. Wavelet Transformations

1) Daubechies Wavelet Transformation

For $N \epsilon \mathbb{N}$, Daubechies wavelet of class D-$2N$ is function $\psi = {}_N\psi \,\epsilon\, L^2(R)$ denoted by

$$\psi(x) \coloneqq \sqrt{2}\sum_{k=0}^{2N-1}(-1)^k h_{2N-1-k}\varphi(2x-k), \qquad (1)$$

where $h_0, \dots, h_{2N-1} \,\epsilon\, R$ are the constant filter coefficients that fulfilling the conditions

$$\sum_{k=0}^{N-1} h_{2k} = \frac{1}{\sqrt{2}} = \sum_{k=0}^{N-1} h_{2k+1}, \qquad (2)$$

similarly, for $l = 0, 1, \dots, N-1$,

$$\sum_{k=2l}^{2N-1+2l} h_k h_{k-2l} = \begin{cases} 1 & if\ l = 0, \\ 0 & if\ l \neq 0, \end{cases} \qquad (3)$$

and where $\varphi = {}_N\varphi : R \rightarrow R$ is the scaling function, given by the recursive equation

$$\varphi(x) = \sqrt{2}\sum_{k=0}^{2N-1} h_k \varphi(2x-k) \qquad (4)$$

Daubechies orthogonal wavelets of classes D2 - D20 (only even index numbers) are the wavelets that generally used [8]. The index number belongs to the number *2N* of coefficient. Single wavelet has a number of vanishing moments equal to half the number of coefficients. In this study we propose to use Daubechies D4 wavelets (DB D4), it has two vanishing moments. With these vanishing moments D4 can encodes polynomial of two coefficients, for example constant and linear signal components.

2) Dyadic Wavelet Transformation

The downsampling wavelet, which samples the scale and translation parameters, is often fails when deal with some assignments such as edge detection, features extraction, and image denoising [9,10]. Different with the downsampling wavelet, the dyadic wavelet samples only the scale parameter of a continuous wavelet transform, and does not samples the translation factor. In one side, it creates highly redundant signal representation, but the other side, since it has shift-invariance ability, this method is a convincing candidate as a feature descriptor method.

Let $L^2(R)$ be the space of square integrable functions on real line R, and define the Fourier transform of the function $\psi \in L^2(R)$ by

$$\hat{\psi}(\omega) = \int_{-\infty}^{\infty} \psi(t) e^{-i\omega t} dt \qquad (5)$$

If there is exist $A > 0$ and B such that

$$A \leq \sum_{-\infty}^{\infty} \left|\hat{\psi}(2^j \omega)\right|^2 \leq B, \tag{6}$$

then $\psi(t)$ is called dyadic wavelet function. Dyadic wavelet transform of $f(t)$ is defined using this $\psi(t)$ by

$$Wf(u, 2^j) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{2^j}} \psi(\frac{t-u}{2^j}), \tag{7}$$

from (6), $\hat{\psi}(0) = 0$ must be satisfied, i.e., $\int_{-\infty}^{\infty} \psi(t)dt = 0$. In order to design the dyadic wavelet function, we need a scaling function $\phi(t)$ satisfying a two-scale relation

$$\phi(t) = \sum_k h[k]\sqrt{2}\phi(2t-k). \tag{8}$$

The scaling function $\phi(t)$ is usually normalized as $\int_{-\infty}^{\infty} \phi(t)dt = 1$.

By (8), the Fourier transform of the scaling function resulting

$$\hat{\phi}(\omega) = \frac{1}{\sqrt{2}} \hat{h}\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right), \tag{9}$$

where $\hat{h}$ denotes a discrete Fourier transform

$$\hat{h}(\omega) = \sum_k h[k] e^{-i\omega t} \tag{10}$$

Since $\hat{\phi}(0) = 1$, we can apply (9) and (10) to obtain $\hat{h}(0) = \sqrt{2}$. Using the scaling function and the wavelet filter $g[k]$, we define a dyadic function by $\psi(t) = \sum_k g[k]\sqrt{2}\,\phi(2t-k)$.

The Fourier transform of $\psi(t)$,

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}} \hat{g}\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right), \tag{11}$$

will be needed later.

Let us denote the discrete Fourier transform of the filters $h[k]$, $h[k]$, $\tilde{h}[k]$, and $\tilde{g}[k]$, by $h(\omega)$, $h(\omega)$, $\hat{\tilde{h}}(\omega)$ and $\hat{\tilde{g}}(\omega)$ respectively.

We suppose that these Fourier transforms satisfy the below condition

$$\hat{\tilde{h}}(\omega)\hat{h}^*(\omega) + \hat{\tilde{g}}(\omega)\hat{g}(\omega) = 2, \quad \omega \in [-\pi, \pi], \tag{12}$$

where * denotes complex conjugation. This condition called a reconstruction condition.

Under condition (12), we have

$$a_{j+1}[n] = \sum_k h[k] a_j[n + 2^j k];\ j = 0,1, \ldots, \tag{13}$$

$$d_{j+1}[n] = \sum_k g[k] a_j[n + 2^j k];\ j = 0,1, \ldots, \tag{14}$$

here $a_0[n]$ is given by $a_0[n] = \int_{-\infty}^{\infty} f(t)\phi(t-n)dt$.

The (13) and (14) are dyadic decomposition formula for one-dimensional signals.

3) Dual-tree Complex Wavelet Transformation

Problems such as lack of shift invariances and poor directional selectivity that usually appears in DWT can be solved effectively using Dual-tree Complex Wavelet Transformation (DTCWT). Besides that ability, DTCWT also has limited redundancy and efficient order-N computation [12,13].

By doubling the sampling rate at each level of the tree, we can obtain approximate shift invariance with real DWT. This doubling process is done by eliminating the downsampling by 2 after level 1, and this is equal with two parallel fully-decimated trees. The filters in one tree must supply half a sample different of delays from the other tree. Odd-length in one tree and even-length in the other are required for linear phase. The image below is the dual-tree filters for the CWT.

From Fig. 3, tree A is real DWT that gives real part of wavelet transform, and tree B is real DWT that gives imaginary part. These real DWTs use different sets of filters. *h0*, *h1* and *g0*, *g1* denote low-pass/high-pass filter pair for upper filter bank and low-pass/high-pass filter bank for lower filter bank, respectively.

Different with the real DWT which only have three sub images in total. DTCWT decomposes three sub images for each spectral quadrant 1 and 2 and will have six bandpass sub images of complex coefficient at each level. Because of the complex wavelet filters have the ability to separate positive from negative frequency vertically and horizontally, the orientation for these six sub images will cover $\pm 75^0$, $\pm 45^0$, $\pm 15^0$.

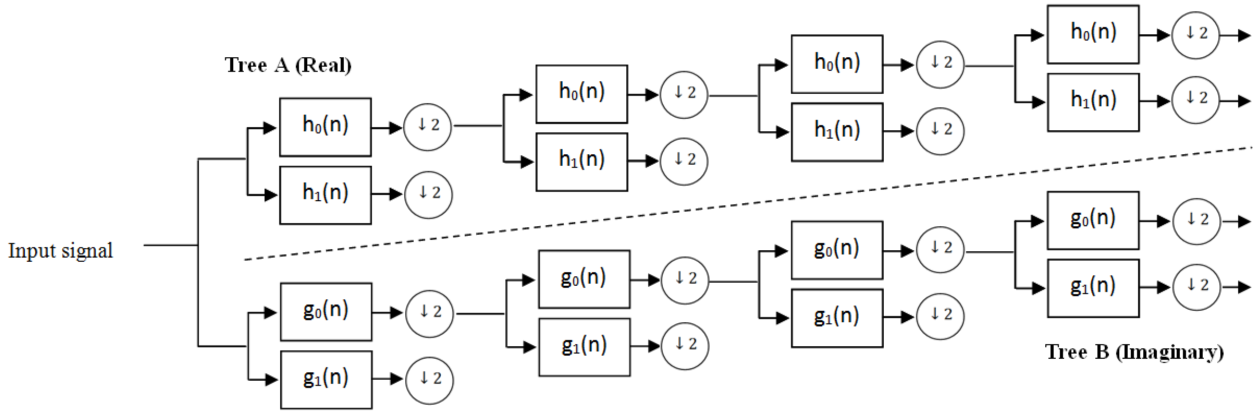

Fig. 3. Dual-tree filters for complex wavelet transformation

The following equation describes the detail:

$$\psi_i(a,b) = \frac{1}{\sqrt{2}}\left(\psi_{1,i}(a,b) - \psi_{2,i}(a,b)\right), \quad (15)$$

$$\psi_{i+3}(a,b) = \frac{1}{\sqrt{2}}\psi_{1,i}(a,b) + \psi_{2,i}(a,b)) \quad (16)$$

for i = 1,2,3, a denotes row, b denotes column, definition of two separable 2-D wavelet bases are below:

$$\begin{aligned}\psi_{1,1}(a,b) &= \phi_h(a)\psi_h(b),\\ \psi_{1,2}(a,b) &= \psi_h(a)\phi_h(b),\\ \psi_{1,3}(a,b) &= \psi_h(a)\psi_h(b),\\ \psi_{2,1}(a,b) &= \phi_g(a)\psi_g(b),\\ \psi_{2,2}(a,b) &= \psi_g(a)\phi_g(b),\\ \psi_{2,3}(a,b) &= \psi_g(a)\psi_g(b)\end{aligned} \quad (17)$$

The purpose of utilization $1/\sqrt{2}$ is to constitute an orthonormal operation from the sum/difference operation. Through the number of advantages from this DTCWT, we are expecting tight competition of performance with the Dyadic wavelet transformation.

B. SVM Classifier

SVM is a powerful tool for data classification. The indicators are the easiness to apply and impose Structural Risk Minimization (SRM). SRM armed the SVM to have strong ability in generalization of data. Its function is to minimize an upper bound on the expected risk. In principle, SVM learns to obtain optimal boundary with maximum margin that able to separate set of objects with different class of membership.

In order to achieve the maximum margin classifier, we have two options. Hard margin and soft margin are the options that totally depend on linearity of the data. Hard margin SVM is applicable to a linearly separable data set. However, often the data is not linearly separable. Soft margin SVM emerged as its solution [14,15]. The optimization problem for the soft margin SVM presented as below:

$$\min_{w,b} \frac{1}{2} \parallel w \parallel^2 + C\sum_{i=1}^{n} \xi_i$$

$$subject\ to\colon y_i(w^T x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0. \quad (18)$$

where w, C, ξ, b are the weight vectors, the penalty of misclassification or margin errors, the margin error, the bias, respectively.

In (18) can lead us to an efficient kernel methods approach. A kernel method is an algorithm that depends on the data only through kernel function, which computes a dot product in some possibly high dimensional data. Using the function ϕ training vector, the input space x is mapped into higher dimensional space. $K(x_i, x_j) = \phi(x_i)^T\phi(x_j)$ is called kernel function. The degree of the polynomial kernel can control the flexibility of resulting classifier [15]. It will be appropriate with this research when we classify 8 classes of leaf. Polynomial kernel is shown in equation (19).

$$K(x_i, x_j) = (\gamma {x_i}^T x_j + r)^d, \gamma > 0. \quad (19)$$

Where γ, r, d are kernel parameters, and i, j denote i^{th}, j^{th} vector in data set.

In this research, we propose to use Sequential Minimal Optimization (SMO). SMO act as efficient solver of the optimization problem in the training of support vector machines. SMO also solves the problems analytically by way of breaks the problems into a series of smallest possible problems.

C. Feature Extraction Components

In this research, we propose to use particular feature extraction components that extracted from the above wavelet transformations. The components are energy, mean, standard deviation and coefficient of variation. These components' values are extracted from four sub images, which are approximation, vertical, horizontal and diagonal sub images for DB D4 and Dyadic wavelet transformations. For DTCWT, those values are obtained from six sub images in real parts tree and six sub images in imaginary parts tree, in total is 12 sub images.

Because wavelet works in the frequency domain, energy is very useful as a feature extraction component. Energy values from several directions and different decomposition levels possibly well-capture the leaf main information which is leaf venation. The other component is mean value, and it can be interpreted as central of tendency of the ornamental leaf data. When standard deviation or predictable dispersion increases in proportion to concentration, we have the other value as a solution called coefficient of variation. Through these comprehensive and fully-related values, we believe that we can gain satisfying results.

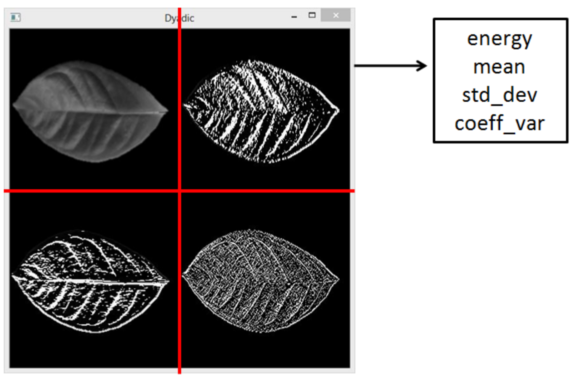

Fig. 4. Feature extraction components for all wavelet transformation

IV. EXPERIMENTS

A. Test Aspects

Test aspects of this research consist of original (no change at all), rotation, scaling, translation, and perspective change. It means we have 5 different datasets with its corresponding test aspect. Inside the rotated dataset, there are images with dissimilar degrees of rotation start from 45^0, 90^0, 135^0, 180^0, 225^0, 270^0, and 315^0. For scaled dataset, we have 30% and 60% downscaled images. Inside the translated dataset, we translated the images to many directions within the original image size. The last is perspective changed dataset with two different perspective changed, are applied to the images.

(a) (b)

(c) (d)

(e)

Fig. 5. Test aspects of ornamental leaf images. (a) original, (b) rotation, (c) scaling, (d) translation, (e) perspective change.

B. Comparison Through Classification

TABLE II. BEST DECOMPOSITION LEVEL FOR DB D4

Decomposition level	Level 1	Level 2	Level 3
Original Dataset	**95.83**	94.17	87.5

TABLE III. BEST DECOMPOSITION LEVEL FOR DYADIC

Decomposition level	Level 1	Level 2	Level 3
Original Dataset	**98.33**	97.5	96.67

TABLE IV. BEST DECOMPOSITION LEVEL FOR DUAL-TREE CWT

Decomposition level	Level 1	Level 2	Level 3
Original Dataset	95.83	96.67	**99.17**

According to the result presented in Table 2 and Table 3, it was obtained that decomposition level 1 is the preferable result comparing to the other levels. Then, the best decomposition level for DTCWT was level 3. The reason for those preferable results are from that level was these wavelet transformation could extract very well the leaf main information called leaf venation, in comparison with the other level.

TABLE V. COMPARISON USING VARIOUS DATASETS

Datasets	Original	Translation	Rotation	Scaling	Perspective	Average
DB D4	95.83	87.5	87.5	90.83	91.67	90.66
Dyadic	98.33	**97.5**	94.17	**96.67**	95	96.33
Dual-tree CWT	**99.17**	95.83	**96.67**	95.83	**95.83**	**96.66**

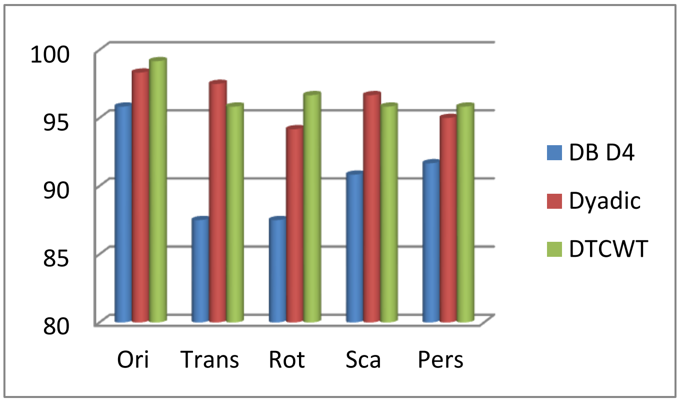

Fig. 6. Comparison of above mentioned feature extraction components using various wavelet transformations.

In Table 5 and Fig. 6 are presented the performance results of previously mentioned test aspects from wavelet transformations. For overall results, DB D4 gave us poor performance in all aspects. As expected, the tight competition came from the Dyadic wavelet transformation and DTCWT. For translation and scaling aspect, the dyadic wavelet transformation is superior to the others. Since sampling only scale parameter of continuous wavelet transformation, made the Dyadic has finer shift invariant property. Because of that reason also, the decomposition image in Dyadic will be exactly same with the input image, and feature extraction components in downscaled image that represent the leaf venation information still could capture.

However, for original, rotation and perspective, the DTCWT is superior. DTCWT has advantages in shift invariant and rotation invariant. Shift invariant property in the DTWCT is not perfect shift invariant but only nearly shift invariant. That reason made the Dyadic results more preferable in the translation dataset.

In the rotation aspect, DTCWT showed its advantages, through the directional selectivity support for six directions in real parts tree and six directions in imaginary parts tree, this DTCWT could give proper results. In original and perspective aspects, the only reason we gained superior results is the effectiveness of utilization 12 feature extraction components. Dissimilar with the DB D4 and Dyadic we only could utilize components from three directions of wavelet decomposition sub images plus one approximation image.

V. CONCLUSIONS

In this research, we have already measured the performance of Daubechies D4, Dyadic, and Dual-tree wavelet transformations. Daubechies D4 included in real DWT, which has several minus points, from the shift variance until lack of directional selectivity. Different with Daubechies, Dyadic and DTCWT has solutions to solve that shift variance problem. Because of the Dyadic sampling only scale parameter of continuous wavelet transformation, resulting highly redundant signals, but this is the important point when we try to extract information from the leaf for the translation and scaling aspects. Besides possess the nearly shift invariant property, DTCWT also support directional selectivity through six directions for real and imaginary parts, make them is superior for original, rotation and perspective aspect, as well as the overall result.

VI. FUTURE WORKS

Hence the results from DTCWT outperform in comparison with the others. We still have a deficiency in translation and scaling aspects. Therefore, to find the best in all aspects we plan to conduct the comparative study involving the Double Density wavelet transformation and Double Density Dual-tree Complex wavelet transformation. Double Density wavelet transform has property nearly equal with the DTWCT, and the Double Density Dual-tree is the combination between both.

REFERENCES

[1] IUCN 2012. Numbers of threatened species by major groups of organisms (1996–2012). <http://www.iucnredlist.org/documents/summarystatistics/2012_2_RL_Stats_Table_1.pdf >. Downloaded on 21 June 2013.

[2] Chapman A. D., "Numbers of living species in Australia and the World". Australian Government, Department of the Environment, Water, Heritage, and the Arts. Canberra, Australia, 2009.

[3] Casanova, D., de Mesquita Sá Junior, J. J. and Bruno, O. M., "Plant leaf identification using Gabor wavelets," Int. J. Imaging Syst. Technol., vol. 19, pp. 236–243, 2009.

[4] D.S. Huang, X.-P. Zhang, G.-B. Huang, "Leaf recognition based on the combination of wavelet Transform and gaussian Interpolation", ICIC 2005, Part I, LNCS 3644, pp. 253–262, 2005.

[5] Arai K., Abdullah I. N., Okumura H. and Kawakami R., "Improvement of automated detection method for clustered microcalcification based on wavelet transformation and support vector machine," International Journal of Advanced Research in Artificial Intelligence(IJARAI), 2(4), 2013.

[6] Arai K., Abdullah I. N., Okumura H., "Image identification based on shape and color descriptors and its application to ornamental leaf", IJIGSP, vol.5, no.10, pp.1-8, 2013.

[7] Hartati Sri, "Tanaman hias berkhasiat obat. IPB Press," 2011. (In Indonesian)

[8] De Vries Andreas, "Wavelets", FH Sudwestfalen University of Applied Sciences, Hagen, Germany, 2006.

[9] Abdukirim T., Nijima K., Takano S. "Lifting dyadic wavelets for denoising," Proceedings of the Third International Workshop on Spectral Methods and Multirate Signal Processing (SMMSP), pp. 147-154. 2003.

[10] Minamoto T., Tsuruta K., Fujii S. "Edge-preserving image denoising method based on dyadic lifting schemes," IPSJ Transactions on Computer Vision and Applications, vol. 2, 48-58, 2010.

[11] Mallat S., "A wavelet tour of signal processing," Academic Press. 1998.

[12] Kingsbury, N. G., "The dual tree complex wavelet transform: a new efficient tool for image Restoration and enhancement," In Proc. EUSIPCO98, Rhodes, Greece, 319-322, 1998.

[13] Selesnick I.W., Baraniuk R.G., and Kingsbury N.G., "The dual-tree complex wavelet transformation", IEEE Signal Processing Magazine, pp. 123-151, November 2015.

[14] Ben-Hur Asa, Weston Jason, "A user's guide to support vector machine. Data Mining Techniques for the Life Science", 223-239, Humana Press, 2010.

[15] Hsu Chih-Wei, Chang Chih-Chuang, Lin Chih-Jen, "A practical guide to support vector classification," Department of Computer Science National Taiwan University, Taiwan, 2010.

[16] Starck Jean-Luc, Murtagh Fionn, Fadili Jalal M. Sparse image and signal processing. Cambridge University Press. 2010.

Human Gait Gender Classification using 3D Discrete Wavelet Transform Feature Extraction

Kohei Arai[1]
[1)]Graduate School of Science and Engineering
Saga University
Saga City, Japan

Rosa Andrie Asmara[1,2]
[2)]Informatics Management Department
State Polytechnics of Malang
Malang, Indonesia

***Abstract*—Feature extraction for gait recognition has been created widely. The ancestor for this task is divided into two parts, model based and free-model based. Model-based approaches obtain a set of static or dynamic skeleton parameters via modeling or tracking body components such as limbs, legs, arms and thighs. Model-free approaches focus on shapes of silhouettes or the entire movement of physical bodies. Model-free approaches are insensitive to the quality of silhouettes. Its advantage is a low computational costs comparing to model-based approaches. However, they are usually not robust to viewpoints and scale. Imaging technology also developed quickly this decades. Motion capture (mocap) device integrated with motion sensor has an expensive price and can only be owned by big animation studio. Fortunately now already existed Kinect camera equipped with depth sensor image in the market with very low price compare to any mocap device. Of course the accuracy not as good as the expensive one, but using some preprocessing method we can remove the jittery and noisy in the 3D skeleton points.**

Our proposed method is to analyze the effectiveness of 3D skeleton feature extraction using 3D Discrete Wavelet Transforms (3D DWT). We use Kinect Camera to get the depth data. We use Ipisoft mocap software to extract 3d skeleton model from Kinect video. From the experimental results shows 83.75% correctly classified instances using SVM.

Keywords—gender gait classification; 3D Skeleton Model; SVM; Biometrics; 3D DWT

I. Introduction

In recent years, there has been an increased attention on effectively identifying individuals for prevention of terrorist attacks. Many biometric technologies have emerged for identifying and verifying individuals by analyzing face, fingerprint, palm print, iris, gait or a combination of these traits [1]–[3].

Human Gait as the classification and recognition object is the famous biometrics system recently. Many researchers had focused this issue to consider for a new recognition system [4]–[11]. Human Gait classification and recognition giving some advantage compared to other recognition system. Gait classification system does not require observed subject's attention and assistance. It can also capture gait at a far distance without requiring physical information from subjects.

There is a significant difference between human gait and other biometrics classification. In human gait, we should use video data instead of using image data as other biometrics system used widely. In video data, we can utilize spatial data as well as temporal data compare to image data.

There are 2 feature extraction method to be used in gait -classification: model based and free model approach [12]. Model-based approaches obtain a set of static or dynamic skeleton parameters via modeling or tracking body components such as limbs, legs, arms and thighs. Gait signatures derived from these model parameters employed for identification and recognition of an individual. It is obvious that model-based approaches are view-invariant and scale-independent. These advantages are significant for practical applications, because it is unlikely that reference sequences and test sequences taken from the same viewpoint. Model-free approaches focus on shapes of silhouettes or the entire movement of physical bodies. Model-free approaches are insensitive to the quality of silhouettes. Its advantage is a low computational costs comparing to model-based approaches. However, they are usually not robust to viewpoints and scale [13].

Gender classification along with human gait recognition has getting the researchers to find its best methods. Wide implementation make they seem so attractive research. The implementation will not only enhance existing biometrics systems but can also serve as a basis for passive surveillance and control in "smart area" (e.g., restricting access to certain areas based on gender) and collecting valuable demographics (e.g., the number of women entering a retail store, airports, post office, or public smoking area etc. on a given day)

Imaging technology developed quickly this decades. Motion capture (mocap) device integrated with motion sensor has an expensive price and can only be owned by big animation studio. Fortunately now already existed Kinect camera equipped with depth sensor image in the market with very low price compare to any mocap device. Of course the accuracy not as good as the expensive one, but using some preprocessing we can remove the jittery and noisy in the 3D skeleton points. Our proposed method is part of model based feature extraction and we call it 3D Skeleton model. 3D skeleton model for extracting gait itself is a new model style considering all the previous model is using 2D skeleton model. The advantages itself is getting accurate coordinate of 3D point for each skeleton model rather than only 2D point. We use Kinect to get the depth data. We use Ipisoft mocap software to extract 3d skeleton model from Kinect video. Those 3D skeleton model exported to BVH animation standard format file and imported to our programming tool which is Matlab. We use Matlab to extract the feature and use a classifier. We create our own gender gait dataset in 3D environment since there are not exist such a dataset before.

II. Proposed Method

The classification of gender gait quality in this paper consists of three part, preprocessing, feature extraction, and classification. Figure 1 shows the complete overview of proposed human disable gait quality classification.

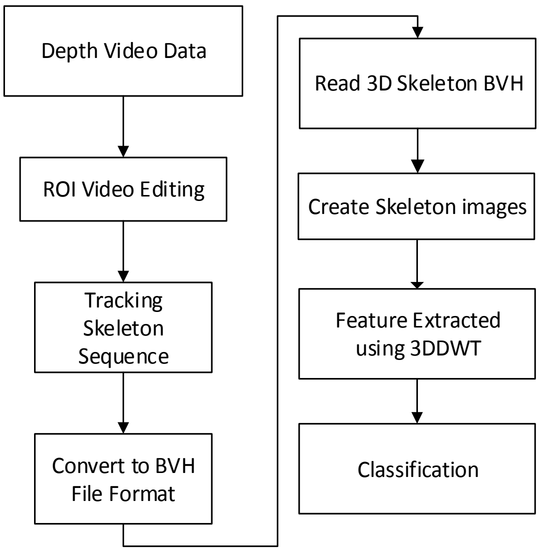

Figure 1: Proposed human gait gender classification

Using Kinect camera have one advantage compare to usual RGB camera. The skeleton created is in 3D space. One can get 2D images from different view angle using only single camera. Figure 2 below shows the 2D skeleton image created from different view angle at the same frame. This is useful to enhance the accuracy of the classification since some paper proposed using multi view image [10], [14]–[16]. However, these papers will only using one view for the analysis.

A. Preprocessing

First, take the Video data using Kinect and IpiRecorder to record the depth data along with RGB video data. To get the video data, there are some recommendation should be considered:

1. Using 9 by 5 feet room space to get best capture.
2. Object should be dressed in casual slim clothing, avoid shiny fabrics.
3. We should ensure that the whole body including arms and legs is visible during the recording states. Beginning from T-Pose and the recording can be started.

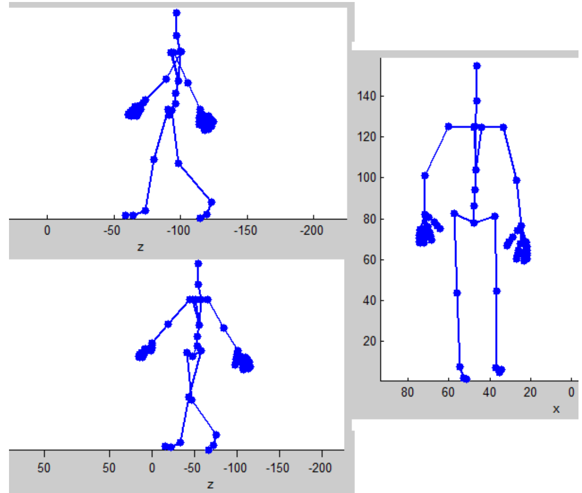

Figure 2. 2D skeleton image created from different view angle at the same frame

Second, processed the depth video data in IPISoft motion capture application. IPISoft will create the 3D skeleton model from video depth recorded using some tracking motion method. The first step is to take only the gait scene, and remove unimportant video scene or we call the Region of Interest (ROI) video. Figure below show the example of video recording.

Third, Create the skeleton 3d model using the tracking motion method, remove the jittery and noises, and export the skeleton model to BVH file format in IPISoft.

Fourth, Read the BVH file, extracted the feature, and classify the feature.

B. Dataset

Unfortunately, there are no Kinect Video Depth gait dataset exists until now. All exist gait dataset is using ordinary RGB camera like USF gait dataset, SOTON gait dataset, and CASIA gait dataset. Figure 3 shows the example of CASIA gait dataset.

Figure 3: Example of CASIA gait dataset

To conduct the experiment, we should prepare the dataset. We will use the Kinect Gait Dataset to analyze and classify gender using gait. The proposed research will search the capability of Kinect and 3D Skeleton model and use their 2D images for gait classification.

Figure 4 below shows the T-pose position before the video recording start. The top right image showing the RGB video sequence. T

he color gradient used to represents the depth in video data. Blue color means the object is close to the camera and red color means the object is far from camera.

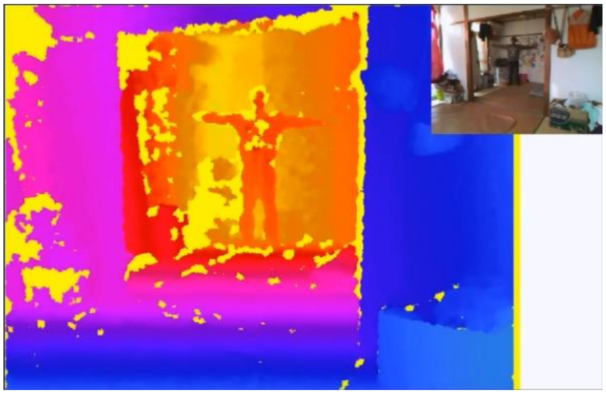

Figure 4: T-Pose Position before the recording begin

Figure 5 below show the 3D skeleton tracking motion sequence. First task is specifying subject's physical parameter like gender and height. IpiSoft will detect the ground plane automatically and provide the 3D skeleton in T-Pose position. Our next job is try to put the T-Pose skeleton in the same position with the subject T-Pose position in the first sequence of video.

This time also we should determine the Region of Interest video to be processed. Instead of all the video sequence that we use, we could only take the most important part of the video sequence. Once we put the skeleton to the same position with the subject, we can refitting pose using the application and start tracking. Jittery removal and Trajectory filtering can be done after the tracking finished.

The skeleton sequence result can be import to BVH file standard. Figure 6 and 7 below shows the BVH file result and preview in BVH file viewer and Matlab.

Figure 5: Skeleton motion tracking sequence

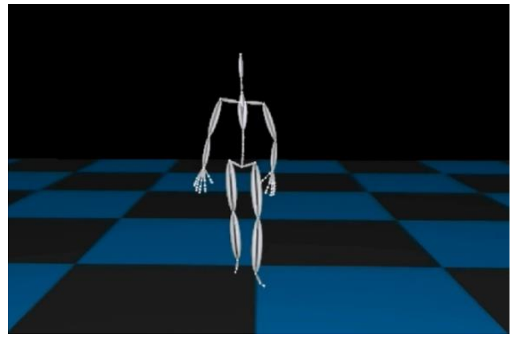

Figure 6: BVH skeleton results

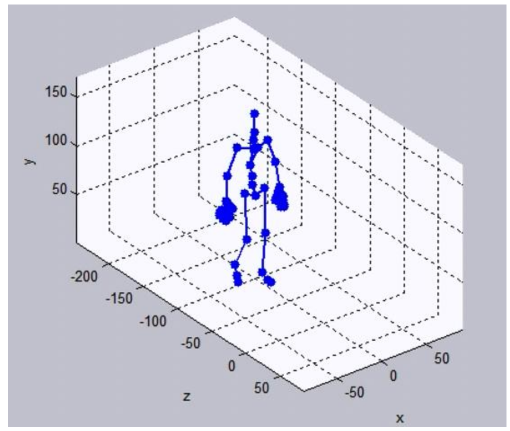

Figure 7: Skeleton model imported in MATLAB

C. *3D Discrete Wavelet Transforms (3D DWT)*

3D version of Discrete Wavelet Transform is specially used in volume and video processing. In the 3D case, the 1D analysis filter bank is applied in turn to each of the three dimensions. If the data is of size N1 by N2 by N3, after applying the 1D analysis filter bank to the first dimension one have two sub band data sets, each of size N1/2 by N2 by N3. After applying the 1D analysis filter bank to the second dimension one have four sub band data sets, each of size N1/2 by N2/2 by N3. Applying the 1D analysis filter bank to the third dimension gives eight sub band data sets, each of size N1/2 by N2/2 by N3/2. This is illustrated in the figure 8 below.

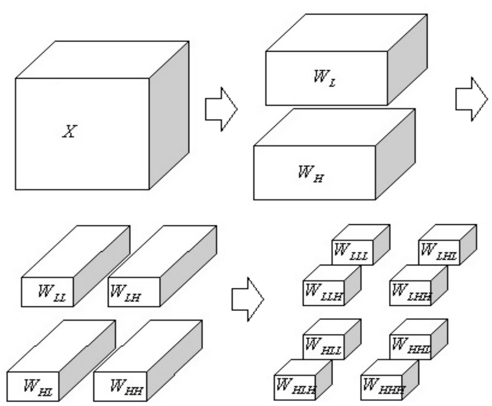

Figure 8. The resolution of a 3-D signal is reduced in each dimension

The block diagram of the 3D analysis filter bank is shown in Figure 9.

D. Feature Extraction

To extract the feature using 3D Discrete Wavelet Transform, we can prepare 2 kind of data. The first data is raw data and the second data is the resized data. The effectiveness and Classification accuracy of each data using statistical feature will be shown in this paper. The resized data have an advantage over the raw data. In Resize data, we can use whole sub bands or decomposition data and process them in the classifier directly.

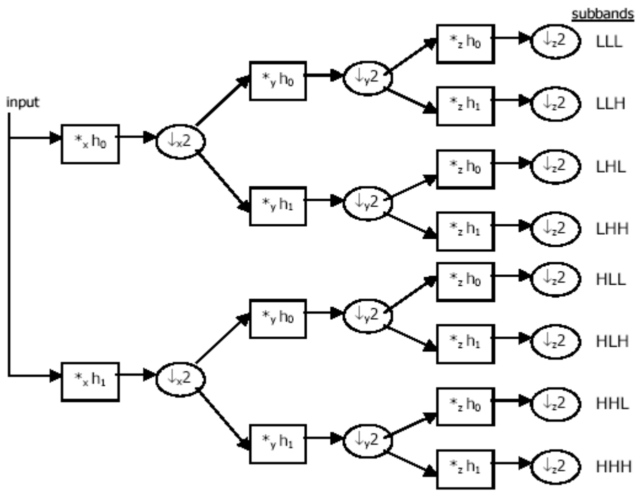

Figure 9. Block diagram of a single level decomposition for the 3D DWT

This can be done because all the dataset have the same dimension. This paper will cover the experimental result of resized data and extracted statistical feature, but this paper will not discuss analyzing of using whole the sub bands feature extraction. To extract the feature in the dataset created, one have to consider about the image size. If all the image is used, it will be costly. One have to extract the skeleton image only and not all the image, thus call it Region of Interest (ROI). One can do automatic ROI using simple image detection since the image is in binary space. After the ROI done, one can extract feature from the data directly or one can resize the data. Thus, there are two kind of data which is raw data and resized data.

Data have to same in image file and frame amount to be used as a resized data. The method used in this paper to create resized data is image resizing and frame cropping. In image resizing, biggest skeleton image ROI will be used as a reference because in this method we don't have to remove some amount of data and those removed data could be valuable information to the system. After finish the image resizing, one can start to crop the frame. This paper will crop the frame based on the smallest amount of video frame, thus all the data have same amount of frame. This paper using middle part of smallest frame amount as a cropped frames. If the smallest frame amount is x, then the video frame crop start at y is round ((total_frame – x)/2) and end in z = y+x.

This paper will used some famous classifier to compare and analyze their best correct classification rate. The decomposition result of Level 1 from 3D DWT of Haar Wavelet will get 8 sub bands which is LLL, LLH, LHL, LHH, HLL, HLH, HHL, and HHH. We will use 3 statistic feature that was used in previous research which is mean, standard deviation, and energy.

The formula for the energy used in Eq.(1) below.

$$E_3 = \sqrt{\sum_{k=1}^{tot_{frame}} \left(\sqrt{\sum_{j=1}^{tot_{col}} \left(\sqrt{\sum_{i=1}^{tot_{row}} (|Wave\ Coef|)^2} \right)^2 } \right)^2 } \quad (1)$$

From above explanation, we can extract 24 attribute from first and second data type. First 8 attributes is a mean for each sub bands. Second 8 attributes is a standard deviation for each sub bands, and third 8 attributes is an energy for each sub bands. All those 24 attributes namely as mean1_LLL, mean2_LLH, mean3_LHL, mean4_LHH, mean5_HLL, mean6_HLH, mean7_HHL, mean8_HHH, std1_LLL, std2_LLH, std3_LHL, std4_LHH, std5_HLL, std6_HLH, std7_HHL, std8_HHH, e3_LLL, e3_LLH, e3_LHL, e3_LHH, e3_HLL, e3_HLH, e3_HHL, and e3_HHH.

E. Classification

This paper will use two famous classifier which is Naïve Bayes and SVM to analyze and compared the results. SVM (Support Vector Machine) are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. The basic SVM takes a set of input data and predicts, for each given input, which of two possible classes forms the output, making it a non-probabilistic binary linear classifier.

Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

III. Experimental Result

We start with using Resized data. From 80 dataset with 40 male and 40 female video, we will use SVM and Naïve Bayes as the classifier. First step is we are selecting the best feature in each classifier. There are two method to be use which is Wrapper Method and Ranked Method. We also conduct those two method in two kind of data preprocessing, which is without preprocessing and data after discretization filter preprocessing.

Table 1 is the result of the selected feature using both methods without preprocessing data (raw data type).

TABLE 1. FEATURE SELECTION WITHOUT PREPROCESSING DATA FOR RAW DATA TYPE

	Wrapper Method	Ranked Method
SVM	e3_LLL, e3_LLH	Std1_LLL, e3_HLH, e3_LLL, e3_HLL, e3_HHL, e3_HHH, e3_LHH
Naïve Bayes	mean2_LLH, mean3_LHL, mean6_HLH, mean7_HHL, mean8_HHH, std1_LLL, std3_LHL, std8_HHH, e3_LLL	

Table 2 is the result of the selected feature using both methods with discretization data (raw data type). Table 3 is the result of the selected feature using both methods without preprocessing data (resized data type). Table 4 is the result of the selected feature using both methods with discretization data (resized data type).

TABLE 2. FEATURE SELECTION WITH DISCRETIZATION FILTER DATA FOR RAW DATA TYPE

	Wrapper Method	Ranked Method
SVM	mean5_HLL, mean6_HLH, mean8_HHH, std1_LLL, e3_HLH	e3_HLH, std1_LLL, e3_HHH, e3_LLL, e3_HLL, e3_HHL, std2_LLH
Naïve Bayes	mean2_LLH, mean3_LHL, mean6_HLH, mean8_HHH, std1_LLL, e3_LLL, e3_HLH	

TABLE 3. FEATURE SELECTION WITHOUT PREPROCESSING DATA FOR RESIZED DATA TYPE

	Wrapper Method	Ranked Method
SVM	e3_LLL, e3_LLH, e3_HLL	std6_HLH, std8_HHH
Naïve Bayes	mean6_HLH, std8_HHH, e3_LHL	

TABLE 4. FEATURE SELECTION WITH DISCRETIZATION FILTER DATA FOR RESIZED DATA TYPE

	Wrapper Method	Ranked Method
SVM	mean5_HLL, mean8_HHH, std1_LLL, std7_HHL, std8_HHH, e3_HLL	std6_HLH, std8_HHH, std1_LLL, std4_LHH, e3_HLL, mean5_HLL, e3_LLL, std5_HLL
Naïve Bayes	mean3_LHL, mean4_LHH, mean5_HLL, mean8_HHH, std4_LHH, std8_HHH, e3_LHL, e3_LHH, e3_HHL	

Table 5 shows the result of Correct Classification Rate (CCR) for the selected features (raw data type). Table 6 shows the result of Correct Classification Rate (CCR) for the selected features (resized data type).

TABLE 5. CORRECT CLASSIFICATION RATE FOR EACH SELECTED FEATURE IN EACH CLASSIFIER FOR RAW DATA TYPE

	Wrapper Method		Ranked Method	
	No preprocess	Discretize	No preprocess	Discretize
SVM	68.75%	80%	48.75%	83.75%
Naïve Bayes	77.5%	68.75%	72.5%	76.25%

TABLE 6. CORRECT CLASSIFICATION RATE FOR EACH SELECTED FEATURE IN EACH CLASSIFIER FOR RAW AND RESIZED DATA TYPES

	Wrapper Method		Ranked Method	
	No preprocess	Discretize	No preprocess	Discretize
SVM	45%	76.25%	58.75%	75%
Naïve Bayes	56.25%	61.25%	67.5%	70%

As seen in the tables above, the best CCR is in raw data type using SVM and discretized data and ranked method selected feature. Table 7 is detail accuracy by class using SVM classifier.

TABLE 7. DETAIL ACCURACY BY CLASS USING SVM CLASSIFIER

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
Male	0.775	0.1	0.886	0.775	0.827	0.838
Female	0.9	0.225	0.8	0.9	0.847	0.838
Weighted Average	0.838	0.163	0.843	0.838	0.837	0.838

IV. CONCLUSION

The proposed method uses Kinect depth sensor camera and Ipisoft motion capture software to generate 3D skeleton model. Ipisoft itself is special purpose application to create skeleton so user can use the motion to their computer generated character motion.

The 3D skeleton generated will then extract the 2D image in one view angle and create 2 model data type which is raw and resized video data type. Using Level 1 Haar 3D DWT, we got 8 sub bands and using 3 statistical feature for all 8 sub bands (Mean, Standard deviation, and Energy). By selecting the best feature and classify the results using SVM and Naïve Bayes, the result shows is Table 5 and Table 6. The best result achieved in raw data type using Ranked method feature selection and discretized data which is 83.75% CCR.

ACKNOWLEDGMENT

Portions of the research in this paper is funded by Saga University Japan. Authors would like to thank to Saga University for their providing of the Laboratory and Infrastructures.

REFERENCES

[1] X. Qinhan, “Technology review - Biometrics-Technology, Application, Challenge, and Computational Intelligence Solutions,” *IEEE Comput. Intell. Mag.*, vol. 2, pp. 5–25, 2007.

[2] E. Yih, G. Sainarayanan, and A. Chekima, “Palmprint Based Biometric System: A Comparative Study on Discrete Cosine Transform Energy, Wavelet Transform Energy and SobelCode Methods,” *Biomed. Soft Comput. Hum. Sci.*, vol. 14, no. 1, pp. 11–19, 2009.

[3] Z. Yang, M. Li, and H. Ai, “An experimental study on automatic face gender classification,” *Pattern Recognition, 2006. ICPR 2006. 18th Int. Conf.*, vol. 3, pp. 1099 – 1102, 2006.

[4] N. V. Boulgouris, D. Hatzinakos, and K. N. Plataniotis, “Gait recognition: A Challenging Signal Processing Technology For Biometric Identification,” *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 78–90, 2005.

[5] M. S. Nixon and J. N. Carter, “Automatic Recognition by Gait,” *Proc. IEEE*, vol. 94, pp. 2013–2024, 2006.

[6] D. Cunado, M. S. Nixon, and J. N. Carter, “Automatic Extraction and Description of Human Gait Models for Recognition Purposes,” *Comput. Vis. Image Underst.*, vol. 90, no. 1, pp. 1–41, 2003.

[7] X. Li, S. Maybank, and S. Yan, “Gait components and their application to gender recognition,” *Syst. Man, Cybern. Part C Appl. Rev. IEEE Trans.*, vol. 38, no. 2, pp. 145–155, 2008.

[8] K. Arai and R. Asmara, “Human Gait Gender Classification using 2D Discrete Wavelet Transforms Energy,” *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, pp. 62–68, 2011.

[9] K. Arai and R. Asmara, “Human Gait Gender Classification in Spatial and Temporal Reasoning,” *IJARAI Int. J. Adv. Res. Artif. Intell.*, vol. 1, no. 6, pp. 1–6, 2012.

[10] G. Huang and Y. Wang, “Gender classification based on fusion of multi-view gait sequences,” *Comput. Vision–ACCV 2007*, vol. 4843, pp. 462–471, 2007.

[11] L. Lee and W. Grimson, “Gait Analysis for Recognition and Classification,” in *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, 2002, pp. 148–155.

[12] J. Wang, M. She, S. Nahavandi, and A. Kouzani, “A Review of Vision-Based Gait Recognition Methods for Human Identification,” *2010 Int. Conf. Digit. Image Comput. Tech. Appl.*, pp. 320–327, Dec. 2010.

[13] K. Arai and R. Asmara, “3D Skeleton model derived from Kinect Depth Sensor Camera and its application to walking style quality evaluations,” *Int. J. Adv. Res. Artif. Intell.*, vol. 2, no. 7, pp. 24–28, 2013.

[14] H.-C. Lian and B.-L. Lu, “Multi-view gender classification using multi-resolution local binary patterns and support vector machines.,” *Int. J. Neural Syst.*, vol. 17, no. 6, pp. 479–87, Dec. 2007.

[15] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li, “Multi-view Gait Recognition Based on Motion Regression Using Multilayer Perceptron,” *2010 20th Int. Conf. Pattern Recognit.*, pp. 2186–2189, Aug. 2010.

[16] Y. Wang, S. Yu, and T. Tan, “Gait recognition based on fusion of multi-view gait sequences,” *Adv. Biometrics*, vol. 3832, pp. 605–611, 2005.

AUTHORS PROFILE

Kohei Arai 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 29 books and published 500 journal papers.

Rosa A. Asmara received the B.E. degree in electronics engineering from Brawijaya University, and the M.S. degree in Multimedia engineering, from Institute of Technology Sepuluh Nopember, Surabaya, Indonesia, in 2004 and 2009, respectively. He is currently a researcher at Information Science in Saga University, Japan. His research interests include signal processing, image processing, parallel processing, pattern recognition, and computer vision.

Method for Traffic Flow Estimation using On-dashboard Camera Image

Kohei Arai
Graduate School of Science and Engineering
Saga University
Saga, Japan

Steven Ray Sentinuwo
Department of Electrical Engineering
Sam Ratulangi University
Manado, Indonesia

***Abstract*—This paper presents the method to estimate the traffic flow on the urban roadway by using car's on-dashboard camera image. The system described, shows something new which utilizes only road traffic photo images to get the information about urban roadway traffic flow automatically.**

Keywords—traffic flow estimation; on-dashboard camera; computer vision.

I. INTRODUCTION

Due to the increasing of the traffic densities nowadays, there exists a growing demand for advanced systems that can provide the essential traffic and travel information for the drivers to improve the traffic quality and travel optimization. The comprehensive and accurate of traffic information is the important aspect in order to manage the roadway network and to provide the navigation service for the road users. The density value measurement is one of the important aspect of transportation management. Intelligent Transportation System is a breakthrough technology that combines components of information management systems to create better transportation system. Currently, demands on the system are growing rapidly with an estimated travel demand increase of 30% over the next ten years[1]. The efficient of transportation system can be an alternative to increase the roadway capacity in order to prevent the traffic congestion at current levels from getting worse.

Vehicle detection and surveillance technologies are an integral part of ITS, since they gather all or part of the data that is used in ITS. There is an estimation that an investment in ITS will allow for fewer miles of road to be built, thus reducing the cost of mitigating recurring congestion by approximately 35 percent nationwide. New vehicle detection and surveillance technologies are constantly being developed and existing technologies improved, to provide speed monitoring, traffic counting, presence detection, headway measurement, vehicle classification, and weigh-in-motion data.

There are many methods and technologies for traffic measurement. However, the conventional measurement methods, such as "in-situ" technologies which are measure traffic data by locating detectors or sensors along the roadway, have some drawbacks. The use of these conventional methods for collecting data is necessary but not sufficient because of their limited coverage area and expensive costs of implementation and maintenance. Then as the alternative to the conventional methods there are alternative methods, such as based on the vehicle location or Floating Car Data method (FCD), and collecting data from "in-vehicle" devices.

This paper proposes a method for traffic measurement using the approach of collecting data from "in-vehicle" device which is by utilizing camera device that located on the car`s dashboard. The image that captured by digital camera photo then was analyzed to get the estimation value of roadway traffic flow condition. A traffic measurement using digital camera device can be a practical and low cost solution compare to the conventional traffic measurement methods.

This paper is organized as follow. The brief description about traffic flow measurement methods and procedures is quick review in Section 2. Section 3 described some previous research that taking into account traffic monitoring models. In Section 4, the proposed method of traffic flow measurement using on-dashboard camera image is explained. Section 5 and Section 6 show the results of simulation study and experiment in the real traffic condition, respectively. Finally, in Section 7, we present the conclusion and the future works.

II. TRAFFIC FLOW MEASUREMENT

Researchers in the area of Intelligent Transportation System have been greatly interested in the various of traffic management and monitoring applications. For several years, under growing pressure for improving traffic management, a wide variety of applications have been developed. Some researchers focus on the theoretical aspect that involves driver behavior into the transportation model [2][3][4][5]. While the other focus on the development of application and tools of intelligent transportation system.

There are some items of interest in traffic theory, have been the following[6]:

- Rates of flow, which evaluates the number of vehicles per unit of time;
- Speeds, which evaluates the distance per unit time;
- Travel time over a known length of road, or sometimes refer to the invers of speed;
- Occupancy, that is percent of time a point on the road is occupied by vehicles;
- Density, which is refer to the number of vehicles per unit distance.
- Time headway between vehicles (time per vehicle);

- Spacing, or space headway between vehicles (distance per vehicles);
- Concentration, that is measured by density of occupancy.

The capability of traffic measurement to gather transportation data have changed over the nearly sixty year span of interest in traffic flow measurement. And to be more growing which there have been a large number of freeway and number of population. Indeed, traffic measurement methods are still changing, there are five common procedures in this area:

- Measurement at a point;
- Measurement over a short section, by which meant less than about 10 meters (m);
- Measurement over a length of road, usually at least 0.5 kilometers (km);
- Using moving observer along the traffic stream;
- Wide-area samples obtained simultaneously from a number of vehicles, as part as the Intelligent Transportation System (ITS).

In moving observer method, there are two common approaches[6]. The first one is the simple floating car data procedure (FCD). This approach intends to record speeds and travel times as the function of time and location along the road. The intention of this approach is the floating car as the observer car behaves as an average vehicle within the traffic stream. However, this approach cannot give provide average speed data. This one just effective for producing a qualitative information about roadway conditions and operations. One form of this approach uses a person in the floating car to record speeds and travel times. The second form uses a modified recording speedometer of the type regularly used in long-distance trucks or buses. While the drawback of this approach is that it means there are usually significantly fewer speed observations than volume observations.

The second one introduced by Wardrop et.al.[7] for urban traffic measurements. This approach intends to obtain both speed and volume measurements simultaneously. This method is based on an observation vehicle that travels in both direction on the roadway. The formulae allow one to estimate both speeds and flows for one direction of travel. The formulae are as follow :

$$q = \frac{(x+y)}{(t_a+t_w)} \quad (1)$$

$$\bar{t} = t_w - \frac{y}{q} \quad (2)$$

where,

- q, is the estimated flow on the road in the direction of interest,
- x, is the number of vehicles traveling in the direction of interest, which are met by the observation car while traveling in the opposite direction,
- y, is the net number of cars that overtake the observation car while traveling in the direction of interest,
- ta , is the travel time taken for the trip against the stream,
- tw , is the travel time for the trip with the stream,
- $\bar{t}$, is the estimate of mean travel time in the direction of interest.

In 1973, Wright revised the theory behind this method[8]. His paper finds that the method gives biased results, although the degree of bias is not significant in practice, and can be overcome. His paper proposed that the driver should fix the journey time in advance then stops along the way would not matter. He found also that the turning traffic (exiting or entering) can upset the calculation done using this method. A suggestion also said that a large number of observations are required for reliable estimation of speeds and flow rates in order to get the precision value.

III. Road traffic Monitoring Models

Most of the road traffic monitoring system are based on motion detection to make a segmentation to the region of the image. Zhu et.al[9]. presented VISATRAM, a system for automatic traffic monitoring using 2D spatio-temporal images. A TV cam- era is mounted above a highway to monitor the traffic through two slice windows, and a panoramic view image and an epipolar plane image are formed for each lane. If the regions contain the appropriate characteristics, vehicles are considered and can be counted or tracked as desired. Out of the motion detection techniques defined so far, the two most frequently used in road-traffic monitoring are the image difference method and the motion detection technique based on features.

The other system uses image difference technique[10]. This technique based on the fact that the differences between two frames captured at different time instants reveal regions in motion. An image difference I is generated by calculating the absolute difference between two frames (I_1 and I_2) and thresholding the result.

$$I(x,y) = \begin{cases} 0, & \text{if } |I_1(x,y) - I_2(x,y)| \leqslant \theta \\ 1, & \text{otherwise} \end{cases} \quad (3)$$

where θ is an appropriate threshold. In the case of traffic monitoring, it is usual for I_1 to be the input frame and I_2 to be the reference frame (background). The reference frame is an image from the scene, without cars. The purpose of the threshold is to reduce the effects of noise and changes in the scene's lighting. The latter is a great problem in computer vision and it is usually necessary to use methods of dynamic updating of the reference frame to adapt to the scene's lighting.

Some complete traffic monitoring systems are available to date. The system that called TRIP (Traffic Research using Image Processing)[11] is designed to count vehicles traveling in a two-lane, two- way highway. The camera is placed in the highway, looking vertically onto the scene.

By using a reference image, the input images are differentiated and the result is thresholded. This leads to a binary image where vehicles in motion should appear.

The greatest effort in the development of ITS was probably made by the U.S. Department of Transportation when it requested Mitretek Systems, Inc.[1] to carry out several ITS projects. The ones that stand out are those meant for a metropolitan ITS infrastructure, to which the following systems belong: Arterial Roads Control System, Highway Control System, Transit Control System (meant for public transportation in cities), Incident Control System, Emergency Control System, Electronic Collection System for Toll Roads (freeways), Electronic Fares Payment System (in-cities), Intersections between Freeways and Railroad Tracks, Regional Information for travelers and Integrated Systems. The subject continues to arouse a lot of interest, as can be seen in the article on vehicle detection and classification[12], where a six-stage system is proposed: (1) segmentation, (2) region tracking, (3) recovery of vehicle parameters, (4) vehicle identification, (5) vehicle tracking and (6) vehicle classification.

IV. TRAFFIC FLOW ESTIMATION USING ON-DASHBOARD CAMERA IMAGE

As stated before, the purpose of this paper is to present a visual system which estimate the traffic flow by utilize digital camera which located on the car's dashboard to capture the road scenery image. Figure 1 shows the schematic diagram of the concept.

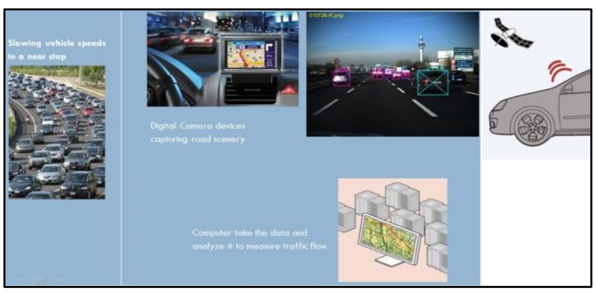

Figure. 1. Schematic diagram of the concept

This system uses a camera mounted on the car's dashboard which capture real traffic image.

V. REAL TRAFFIC EXPERIMENT

Empirical data is used in this experiment. Figure 2 shows the location of this experiment. The on-dashboard camera then capture the front scenery information. In this example, which is showed by Figure 3, we captured the real traffic situation and investigated for the opposite lane direction of the observer car.

This system uses four sub-step to identified the car. In the initial step, this system crop the raw image to get only the opposite direction lane image. This preprocessed image is showed by Figure 3. In the second step, the preprocessed image is gray-scaled. After that the threshold is applied to the image. Then the car is detected by using blob detection and some morphology step.

Figure. 2. The Location of Observation

Figure. 3. Road traffic monitoring image

Figure. 4. Preprocessed Image

In the Figure 6, the difference matrix is applied to a threshold. The gray levels greater and lower than the threshold is updated as 1 and 0, respectively. Which leads the car objects to be represented as black pixels.

Figure. 5. Real Image in gray scales

Figure. 6. Segmented Image

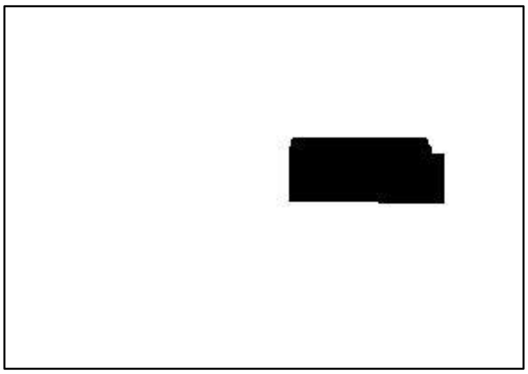

Figure. 7. Processed Image

(a)Original

(b)Canny filter summation

Figure 8. Example of images of which it is hard to count the number of cars in front of the car in concern

As shown in Figure 8 (a) and (b), it is hard to count the number of cars in front of the car in concern because some cars are overlapped just a behind of the car just in front of the car in concern. It, however, the hidden cars (occluded cars) are appeared when they are observed in curved roads or in sloped roads as shown in Figure 9 (a) and (b). Where the image (a) shows the original image while the image (b) shows Canny filter with summation is applied for segmentations.

(a)Original

(b)Canny filter summation

Figure 9. Example of images of which it is possible to count the number of cars in front of the car in concern

VI. Conclusion and Future Works

Automatic traffic flow estimation through on-dashboard camera is the new approach for the intelligent transportation system domain. This paper proposes a method for traffic measurement using the approach of collecting data from "in-vehicle" device which is by utilizing camera device that located on the car`s dashboard. The image that captured by digital camera photo then was analyzed to get the estimation value of roadway traffic flow condition. A traffic measurement using digital camera device can be a practical and low cost solution compare to the conventional traffic measurement methods.

The future step needs the effective solution on how to remove the noise background and shadow that occur on the real traffic image.

REFERENCES

[1] T. Proper, A, "Intelligent transportation system benefits," *US Dep. Transp. Washingt. DC*.

[2] A. Schadschneider, "The Nagel-Schreckenberg model revised," *Eur. Phys. J. B*, vol. 10, no. 3, pp. 573–582, 1999.

[3] K. Arai and S. R. Sentinuwo, "Effect of Driver Scope Awareness in the Lane Changing Maneuvers Using Cellular Automaton Model," *Int. J. Adv. Res. Artif. Intell.*, vol. 2, no. 7, pp. 6–12, 2013.

[4] K. Arai and S. R. Sentinuwo, "Validity of Spontaneous Braking and Lane Changing with Scope of Awareness by Using Measured Traffic Flow," *Int. J. Adv. Res. Artif. Intell.*, vol. 2, no. 7, pp. 13–17, 2013.

[5] K. Arai and S. R. Sentinuwo, "Validity of Spontaneous Braking and Lane Changing with Scope of Awareness by Using Measured Traffic Flow," *Int. J. Adv. Res. Artif. Intell.*, vol. 2, no. 7, pp. 13–17, 2013.

[6] F. Hall, "Traffic stream characteristics," Washington, DC, 1975.

[7] J. . Wardrop and G. Charlesworth, "A Method of Estimating Speed and Flow of Traffic from a Moving Vehicle," in *Proceedings of the Institution of Civil Engineers, Part II, Vol. 3*, 1954, pp. 158–171.

[8] C. Wright, "A Theretical Analysis of the Moving Observer Method," *Transp. Res. 7*, pp. 293–311, 1973.

[9] Z. Zhu, G. Xu, B. Yang, D. Shi, and X. Lin, "VISATRAM: A real-time vision system for automatic traffic monitoring," *Image Vis. Comput.*, 2000.

[10] M. Sonka, V. Hlavac, and R. Boyle, *Image processing, analysis, and machine vision*. Chapman & Hall, 1999.

[11] K. Dickinson and C. Wan, "Road traffic monitoring using the TRIP II system," *Road Traffic Monit. 1989*

[12] S. Gupte and O. Masoud, "Detection and classification of vehicles," *Intell. Transp. Syst.*, vol. 3, no. 1, pp. 37–47, 2002.

AUTHORS PROFILE

Kohei Arai 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 29 books and published 500 journal papers.

Steven Ray Sentinuwo, received the M.Eng. degree in Information Engineering, from University of Indonesia and Dr.Eng. degree in Information Science from Saga University, in 2006 and 2013, respectively. He is currently a Post-Doctoral Fellow at Information Science in Saga University, Japan. His research interest includes robot path planning, modeling and simulation, information system management, and computer vision.

Predicting Quality of Answer in Collaborative Question Answer Learning

Kohei ARAI
Graduate School of Science and Engineering,
Saga University
Saga Japan[1]

ANIK Nur Handayani
Electrical and Information Technology,
State University of Malang
Malang, Indonesia

***Abstract*— Studies over the years shown that students had actively and more interactively involved in a classroom discussion to gain their knowledge. By posting questions for other participants to answer, students could obtain several answers to their question. The problem is sometimes the answer chosen by student as the best answer is not necessarily the best quality answer. Therefore, an automatic recommender system based on student activity, may improve these situations as it will choose the best answer objectively. On the other side, in the implementation of collaborative learning, in addition to sharing information, sometimes students also need a reference or domain knowledge which relevant with the topic. In this paper, we proposed answer quality predictor in collaborative question answer (CQA) learning, to predict the quality of answer either from recommender system based on users activity or domain knowledge as reference information**.

Keywords*— collaborative question answer learning; domain knowledge; answer quality predictor; recommender.

I. Introduction

The concept of Collaborative Learning is two or more people learn or attempt to learn something together than independent. Different with individual learning, in collaborative learning student can exploit and share their resources and skills by asking, evaluating, monitoring one another's information and idea, etc [1]. Collaborative Learning is a model that knowledge can be created by sharing experiences within a population where members actively interact [2] [3]. Including both directly with face-to-face conversations [5] or using computer discussions (online forums, chat rooms, etc.) [6].

In [3] authors indicate that when they found some problem, students learn better when they learn together more frequently than working individually as members in a group. Indeed, the effectiveness of collaborative learning on the internet has been identified by various studies. Interaction among students is fostered as communication over the internet is unpretentious and convenient when addressing to a single user or multiple users. By posting questions for other participants to answer, students could obtain several answers to their question. The problem is sometimes the answer chosen by student as the best answer is not necessarily the best quality answer. The decision of an asker is influenced by subjective reasoning such as the relations between students, the asker's own point of view, his lack on the subject and others [7]. Therefore, an automatic recommender system may improve these situations as it will choose the best answer objectively. On the other side, in the implementation of collaborative learning, in addition to sharing information, sometimes students also need a reference or domain knowledge which relevant with the topic. The function of domain knowledge is used as knowledge about the environment in which the target information operates as a reference. In [11], we had developed collaborative question answer (CQA) using domain knowledge and answer quality predictor. Besides providing answer quality predictor as a recommender, the system also provides an answer that is taken from the domain knowledge as a reference.

In this paper, we proposed answer quality predictor in collaborative question answer (CQA) learning, to predict the quality of answer either from recommender system based on users activity or domain knowledge as reference information. With the proposed system right after collaborative answer, then answer quality predictor will give recommendation from the entire student's answer. And in the same time QA tools will extract answer from domain knowledge. The information from domain knowledge and answer quality predictor will be reprocess in the recommender system to predict as a bad, medium, or good answer. The paper is organized as follows. First is introduction for the question and answering system. Section 2 presents the proposed method. Section 3 explains implementation and result. Finally section 4 is summary and conclusion of this paper.

II. Proposed Method

Our proposed method in this paper consists of four parts. There are data collection, annotator, feature extraction, and coefficient correlation with answers. We explored Decision Trees classifier to get high precision on the target class (Weka framework used in this study [15]). Figure 1 shows the architecture of the system.

A. Data Collection

There are two kind of data, first is data that derived from Indonesian Yahoo! Answers (http://id.answers.yahoo.com/) and choosing the internet and computer category. This data had been processed through the answer quality predictor [9] [10]. For the domain knowledge we used id.wikipedia [8], there were over 100.000 articles in the Indonesian Wikipedia project. We collected 556 data from answer yahoo that could be processed in the QA tools to extract answer from domain knowledge [11].

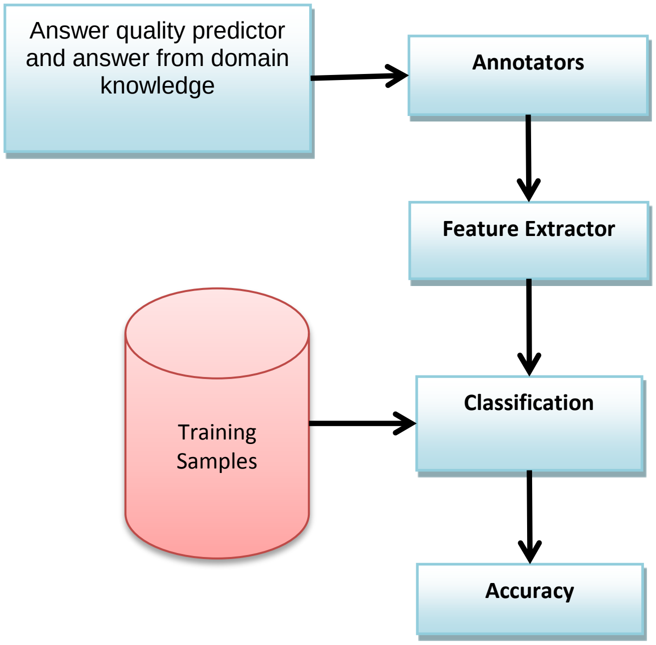

Figure1. Architecture of the System

TABLE 1. DATA COLLECTION

Category	Internet and Computer
Domain Knowledge	id.wikipedia
QA pairs	id.answeryahoo
Question	350
Answer	350
Data Training	250
Data Testing	100

The quality of a Q&A depends on the question part and answer part. For the question part we use most popular resolved question. Users could not get any useful information from bad questions. The reality bad questions always lead to bad quality answers. Therefore we decide to estimate the good answer by using annotators, and for all we got 350 Q&A pair. Table 1 shows the data collection of the research proposed.

B. *Annotator*

The quality of a question answer depends on both the question part and the answer part. It is often impossible to gather evaluative data about answers from the askers themselves, and that was the case here. The information of collaborative question answer is typically complex and subjective. We use annotators for manual judgment of answer quality and relevance. General, good answers tend to be relevant, information, objective, sincere and readable. We may separately measure these individual factors and combine scores to calculate overall the quality of the answer. Our annotators read answers, consider all of the above factors and specify the quality of answers in three levels: Bad, Medium and Good. Consider factors are as follows,

1. This answer provides enough information for the question. (informative)
2. This answer is polite (not offending). (polite)
3. This answer completely answers the whole question. (complete)
4. This is an easy to read answer. (readable)
5. This answer is relevant to the question. (relevant)
6. The answer contains enough detail. (detailed)
7. This answer is useful or helpful to address the question. (helpful)

C. *Feature Extraction*

A number of features have been identified in the literature for predicting the quality of answer. In this research, we used text feature and readability feature to predict the quality of answer. The selection of these two kinds of features based on the first Q&A pairs is already processed in the answer quality predictor; second there is internal reviewer from in term of domain knowledge system.

1) Text features are those extracted from the textual content of the articles used by [12] [13] [14]. The general intuition behind them is that a mature and good quality text is probably neither to short, which could indicate incomplete topic coverage, nor excessively long, which could indicate verbose content. We use the following features :

a) Character length: Number of characters for the answer.

b) World length: Number of words for the answer.

c) Sentences length: Number of sentences for the answer.

2) Readability Feature

These features, first used in [15], are intended to estimate the age or US grade level necessary to comprehend a text. They comprise several metrics combining counts of words and sentences. The intuition behind these features is that good information should be well written, understandable, and free of unnecessary complexity. The features are;

a) Automatic readability index (t_{rari}) : This metric was proposed in [16] and consists of using the average of word per sentence and the average of characters per words to estimate the readability.

$$t_{rari} = 4.71\frac{characters}{words} + 0.5\frac{words}{sentences} - 21.43 \ldots\ldots (1)$$

b) Coleman Liau (t_{rcl}): This metric was proposed in [84] and consists of the average of characters per word and the number of sentences in a fragment of 100 words (w_f).

$$t_{rari} = 5.89\frac{characters}{words} + 0.5\,wf - 15.48 \ldots\ldots\ldots\ldots (2)$$

c) The compound is a compound word or a combination of all of the basic morpheme that existed as a word that has a specifically pattern of phonological, grammatical, and semantic according to the rules of the language. The specific pattern distinguishes as a phrase or combination of words. For example, in Indonesian, kamar mandi is a compound word, while the baju hijau is the phrase while in English; the blackbird is a compound word, while the black bird is the phrase. [17]

d) Loan word is a word that derived from foreign languages that has been integrated into an Indonesian and generally accepted to be used. Indonesian has absorbing many words from other languages, especially those that have direct *contact with the community, either through tarding (Sanskrit, Chinese, Arabic), or colonialism (Portuguese, Dutch, Japanese), as well as the development of science (English). [17]*

e) Abstract noun is a type of noun (to explain the names of objects) which the existence could not be captured using human eyes and can only be imagined. The examples of abstract nouns are science, dreams, ideas, inspiration, happiness and others. [17]

f) Conjunction:In grammar, conjunction (abbreviated conj or cnj) is a part of speech that connects two words, sentences, phrases or clauses.

A discourse connective is a conjunction joining sentences. This definition may overlap with that of other parts of speech, so what constitutes a "conjunction" must be defined for each language. In general, a conjunction is an invariable grammatical particle, and it may or may not stand between the items it conjoins.[17]

D. Correlation Coefficient

The function of the correlation coefficient is to know how closely one variable is related to another variable [18], in this case the correlation between individual features and the annotators scores (good answers have higher scores: Bad = 0, Medium = 1, Good = 2). Table 2 shows coefficient correlation with the quality of answer. Surprisingly, all of feature has the strongest correlation with the quality of the answer, except for Auto read index and Coleman liau index.

From the calculation of Corr, we can see that text feature and readability feature affects the quality of the answers. In this study (Computer and Internet) using a lot of loan word (eg. Computer → komputer, Processor → prosesor, etc) , abstract noun word (eg. Principle →prinsip, definition → definisi, etc), and compound word (database, how it work →carakerja, etc). Auto read index and Coleman liau index is a feature used to calculate English language readability parameter. Several intelligent of United States of America used this parameter to measure readability of electronic letter. This convinced us that the character of one language is different one another.

TABLE 2. COEFFICIENT CORRELATION

Features	**Correlation**
Number of loan word	0.948
Number of Abstract noun word	0.999
Number of conjunction	0.828
Number of compound word	0.861
Number of char	0.971
Number of word	0.961
Number of sentence	0.928
Auto read index	0.0088
Coleman liau index	0.0002

The formula for Pearson's Correlation Coefficient:

$$r_{xy} = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{n_x}\right]\left[\sum Y^2 - \frac{(\sum Y)^2}{n_y}\right]}} \quad (3)$$

III. IMPLEMENT AND RESULT

We will implement the proposed methods to the Q&A pair of data. There are four kind data for the classification, first is data that acquired from the entire feature, data with high correlation (> 0.1 and > -0.1), text feature data, and readability feature data. We build the predictor using 250 training data and 100 testing data.

TABLE 3. ACCURACY OF ALL FEATURE AND CORR FEATURE

Feature	Training Set	Data Test	CV		
			5	10	15
All Feature	93.6	72.8	88.4	88.4	88.4
Corr Feature	93.2	71.2	88	88.4	88.4
Text Feature	90.4	70.4	84.4	84	83.2
Readability Feature	93.2	73.6	87.6	86.4	87.6

Table 3 shows prediction accuracy for the different implementation of answer quality, in particular comparing the choice in classifier algorithm, feature sets (using all feature, Correlation feature, text feature, and readability feature) and test option. By using C4.5 results, the best performance of the entire variant feature is all features with 93.6 of accuracy slightly adrift of 0.4 with Corr feature. We can conclude that text feature is a part of readability feature because some text feature parameter are count character, word, and sentence. For the count word parameter, in the computer and internet subject mostly users used Loan word. Another interesting result from the table 3 we could see that the differences between all features, Correlation feature and readability feature, is not too significant for accuracy it is about 0,4. This indicates that feature which does not have high correlation is not too pretty significant impact for classification results. Figure.2 shows the result classification on training data using Weka framework.

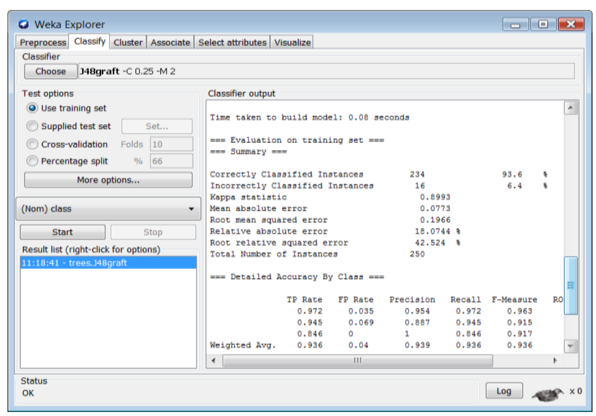

Figure 2. Result of Classification on Training Data

IV. CONCLUSION

In this paper we presented our knowledge to quantify and predict quality of answer in collaborative question answer (CQA) learning, especially for Indonesian. Beyond developing models to select best answer and evaluate the quality of answers, there are several important lessons to learn here for measuring content quality in CQA. We find that domain knowledge information sometime isn't providing good answer. On the other side students answer better than domain knowledge answer.

With appropriate features, we could build models that could have significantly higher probability of identifying the best answer class than classifying a non-best answer. From the entire system from 9 features, we conclude as following:

1) *From the four existing feature, the highest accuracy exist on all feature set (comparing with correlation coefficient set, text feature set and readability feature set).*

2) *The best performance for all feature set by using C4.5 classifier, with averaged accuracy 93.6 for training set, 72.8 data test and 88.4 for cross validation.*

In the future our models and predictions could be useful for predictor quality information as a recommender system to complete collaborative question answer learning.

ACKNOWLEDGMENT

The authors wish to thank to the students of Electrical and Information Technology State University of Malang Indonesia, who contributed and supported to the experiments.

REFERENCES

[1] Dillenbourg, P. (1999). Collaborative Learning: Cognitive and Computational Approaches. Advances in Learning and Instruction Series. New York, NY: Elsevier Science, Inc.

[2] Chiu, M. M. (2000). Group problem solving processes: Social interactions and individual actions. Journal for the Theory of Social Behavior, 30, 1, 27-50.600-631.

[3] Chiu, M. M. (2008).Flowing toward correct contributions during groups' mathematics problem solving: A statistical discourse analysis. Journal of the Learning Sciences, 17 (3), 415 – 463

[4] Mitnik, R., Recabarren, M., Nussbaum, M., & Soto, A. (2009). Collaborative Robotic Instruction: A Graph Teaching Experience. Computers & Education, 53(2), 330-342.

[5] Chiu, M. M. (2008). Effects of argumentation on group micro-creativity. Contemporary Educational Psychology, 33, 383 – 402.

[6] Chen, G., & Chiu, M. M. (2008). Online discussion processes. Computers and Education, 50, 678 – 692

[7] Wang, C.C., Hung J.C., Yang C.Y., Shih T.K. (2006). An Apllication of Question Answering System for Collaborative Learning. IEEE Conference on ICDCSW'06

[8] Kohei Arai, Anik Nur Handayani, Question Answering System for an Effective Collaborative Learning. (IJACSA) Vol 3. No.1, 2012. Page 60-64.

[9] Kohei Arai, Anik Nur Handayani, Predicting Quality of Answer in Collaborative Q/A Community. (IJARAI), Vol 2. No.3, 2013. Page 21-25.

[10] Kohei Arai, Anik Nur Handayani, Question Answering for Collaborative Learning with Answer Quality Predictor. (IJMECS), Vol 5. No.5, 2013. Page 12-17.

[11] Kohei Arai, Anik Nur Handayani, Collaborative Question Answering System Using Domain Knowledge and Answer Quality Predictor (IJMECS), Vol. 5 Number 11, 2013. Page 21-27

[12] Watzlawick, P. (1967) Pragmatics of Human Communications: A Study of Interactional Patterns. Pathologies and Paradoxes. W.W. Norton, New York.

[13] Smith, B. L., & MacGregor, J. T. (1992). "What Is Collaborative Learning?". National Center on Postsecondary Teaching, Learning, and Assessment at Pennsylvania State University

[14] Chiu, M. M. (2004). Adapting teacher interventions to student needs during cooperative learning. American Educational Research Journal, 41, 365-399.

[15] L. Rassbach, T. Pincock, and B. Mingus. Exploring the feasibility of automatically rating online article quality. http://upload.wikimedia.org/wikipedia/wikimania2007/d/d3/RassbachPincockMingus07.pdf.

[16] E. A. Smith and R. J. Senter. Automated readability index. 1967.

[17] http://id.wiktionary.org/wiki/Kategori:Kata_bahasa_Indonesia

[18] Chiu, M. M. (2000). Group problem solving processes: Social interactions and individual actions. for the Theory of Social Behavior, 30, 1, 27-50.600-631.

AUTHORS PROFILE

Kohei Arai 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 26 books and published 500 journal papers.

AnikNurHandayanireceived the B.E. degree in electronics engineering from Brawijaya University, and the M.S. degree in ElectricalEngineering, from Institute of Technology SepuluhNopember, Surabaya, Indonesia, in 2004 and 2008, respectively. She is currently a PhD Student at Information Science in Saga University, Japan.

Preliminary Study on Phytoplankton Distribution Changes Monitoring for the Intensive Study Area of the Ariake Sea, Japan Based on Remote Sensing Satellite Data

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

Toshiya Katano
Institute of Lowland and Marine Research
Saga University
Saga City Japan

***Abstract*—Phytoplankton distribution changes in the Ariake Sea areas, Japan based on remote sensing satellite data is studied. Through experiments with Terra and AQUA MODIS data derived chlorophyll-a concentration and suspended solid as well as truth data of chlorophyll-a concentration together with meteorological data and tidal data which are acquired 7 months from October 2012 to April 2013, it is found that strong correlation between the truth data of chlorophyll-a and MODIS derived chlorophyll-a concentrations with R square value ranges from 0.677 to 0.791. Also it is found that the relations between ocean wind speed and chlorophyll-a concentration as well as between tidal effects and chlorophyll-a concentration. Meanwhile, there is a relatively high correlation between sunshine duration a day and chlorophyll-a concentration.**

Keywords-chlorophyl-a concentration; suspended solid; ocean winds.

I. INTRODUCTION

Due to red tide contaminations, water color is changed by an algal bloom. In accordance with increasing of phytoplankton concentration, sea surface color changes from blue to green as well as to red or brown depending on the majority of phytoplankton (Dierssen et al, 2006) so that it is capable to detect red tide using this color changes [1].

MODIS ocean color bands data is used for red tide detection. An iterative approach (Arnone et al., 1998 [2]; Stumpf et al., 2003 [3]) for sediment-rich waters, based on the Gordon and Wang (1994) algorithm [4], is used to correct for the atmospheric interference in the six ocean color bands in turbid coastal waters to obtain water leaving radiance, which are then used in the band-ratio algorithm (O'Reilly et al., 2000 [5]) to estimate Chlorophyll in unit of mg m^{-3}. Also suspended solid is estimated with two bands algorithm (visible minus near infrared bands data). The multi-channels of red tide detection algorithms (in the formula of $C=(R_i-R_j)/(R_k-R_l)$ where R_i, R_j, R_k and R_l are the reflectivity derived from bands *i*, *j*, *k* and *l*.) are proposed. Also learning approaches based on k-nearest neighbors, random forests and support vector machines have been proposed for red tide detection with Moderate Resolution Imaging Spectro-radiometer: MODIS satellite images (Weijian C.,et al.,2009) [6].

Satellite based red tide detection does work under a fine weather condition but not under cloudy and rainy conditions obviously. Furthermore, revisit period of fine resolution of radiometer onboard satellite orbits are longer than typical red tide propagations so that it is not enough observation frequency if only remote sensing satellite is used for red tide detections. Therefore satellite-and ground-based red tide monitoring system is proposed [7]. In the ground based red tide monitoring system, green colored filtered camera and polarization camera are featured for detection of red tide and discrimination of red tide types [8].

The Ariake Sea is the largest productive area of Nori (*Porphyra yezoensis*) in Japan. In winters of 2012 and 2013, a massive diatom bloom occurred in the Ariake Sea, Japan (Ito et al. 2013). In case of above red tides, bloom causative was *Eucampia zodiacus*. This bloom had been firstly developed at the eastern part of the Ariake Sea. However, as the field observation is time-consuming, information on the developing process of the red tide, and horizontal distribution of the red tide has not yet been clarified in detail. To clarify the horizontal distribution of red tide, and its temporal change, remote sensing using satellite data is quite useful.

In this paper, the chlorophyll-a concentration algorithm developed for MODIS is firstly validated. Then apply the algorithm to MODIS data which are acquired at the Ariake Sea areas, Japan specifically. Also the relations between tidal effects and chlorophyll-a concentration as well as between ocean wind speed and chlorophyll-a concentration together with between sunshine duration a day and chlorophyll-a concentration.

In the next section, the method and procedure of the experimental study is described followed by experimental data and estimated results. Then conclusion is described with some discussions.

II. METHOD AND PROCEDURE

A. *The Procedure*

The procedure of the experimental study is as follows, *1) Gather the truth data of chlorophyll-a concentration measured at the observation towers in the Ariake Sea areas together with the corresponding areas of MODIS derived chlorophyll-a concentration,*

2) Gather the meteorological data which includes sunshine duration a day, ocean wind speed and direction, tidal heights,

3) Correlation analysis between the truth data and MODIS derived chlorophyll-a concentration as well as between geophysical parameters, ocean wind speed, sunshine duration a day, tidal heights and chlorophyll-a concentration is made.

B. MODIS

MODIS stands for The Moderate Resolution Imaging Spectro-radiometer which allows observations of the earth's surface with a variety of wavelength regions ranges from visible to thermal infrared with spatial resolution of 250, 500 and 1000 meters. Major specifications of MODIS are shown in Table 1 and 2.

MODIS is carried on Terra and AQUA satellites with local mean time of 10:30 a.m. and 1:30 p.m. with 2330km of swath width. Therefore, it may cover entire globe within a day. MODIS data is applicable to land surface, ocean surface, and atmosphere observations as is shown in Table 2. In particular for ocean surface observations, bands 8 to 16 data are useful.

TABLE I. MAJOR SPECIFICATIONS OF MODIS[1]

Orbit:	705 km, 10:30 a.m. descending node (Terra) or 1:30 p.m. ascending node (Aqua), sun-synchronous, near-polar, circular
Scan Rate:	20.3 rpm, cross track
Swath Dimensions:	2330 km (cross track) by 10 km (along track at nadir)
Telescope:	17.78 cm diam. off-axis, afocal (collimated), with intermediate field stop
Size:	1.0 x 1.6 x 1.0 m
Weight:	228.7 kg
Power:	162.5 W (single orbit average)
Data Rate:	10.6 Mbps (peak daytime); 6.1 Mbps (orbital average)
Quantization:	12 bits
Spatial Resolution:	250 m (bands 1-2) 500 m (bands 3-7) 1000 m (bands 8-36)
Design Life:	6 years

TABLE II. APPLICATION OF MODIS SENSOR DATA[2]

Primary Use	Band	Bandwidth[1]	Spectral Radiance[2]	Required SNR[3]
Land/Cloud/Aerosols Boundaries	1	620 - 670	21.8	128
	2	841 - 876	24.7	201
Land/Cloud/Aerosols Properties	3	459 - 479	35.3	243
	4	545 - 565	29.0	228
	5	1230 - 1250	5.4	74
	6	1628 - 1652	7.3	275
	7	2105 - 2155	1.0	110
Ocean Color/ Phytoplankton/ Biogeochemistry	8	405 - 420	44.9	880
	9	438 - 448	41.9	838
	10	483 - 493	32.1	802
	11	526 - 536	27.9	754
	12	546 - 556	21.0	750
	13	662 - 672	9.5	910
	14	673 - 683	8.7	1087
	15	743 - 753	10.2	586
	16	862 - 877	6.2	516
Atmospheric Water Vapor	17	890 - 920	10.0	167
	18	931 - 941	3.6	57
	19	915 - 965	15.0	250
Primary Use	**Band**	**Bandwidth[1]**	**Spectral Radiance[2]**	**Required NE[delta]T(K)[4]**
Surface/Cloud Temperature	20	3.660 - 3.840	0.45(300K)	0.05
	21	3.929 - 3.989	2.38(335K)	2.00
	22	3.929 - 3.989	0.67(300K)	0.07
	23	4.020 - 4.080	0.79(300K)	0.07
Atmospheric Temperature	24	4.433 - 4.498	0.17(250K)	0.25
	25	4.482 - 4.549	0.59(275K)	0.25
Cirrus Clouds Water Vapor	26	1.360 - 1.390	6.00	150(SNR)
	27	6.535 - 6.895	1.16(240K)	0.25
	28	7.175 - 7.475	2.18(250K)	0.25
Cloud Properties	29	8.400 - 8.700	9.58(300K)	0.05
Ozone	30	9.580 - 9.880	3.69(250K)	0.25
Surface/Cloud Temperature	31	10.780 - 11.280	9.55(300K)	0.05
	32	11.770 - 12.270	8.94(300K)	0.05
Cloud Top Altitude	33	13.185 - 13.485	4.52(260K)	0.25
	34	13.485 - 13.785	3.76(250K)	0.25
	35	13.785 - 14.085	3.11(240K)	0.25
	36	14.085 - 14.385	2.08(220K)	0.35

[1] Bands 1 to 19 are in nm; Bands 20 to 36 are in µm
[2] Spectral Radiance values are (W/m^2 -µm-sr)
[3] SNR = Signal-to-noise ratio
[4] NE(delta)T = Noise-equivalent temperature difference

Note: Performance goal is 30-40% better than required

There are following many ocean related products[3],

Angstrom Exponent
Aerosol Optical Thickness

[1] http://modis.gsfc.nasa.gov/about/specifications.php
[2] http://modis.gsfc.nasa.gov/about/specifications.php
[3] http://oceancolor.gsfc.nasa.gov/

Chlorophyll a
Downwelling diffuse attenuation coefficient at 490 nm
Level 2 Flags
Photosynthetically Available Radiation
Particulate Inorganic Carbon
Particulate Organic Carbon
Sea Surface Temperature Quality
Sea Surface Temperature Quality - 4um
Remote Sensing Reflectance
Sea Surface Temperature
Sea Surface Temperature 4um

In this study, chlorophyll-a product is used for the experiments.

C. *The Intensive Study Areas*

Figure 1 shows the intensive study areas in the Ariake Sea area, Kyushu, Japan.

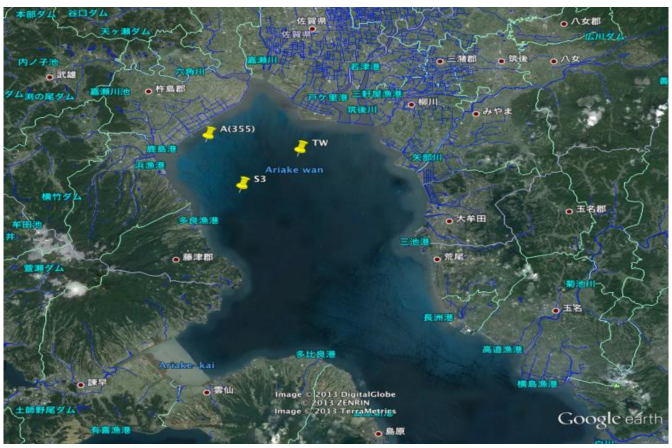

Fig 1. The Intensive study areas (Yellow pins shows the areas)

There are three observation tower points, TW, S, and A. TW is closely situated to the Saga Ariake Airport and is situated near the river mouth. On the other hand, A is situated most closely to the coastal area while S is situated in the middle point of the Ariake Sea width and is situated most far from the coastal areas and river mouths.

III. EXPERIMENTS

A. *The Data Used*

The truth data of chlorophyll-a concentration measured at the observation towers in the intensive study areas in the Ariake Sea areas together with the corresponding areas of MODIS derived chlorophyll-a concentration which area acquired for the observation period of 7 months during from October 2012 to April 2013 are used for the experiments. Also, the meteorological data which includes sunshine duration, ocean wind speed and direction, tidal heights which are acquired for the 7 months are used for the experiments.

Figure 2 shows an example of the chlorophyll-a concentration image which is derived from MODIS data which is acquired on 25 February 2013.

The chlorophyll-a concentration measured at the tower, TW ranges from 52 to 64 ug/l (the highest value in the 7 months). This is red tide (Phytoplankton) blooming period. Such this MODIS derived chlorophyll-a concentration data are available almost every day except cloudy and rainy conditions.

Blooming is used to be occurred when the seawater becomes nutrient rich water, calm ocean winds, long sunshine duration after convection of seawater (vertical seawater current from the bottom to sea surface). Therefore, there must exists relations between the geophysical parameters, ocean wind speed, sunshine duration, tidal heights and chlorophyll-a concentration.

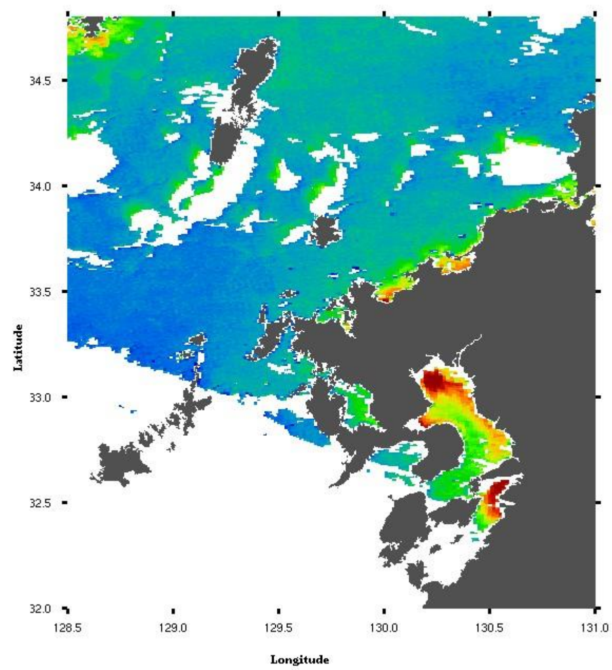

Fig 2. An example of the chlorophyll-a concentration image which is derived from MODIS data which is acquired on 25 February 2013

B. *The Relation Between Truth Data and MODIS Derived Colorophyl-a Concentrations*

In order to check validity of the chlorophyll-a concentration estimation algorithm, the relation between truth and MODIS derived chlorophyll-a concentration is investigated. Figure 3 shows the relation between these for intensive study area of S while Figure 4 shows the relation for intensive study area of TW.

In the figures, L and H denote Low and High chlorophyll-a concentration of the day. Also Ch-a denotes chlorophyll-a concentration in unit of μg/l. Linear (L) denotes linear regression line for minimum chlorophyll-a concentration of the day while Linear (H) denotes that for maximum chlorophyll-a concentration of the day.

C. *Trend Analysis*

MODIS derived chlorophyll-a concentration which are acquired from January 4 2013 to February 28 2013 are shown in Figure 5.

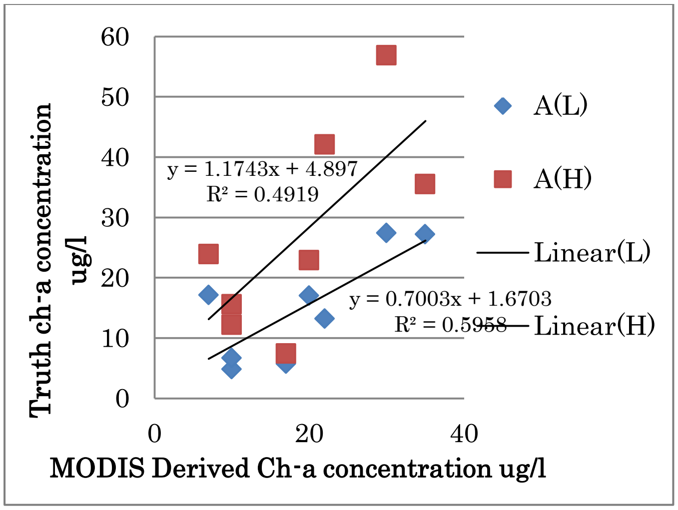

Fig 3. Relation between truth and MODIS derived chlorophyll-a concentrations for the intensive study area of S

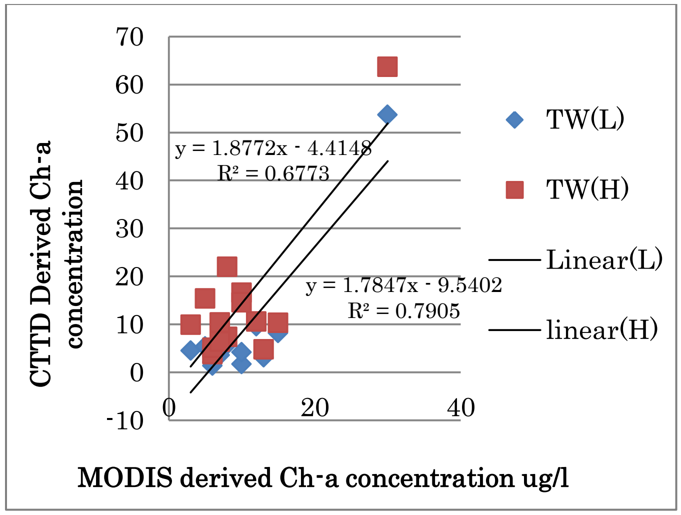

Fig 4. Relation between truth and MODIS derived chlorophyll-a concentrations for the intensive study area of TW

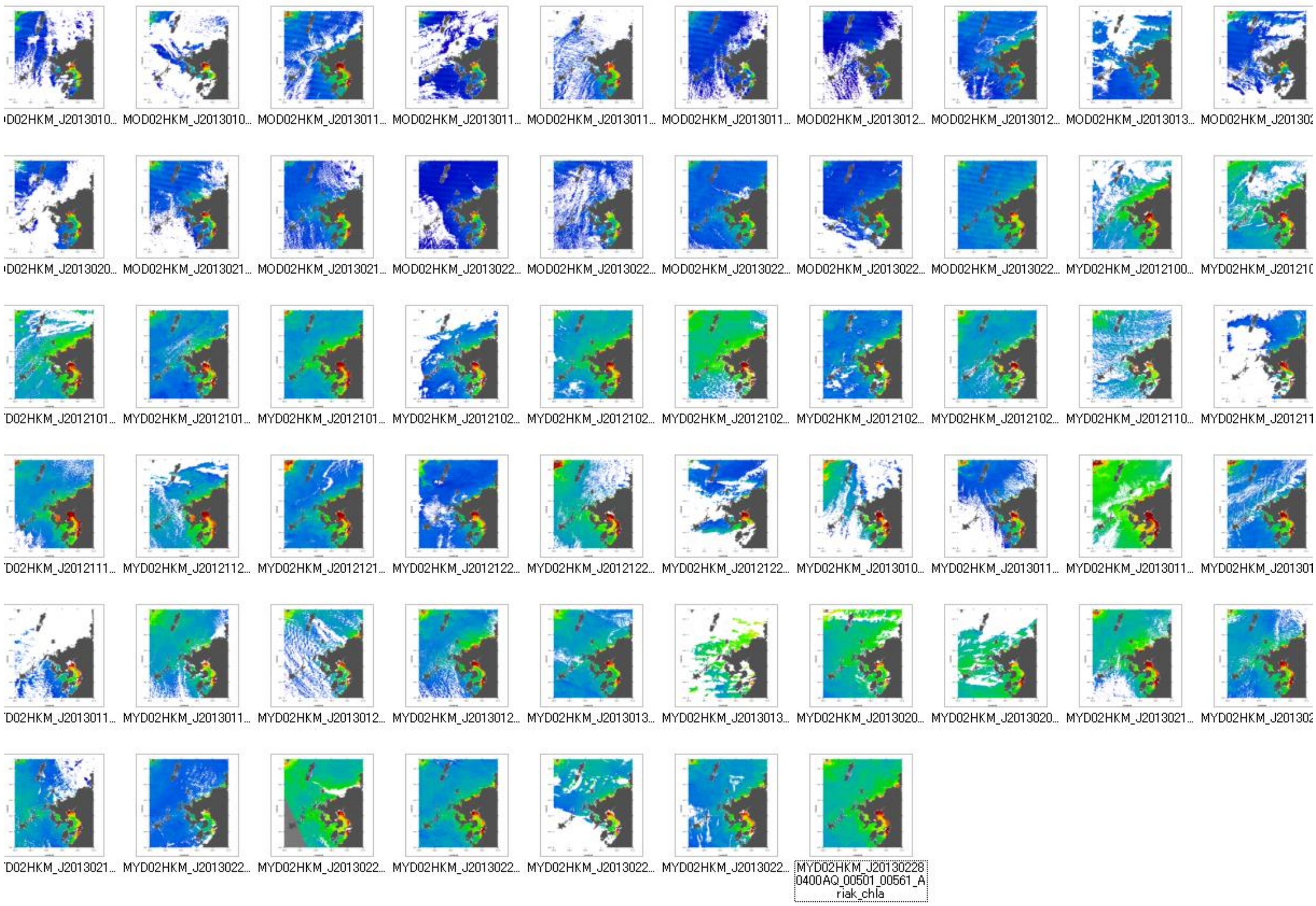

Fig 5. MODIS derived chlorophyll-a concentration in the intensive study areas which are acquired from January 4 2013 to February 28 2013

A massive red tide by large diatom of Eucampia zodiacus occurred from February to April, 2013 in the innermost are of the Ariake Sea. The bloom developing process are clearly detected in Figure 5.

In the first half of the Figure 5, surface chlorophyll-a concentration was low in the sea; however, it obviously increased in the middle of the Figure 5.

Finally, in the later half of the figure, quite high concentrations of chlorophyll-a were detected in the innermost area of the sea, as the Eucampia population developed. To clarify the developing mechanism of this red tide, other environmental variables such as temperature, salinity, nutrient concentrations should be further investigated.

D. Relations between geophysical parameters and chlorophyll-a concentrations

Nutrient rich seawater near the bottom is used to be flown from the sea bottom to the sea surface when the seawater convection is occurred due to ocean wind, and tidal effect. Then chlorophyll-a concentration increases in accordance with sunshine duration with the nutrient rich sea surface water. Therefore, there must exist some relations between the geophysical parameter of ocean wind, tidal effect and sunshine duration.

Figure 6 shows a relatively high relation between accumulated sunshine duration for the past three days and chlorophyll-a concentration. Ocean wind is not so strong, while tidal effect is also not so high. Therefore the relation between accumulated sun appearances time duration is proportional to the truth data of chlorophyll-a concentration.

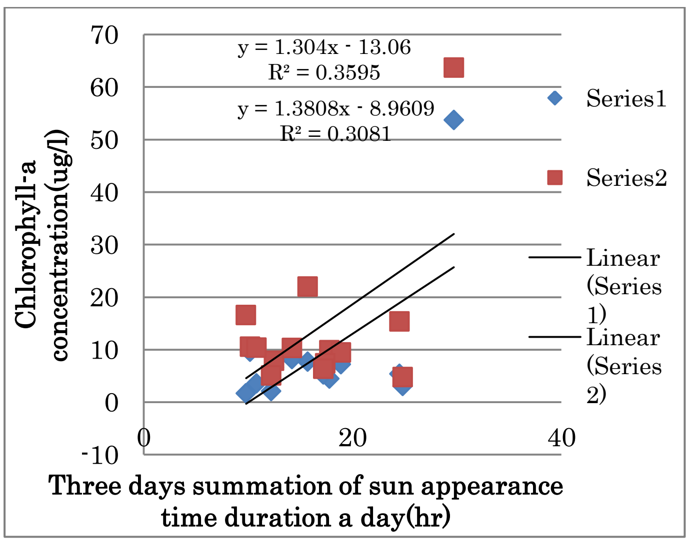

Fig 6. Relation between accumulated sunshine duration a day for the past three days and chlorophyll-a concentration

It can also be said that there are very weak relation between chlorophyll-a concentration and ocean wind as well as sea level difference a day due to the fact that ocean wind is not so high while tidal effect is not so large. Figure 7 shows the relation between chlorophyll-a concentration and accumulated daily averaged ocean wind speed for the past two days.

Meanwhile, Figure 8 shows the relation between chlorophyll-a concentration and sea surface level difference a day while Figure 9 shows the relation between chlorophyll-a concentration and the accumulated sea surface level difference for the past two days.

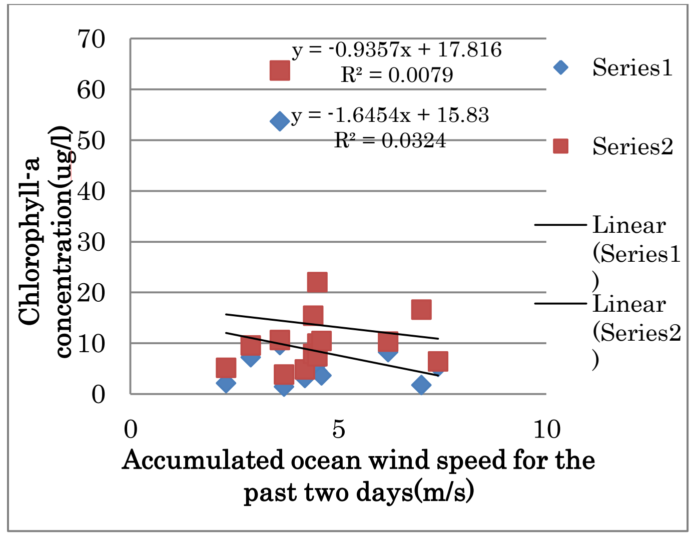

Fig 7. Relation between chlorophyll-a concentration and accumulated daily averaged ocean wind speed for the past two days.

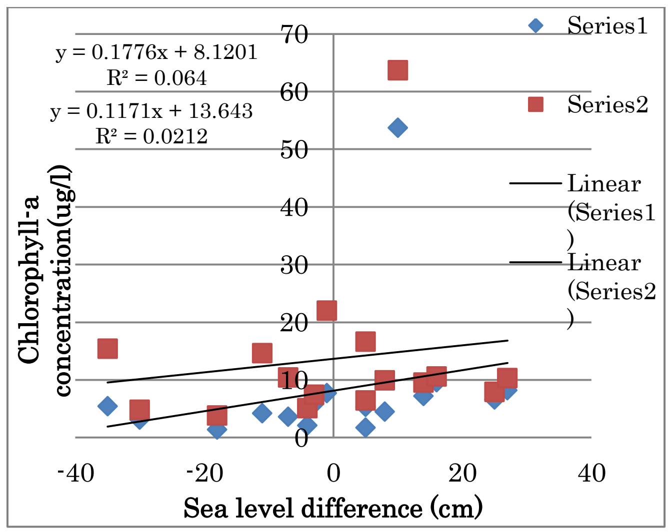

Fig 8. Relation between chlorophyll-a concentration and sea surface level difference a day

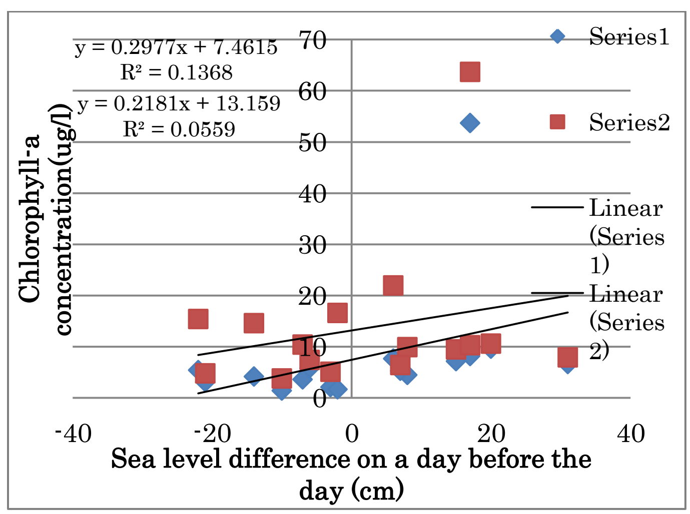

Fig 9. Relation between chlorophyll-a concentration and the accumulated sea surface level difference for the past two days.

It is confirmed that relation between ocean wind speed and chlorophyll-a concentration is weak. Also, it is confirmed that relation between sea surface level difference a day and chlorophyll-a concentration is weak. It may said that relation between two days accumulated sea surface level difference a day and chlorophyll-a concentration is higher than that between one day sea surface level difference and chlorophyll-a concentration. That is because of the fact that there is time delay of chlorophyll-a increasing after the nutrient rich bottom seawater is flown to the sea surface.

IV. CONCLUSION

Preliminary study on phytoplankton distribution changes monitoring for the intensive study area of the Ariake Sea, Japan based on remote sensing satellite data is conducted. Phytoplankton distribution changes in the Ariake Sea areas, Japan based on remote sensing satellite data is studied.

Through experiments with Terra and AQUA MODIS data derived chlorophyll-a concentration and suspended solid as well as truth data of chlorophyll-a concentration together with meteorological data and tidal data which are acquired 7 months from October 2012 to April 2013, it is found that strong correlation between the truth data of chlorophyll-a and MODIS derived chlorophyll-a concentrations with R square value ranges from 0.677 to 0.791.

Also it is found that the relations between ocean wind speed and chlorophyll-a concentration as well as between tidal effects and chlorophyll-a concentration. Meanwhile, there is a relatively high correlation between sunshine duration a day and chlorophyll-a concentration.

As the results from the experiments, it is found that the followings,

1) the accumulated sun appearances time duration is proportional to the truth data of chlorophyll-a concentration

2) it is confirmed that relation between ocean wind speed and chlorophyll-a concentration is weak,

3) it is confirmed that relation between sea surface level difference a day and chlorophyll-a concentration is weak,

4) it may said that relation between two days accumulated sea surface level difference a day and chlorophyll-a concentration is higher than that between one day sea surface level difference and chlorophyll-a concentration,

5) there is a time delay of chlorophyll-a increasing after the nutrient rich bottom seawater is flown to the sea surface.

ACKNOWLEDGMENT

The authors would like to thank Dr. Yuichi Hayami of Institute of Lowland and Marine Research, Saga University for his great supports through the experiments.

REFERENCES

[1] Dierssen H.M., R.M.Kudela, J.P.Ryan, R.C.Zimmerman, Red and black tides: Quantitative analysis of water-leaving radiance and perceived color for phytoplankton, colored dissolved organic matter, and suspended sediments, Limnol. Oceanogr., 51(6), 2646–2659, E 2006, by the American Society of Limnology and Oceanography, Inc., 2006.

[2] Arnone, R. A., Martinolich, P., Gould, R. W., Jr., Stumpf, R., & Ladner, S., Coastal optical properties using SeaWiFS. Ocean Optics XIV, Kailua Kona, Hawaii, USA, November 10–13, 1998. SPIE Proceedings., 1998.

[3] Stumpf, R. P., Arnone, R. A., Gould Jr., R. W., Martinolich, P. M., & Martinuolich, V., A partially coupled ocean-atmosphere model for retrieval of water-leaving radiance from SeaWiFS in coastal waters. In S. B. Hooker, & E. R. Firestone (Eds.), SeaWiFS Postlaunch Tech. Report Series. NASA Technical Memorandum, 2003-206892, vol. 22 (p. 74), 2003.

[4] Gordon, H. R., & Wang, M., Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: A preliminary algorithm. Applied Optics, 33, 443–452, 1994.

[5] O'Reilly, J. E., Maritorena, S., Siegel, D. A., O'Brien, M. C., Toole, D.,Chavez, F. P., et al., Ocean color chlorophyll a algorithms for SeaWiFS, OC2, and OC4: Version 4. In B. Hooker, & R. Firestone (Eds.), SeaWiFS Postlaunch Tech. Report Series. NASA Technical Memorandum 2000-206892, vol. 11 (p. 2000), 2000.

[6] Weijian C., Hall, L.O., Goldgof, D.B., Soto, I.M., Chuanmin H, Automatic red tide detection from MODIS satellite images, Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on SMC, 2009.

[7] Kohei Arai and Yasunori Terayama, Polarized radiance from red tide, Proceedings of the SPIE Asia Pacific Remote Sensing, AE10-AE101-14, Invited Paper, 2010

[8] Kohei Arai, *Red tides*: combining satellite- and ground-based detection. 29 January 2011, *SPIE Newsroom*. DOI: 10.1117/2.1201012.003267, http://spie.org/x44134.xml?ArticleID=x44134

[9] Graham LE, Wilcox LW. (2000) Algae. pp. 232-68. Prentice Hall. ISBN 0-13-660333-5

[10] Kroger N, Deutzmann R, Sumper M (1999). "Polycationic peptides from diatom biosilica that direct silica nanosphere formation". *Science* **286** (5442): 1129-32.

[11] Martin-Jezequel V, Hildebrand M, Brzezinski MA (2000). "Silicon metabolism in diatoms: Implications for growth". *J Phycol* **36**: 821-40.

[12] Adl, Sina M.; *et al.* (2005), "The New Higher Level Classification of Eukaryotes with Emphasis on the Taxonomy of Protists", *J. Eukaryot. Microbiol.* **52** (5): 429–493

[13] Adl, Sina M.; *et al.* (2012), "The Revised Classification of Eukaryotes", *J. Eukaryot. Microbiol.* **59** (5): 429–493

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 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 500 journal papers

Using Unlabeled Data to Improve Inductive Models by Incorporating Transductive Models

ShengJun Cheng
School of Computer Science and Technology
Harbin Institute of Technology
Harbin 150001, China
Email: http://hitwer@gmail.com

Jiafeng Liu
Harbin Institute of Technology
Harbin150001, China

XiangLong Tang
Harbin Institute of Technology
Harbin150001, China

***Abstract*—This paper shows how to use labeled and unlabeled data to improve inductive models with the help of transductive models. We proposed a solution for the self-training scenario. Self-training is an effective semi-supervised wrapper method which can generalize any type of supervised inductive model to the semi-supervised settings. it iteratively refines a inductive model by bootstrap from unlabeled data. Standard self-training uses the classifier model(trained on labeled examples) to label and select candidates from the unlabeled training set, which may be problematic since the initial classifier may not be able to provide highly confident predictions as labeled training data is always rare. As a result, it could always suffer from introducing too much wrongly labeled candidates to the labeled training set, which may severely degrades performance. To tackle this problem, we propose a novel self-training style algorithm which incorporate a graph-based transductive model in the self-labeling process. Unlike standard self-training, our algorithm utilizes labeled and unlabeled data as a whole to label and select unlabeled examples for training set augmentation. A robust transductive model based on graph markov random walk is proposed, which exploits manifold assumption to output reliable predictions on unlabeled data using noisy labeled examples. The proposed algorithm can greatly minimize the risk of performance degradation due to accumulated noise in the training set. Experiments show that the proposed algorithm can effectively utilize unlabeled data to improve classification performance.**

***Keywords*—Inductive model, Transductive model, Semi-supervised learning, Markov random walk.**

I. Introduction

Traditional inductive models like Naive Bayes, CARTs[1], Support Vector Machines are always in supervised settings, which means these model can only be trained on labeled data. Training a good inductive model needs enough labeled examples. Unfortunately, preparing labeled data for such task is often expensive and time consuming, while unlabeled data are readily available. This was the major motivation that led to the arise of semi-supervised paradigm which utilizes few labeled examples and vast amounts of cheap unlabeled examples to learns a model. Semi-supervised learning has achieved considerable success in a wide variety of domains, existing semi-supervised learning methods can be roughly categorized into several paradigms[2], including generative models, semi-supervised support vector machines (S3VMs), graph based methods and bootstrapping wrapper method.

Self-training[3] is a simple and effective semi-supervised algorithm which has been successfully applied to various real-world tasks. It is an wrapper method, which means it can generalize any type of supervised inductive model to the semi-supervised settings[4]. Self-training initially trains a classifier on labeled data and then iteratively augments its labeled training set by adding several newly pseudo-labeled unlabeled examples with most confident predictions of its own. Standard self-training uses the classifier model(trained on labeled examples) to label and select candidates from the unlabeled training set, which may be problematic since the initial classifier may not be able to provide highly confident predictions as labeled training data is always rare. In addition, since self-training utilizes unlabeled data in an incremental manner, early noise introduced to training sets would be reinforced round by round, resulting in severe performance degradation. Although some techniques, e.g. data editing[5], have been employed to alleviate this noise-related problem[6], results are yet undesirable. As a result, it could always suffer from introducing too much wrongly labeled candidates to the labeled training set, which may severely degrades performance. Another drawback of self-training is that the newly added examples are not informative to the current classifier, since they can be classified confidently[7]. As a result, they may only help increase the classification margin, without actually providing any novel information to the current classifier.

In this paper, we show how to use unlabeled data to improve inductive models with the help of transductive models. We proposed a solution for the self-training scenario, a novel self-training style algorithm is proposed. Generally, unlike traditional self-training only using labeled data to label and select unlabeled example for training set augmentation, our algorithm utilizes both labeled and unlabeled data to facilitate the self-labeling process. In detail, all the labeled and unlabeled examples are presented as a graph, where a novel markov random walk with constrains is proposed to label all examples on graph in a transductive setting[8]. This graph-based method satisfy *manifold assumption*that examples with high similarities in the input space should share similar labels. Typically, Most graph based methods output label information to unlabeled data in a transductive setting such as Label propagation, markov random walks, Low density separation[9]. Those methods are designed to utilize unlabeled by representing both the data as a graph, with examples as vertices and similarities of examples as edges. Existing transductive graph-based methods assume all labels on labeled data correct, can not work under training sets subject to noise. While our transductive model can naturally deals with noisy labeled data, which utilize ”label smooth" to automatically adjust the potential wrong labels. By incorporating this transductive model to the self-training process, we expect any applied supervised inductive model can be greatly improved.

The main contribution of this paper can be summarized as follows:

- We show that incorporating transductive models to inductive models in semi-supervised settings can improve classification performance.
- We propose a novel self-training algorithm which utilizes a graph-based transductive model for using both labeled and unlabeled data to label and select unlabeled example for training set augmentation.
- We propose a novel transductive model based on graph random walk with constrains. This transductive model can deal with labeled training set with noise and provide more reliable predictions for all unlabeled examples, has strong tolerance to noise in the training set.
- We conduct extensive experiments on several UCI benchmark data sets to evaluate its performance with 3 different inductive model and empirically demonstrate that our algorithm can effectively exploit unlabeled data to achieve better generalization performance.

The rest of this paper is organized as follows. Section 2 describes the problem and gives the algorithm in detail. Section 3 presents the experimental results on UCI data sets when various inductive models are utilized. A short conclusion and future work are presented in Section 4.

mds

January 11, 2007

II. The Proposed Algorithm

A. Problem Description and Notation

Let L denote the labeled training set with size $|L|$ and U denote the unlabeled training set with size $|U|$. The goal of our algorithm is to learn a classifier from $L \cup U$ to classify unseen examples. Generally, initial labeled examples are quite few, i.e. $|L| \ll |U|$.

The proposed algorithm learns a inductive model f from labeled and unlabeled data as follows: 1)initialize the model f using labeled set L; 2)use f to predict labels on unlabeled set U; 3)select a subset S from U for which f has the most confident predictions; 4)construct a neighborhood graph G with $L \cup U$ under certain similarity measure; 5)incorporate a transductive model into the self-labeling process: knowing the prior information about labels on $L \cup U$, start a constrained random walk on Graph G to label all the unlabeled examples in U; 6)choose k most confident examples from U for labeled training set augmentation according to the output of random walk; 7)refine f with augmented labeled data. The procedure goes on until there are no unlabeled examples left.

The key distinction between the proposed algorithm and standard self-training is the incorporated transductive model that utilizes both labeled and unlabeled examples to give prediction on unlabeled data. Most graph based semi-supervised methods are transductive, which are nonparametric and can deal with multi-classification problems. We proposed a novel constrained markov random walk for the transduction purpose. The most desirable property of the proposed transdutive model is that it can work well even if training set contains label-noise. Therefore, it is perfectly suitable for the self-training process, as the pseudo-labeled set S may contain some wrongly labeled examples. At this step, it is expected to yield more reliable predictions on unlabeled data than the classifier does with training set subject to label-noise. Next, we will present the details of the proposed transductive graph-based model.

B. Markov Random Walk with Constrains

Markov random walk is regarded as a transductive graph based approach which exploits manifold assumption to label all the unlabeled examples. Typically, it is given an undirected graph $G = (V, E, W)$, where a node $v \in V$ corresponds to an example in $L \cup U$, an edge $e = (a, b) \in V \times V$ indicates that the label of the two vertices a,b should be similar and the weight W_{ab} reflects the strength of this similarity. In this paper, graph is constructed by using the k nearest neighbor criterion. For each example $v \in L \cup U$, Let $\mathcal{C} = \{1, , m\}$ be the set of possible labels. Two row-vectors $\mathbf{Y}_v, \hat{\mathbf{Y}}_v$ are presented. The first vector $\mathbf{Y}_v$ is the input. The lth element of the vector $\mathbf{Y}_v$ encodes the prior knowledge about label l for example v. For instance, a labeled example v with label c has $\mathbf{Y}_{vc}$ set to 1, and the remaining $m-1$ elements of $\mathbf{Y}_v$ set to 0. Unlabeled examples have all their elements set to 0, that is $\mathbf{Y}_{vl} = 0$ for $l = 1...m$. The second vector $\hat{\mathbf{Y}}_v$ is the output of the algorithm, using similar semantics as $\mathbf{Y}_v$. For instance, a high value of $\hat{\mathbf{Y}}_{vl}$ indicates that algorithms believe that the vertex(example) v should have label l.

The constrains of random walks is formalized via three possible actions: inject, continue and abandon(denoted by $inj, cont, abnd$ with pre-defined probabilities $p_v^{inj}, p_v^{cont}, p_v^{abnd}$. Clearly, their sum is unit: $p_v^{inj} + p_v^{cont} + p_v^{abnd} = 1$. To label any example v(either labeled or unlabeled), we initiate a random-walk starting at v facing three options: with probability p_v^{inj} the random-walk stops and return(i.e. $inject$) the pre-defined vector information $\mathbf{Y}_v$. We constrain p_v^{inj} for unlabeled examples. Second, with probability p_v^{abnd} the random-walk abandons the labeling process and returns the all-zeros vector $\mathbf{0}_m$. Third, with probability p_v^{cont} the random-walk continues to one of vs neighbors v' with probability proportional to $W_{v'v}$. Note that by definition $W_{v'v} = 0$ if $(v', v) \notin E$. We summarize the above process with the following set of equations. The transition probabilities are,

$$\Pr[v'|v] = \begin{cases} \dfrac{W_{v'v}}{\sum\limits_{u:(u,v)\in E} W_{uv}}, (v', v) \in E \\ 0 \text{, otherwise} \end{cases} \tag{1}$$

The expected score $\hat{\mathbf{Y}}_v$ for node $v \in V$ is given by,

$$\hat{\mathbf{Y}}_v = p_v^{inj} \times \mathbf{Y}_v + p_v^{cont} \times \sum_{v':(v',v)\in E} \Pr[v'|v]\, \hat{\mathbf{Y}}_{v'} + p_v^{abnd} \times \mathbf{0}_m \tag{2}$$

In this paper, the three probabilities $p_v^{inj}, p_v^{cont}, p_v^{abnd}$ are set using the same heuristics adapted from [10], which are defined by,

$$p_v^{cont} = \frac{c_v}{z_v} \quad ; \quad p_v^{inj} = \frac{d_v}{z_v} \quad ; \quad p_v^{abnd} = 1 - p_v^{cont} - p_v^{inj} \tag{3}$$

$c_v$is monotonically decreasing with the number of neighbors for node v in graph G. Intuitively, the higher the value of c_v, the lower the number of neighbors of vertex v and higher the information they contain about the labeling of v. The other quantity $d_v$is monotonically increasing with the entropy (for labeled vertices). It is noteworthy that abandonment occurs only when the continuation and injection probabilities are low enough. This is most likely to happen at unlabeled nodes with high degree. In effect, high p_v^{abnd} prevents the algorithm from propagating information through high degree nodes.

The final labeling information for all $v \in L \cup U$ can be computed through random walk based on Eq.(2). The algorithm converges when label distribution on each node ceases to change. Note that initial labeled data set L assumes to be noise-free, while the pseudo-labeled dataset S may contain classification noise, hence, certain modification about the transition probabilities needs to be made:

- Since labels on L, which are considered noise-free, should not change during the random walk. For example $v \in L$, the transition probabilities should be fixed as follows: $p_v^{inj} = 1$, $p_v^{cont} = 0$, $p_v^{abnd} = 0$;
- Since examples in S may be wrongly labeled by the classifier, labels on S are allowed to change. For $\forall v \in S$, the transition probabilities should be computed according to Eq.(3);
- For unlabeled example $u \in U - S$, we only constrain $p_u^{inj} = 0$.

Note that the predicted label y_u and labeling confidence CF(u, y_u) of each example $u \in U - S$ can be easily obtained from $\mathbf{Y}_u$:

$$y_u = \arg\max_l \hat{\mathbf{Y}}_{ul} \;,\; l = 1, ...m \tag{4}$$

$$\text{CF}(u, y_u) = \hat{\mathbf{Y}}_{uc} \,, c = y_u \tag{5}$$

In this paper, our strategy is to incorporates such transductive model into the standard self-training's labeling process, concrete procedures of the proposed algorithm is outlined in **Algorithm 1.** It is noteworthy that size of S only has mediate and minor impact on the final performance. For convenience, $|S|$ is empirically set equal to the number of initial labeled examples,i.e. $|L|$, and we also set k equal to $|L|$, The maximum iteration number M is set to 50.

Algorithm 1 Proposed Algorithm

Input:
  -$L \cup U$: training sets
  -$Learner$: learning algorithm for inducing a classifier
  -M: number of iteration
Output:
  -f: the returned classifier
  Construct neighborhood graph $G = (V, E, W)$;
  initialize $p_v^{inj}, p_v^{cont}, p_v^{abnd}$;
  $L' \leftarrow NULL$;
  while $Iter \leq M$ **do**
    Use f to make predictions on U;
    Select S from U { $|S|$ most confident predictions of f};
    Recompute $p_v^{inj}, p_v^{cont}, p_v^{abnd}$ for $v \in S$ using Eq.(3);
    Reset prior labeling knowledge $\mathbf{Y}_v$;
    Output $\hat{\mathbf{Y}}_u$ for all $u \in U - S$ by constrained random walk on Graph G;
    Compute y_u, CF(u, y_u) for all $u \in U$ using Eq.(4)and Eq.(5)
    Choose the k most confident examples from U based on CF(u, y_u);
    Add the chosen pseudo-labeled examples to L';
    $f \leftarrow Learn(L \cup L')$
  end while

TABLE I: Data set summary

Data set	Attribute	Size	Class	Class distribution(%)
australian	14	690	2	55.5/44.5
bupa	6	345	2	42.0/58.0
colic	22	368	2	63.0/37.0
diabetes	8	768	2	65.1/34.9
german	20	1000	2	70.0/30.0
hypothyroid	25	3163	2	4.8/95.2
ionosphere	34	351	2	35.9/64.1
kr-vs-kp	36	3196	2	52.2/47.8
sick	29	3772	2	6.1/93.9
tic-tac-toe	9	958	2	65.3/34.7
vehicle	18	846	4	25.1/25.7/25.7/23.5
wdbc	13	178	3	33.1/39.9/27.0

III. Experiments and Discussion

In this section, we design experiments to verify the efficacy of our algorithm. We mainly focus on the self-training framework, trying to find out how transductive model can improve the semi-supervised inductive model. 12 UCI data sets are used in the experiments[11]. Information on these data sets is shown in Table 1. For each data set, about 25% data are kept as test examples. 10% of the remaining data set is used as the labeled training set L; the rest examples are treated as the unlabeled set U.

The proposed algorithm is compared with standard self-training and SETRED[6]. SETRED is an improved self-training algorithm by incorporating data editing techniques to help identify and remove wrong labels from the training sets during the self-training process. For fair comparison, the termination criteria used by self-training and SETRED are similar to that used by our algorithm.

we used three supervised inductive model as base learners to perform classifier induction, aiming to investigate how each comparing algorithm behaves along with base learners bearing diverse characteristics. Specifically, Naive Bayes,

TABLE II: Classification error rates of 3 compared algorithms on 12 datasets using naive bayes

Dataset	Classification error rates: Naive Bayes						
	initial	proposed algorithm		SETRED		Self-training	
		final	improve/%	final	improve/%	final	improve/%
australian	0.243	0.224	**7.9**	0.234	3.7	0.236	2.9
bupa	0.481	0.442	8.1	0.459	4.6	0.438	**8.9**
colic	0.217	0.207	**4.6**	0.212	2.3	0.221	-1.8
diabetes	0.267	0.257	3.7	0.245	**8.2**	0.264	1.1
german	0.285	0.281	1.4	0.276	**3.2**	0.298	-4.6
hypothyroid	0.024	0.021	**12.5**	0.023	4.2	0.021	**12.5**
ionosphere	0.155	0.129	**16.8**	0.151	2.6	0.166	-7.1
kr-vs-kp	0.142	0.137	3.5	0.143	-0.7	0.128	**9.9**
sick	0.089	0.084	5.6	0.084	5.6	0.079	**11.2**
tic-tac-toe	0.343	0.324	**5.5**	0.328	4.4	0.346	-0.9
vehicle	0.398	0.322	**19.1**	0.367	7.8	0.429	-7.8
wine	0.103	0.096	6.8	0.089	**13.6**	0.116	-12.6
average	–	–	**8.0**	–	4.9	–	1.0

TABLE III: Classification error rates of 3 compared algorithms on 12 datasets using CART

Dataset	Classification error rates: CART						
	initial	proposed algorithm		SETRED		Self-training	
		final	improve/%	final	improve/%	final	improve/%
australian	0.222	0.193	**13.0**	0.199	10.4	0.205	7.7
bupa	0.399	0.368	**7.8**	0.391	2.0	0.388	2.8
colic	0.181	0.163	**10.0**	0.174	3.9	0.168	7.2
diabetes	0.316	0.288	8.8	0.272	**13.9**	0.289	8.5
german	0.351	0.324	**7.6**	0.336	4.3	0.336	4.3
hypothyroid	0.012	0.011	**8.3**	0.016	-33.3	0.021	-75
ionosphere	0.155	0.121	**22.0**	0.144	7.1	0.149	3.9
kr-vs-kp	0.035	0.025	28.6	0.018	**48.6**	0.022	37.1
sick	0.024	0.021	**12.5**	0.026	-8.3	0.023	4.2
tic-tac-toe	0.292	0.258	**11.6**	0.268	8.2	0.277	5.1
vehicle	0.254	0.208	**18.1**	0.223	12.2	0.234	7.9
wine	0.085	0.061	**28.2**	0.072	15.3	0.064	24.7
average	–	–	**14.7**	–	7.0	–	3.2

CART, SVM models are used in the experiments. We use LibSVM[12] implementation for SVM. Note that only Naive Bayes is generative model that can yield probabilistic outputs, CART uses the proportion of dominating class in leaf node as probabilistic output and LibSVM is configured to give probabilistic estimates by using the training option ”-b 1”. For our algorithm and SETRED, we choose a medium number of nearest neighbors, i.e 8 for graph construction, we utilize EUCLIDEAN distance as the similarity measure mainly based on its simplicity and empirical evidences.

Experiments are carried out on each data set for 100 runs under randomly partitioned labeled/unlabeled/test splits. TableII to TableIV present classification errors of these compared algorithms under different inductive models. The ”*initial*” column denotes the average error rates of classification with labeled data only. Columns denoted by ”*final*” and ”*improve*” represent the average error rates and performance improvements of each algorithm respectively.

TableII to TableIV show that proposed algorithm can effectively improve the performance with all the underlying in-

TABLE IV: Classification error rates of 3 compared algorithms on 12 datasets using LibSVM

Dataset	Classification error rates: LibSVM						
	initial	proposed algorithm		SETRED		Self-training	
		final	improve/%	final	improve/%	final	improve/%
australian	0.268	0.229	**14.6**	0.234	12.7	0.241	10.1
bupa	0.382	0.353	**7.6**	0.385	-0.8	0.390	-2.1
colic	0.267	0.176	**34.1**	0.198	25.6	0.235	12.1
diabetes	0.421	0.345	**18.0**	0.428	-1.6	0.448	-6.4
german	0.214	0.188	12.1	0.202	5.6	0.179	**16.4**
hypothyroid	0.029	0.024	17.2	0.021	**27.6**	0.026	10.3
ionosphere	0.183	0.164	10.4	0.142	**22.4**	0.181	1.1
kr-vs-kp	0.034	0.027	20.6	0.031	8.8	0.023	**32.4**
sick	0.036	0.034	4.0	0.031	**13.9**	0.033	8.3
tic-tac-toe	0.071	0.036	49.0	0.056	21.1	0.030	**57.7**
vehicle	0.398	0.316	**20.6**	0.367	7.8	0.429	-7.8
wine	0.103	0.048	**53.4**	0.066	36	0.116	-12.6
average	–	–	**21.8**	–	14.9	–	10.0

ductive model. he two-tailed paired t-test under the significant level of 95% shows that all the improvements of performance are significant. The biggest improvements achieved by these three self-training style algorithms have been boldfaced. In fact, if the improvements are averaged across all the data sets, base learners, it can be found that the average improvement of our algorithm is about 14.8%. It is impressive that with all the employed inductive models our algorithm has achieved the biggest improvement among the other 2compared algorithms. Moreover, if the algorithms are compared through counting the number of winning data sets, i.e. the number of data sets on which an algorithm has achieved the biggest improvement among the compared algorithms, our algorithm is almost always the winner. In detail, when Naive Bayes are used, it has 6 winning data sets while self-training has 4 and SETRED has 3; when CARTs are used, our algorithm and SETRED has 10 and 3 winning data sets respectively, while self-training do not have winning data sets; when SVMs are used, our algorithm, SETRED and self-training has 6, 3 and 3 winning data sets respectively.

In particular, comparing to SETRED which utilizes a specific data editing technique to actively identify wrongly-labeled examples in the enlarged training set, the proposed algorithm has achieved better results with no effort for cleaning the training set. This evidence supports our arguments that the incorporated transductive model is robust to noises introduced by the self-labeled process, thus it can achieve stable performance. Moreover, empirical results of on the 2 multi-class datasets(*vehicle,wine*) suggest that our algorithm is superior to self-training and SETRED when dealing with multi-class classification problems. This is mainly due to the fact that it can naturally handle multi-class classification by exploiting manifold assumption to yield confident predictions for training set augmentation.

IV. Conclusions

This paper shows the benefits of incorporating transdcutive models into semi-supervised bootstrap inductive models, such as self-training. This strategy utilizes both labeled and unlabeled data to yield more reliable predictions for unlabeled examples. We propose a robust self-training style algorithm which exploits manifold assumption to facilitate the self-training process. We adopt a transductive model based on graph random walks to prevent performance degradation due to classification noise accumulation. Empirical results on 12 UCI datasets show that proposed algorithm can effectively exploit unlabeled data to enhance performance.

Graph construction is vital to our algorithm. In this paper, we only use the common EUCLIDEAN distance as the distance measure, there is no guarantee that this is the optimal choice. Generally, the problem of choosing the best distance measure for a specific learning task is very difficult, and some efforts have been made towards tackling this problem under the name of distance metric learning. How to identify or learn the optimal distance measure and how does it affect the performance are worth further investigation.

Acknowledgment

The research is partially supported by National Natural Science Foundation of China under grant number 61173087,61073128. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

[1] D. Steinberg and P. Colla, "Cart: classification and regression trees," *The Top Ten Algorithms in Data Mining*, pp. 179–201, 2009.

[2] X. Zhu, "Semi-supervised learning literature survey," *Computer Sciences TR 1530, Univ. of Wisconsin*, 2008.

[3] D. Yarowsky, "Unsupervised word sense disambiguation rivaling supervised methods," in *Proceedings of the 33rd annual meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 1995, pp. 189–196.

[4] S. Abney, "Understanding the yarowsky algorithm," *Computational Linguistics*, vol. 30, no. 3, pp. 365–395, 2004.

[5] Y. Wang, X. Xu, H. Zhao, and Z. Hua, "Semi-supervised learning based on nearest neighbor rule and cut edges," *Knowledge-Based Systems*, vol. 23, no. 6, pp. 547–554, 2010.

[6] M. Li and Z.-H. Zhou, "Setred: Self-training with editing," in *Advances in Knowledge Discovery and Data Mining*. Springer, 2005, pp. 611–621.

[7] M.-F. Balcan, S. Hanneke, and J. W. Vaughan, "The true sample complexity of active learning," *Machine learning*, vol. 80, no. 2-3, pp. 111–139, 2010.

[8] T. Joachims *et al.*, "Transductive learning via spectral graph partitioning," in *ICML*, vol. 3, 2003, pp. 290–297.

[9] A. Subramanya and J. Bilmes, "Semi-supervised learning with measure propagation," *The Journal of Machine Learning Research*, vol. 12, pp. 3311–3370, 2011.

[10] S. Baluja, R. Seth, D. Sivakumar, Y. Jing, J. Yagnik, S. Kumar, D. Ravichandran, and M. Aly, "Video suggestion and discovery for youtube: taking random walks through the view graph," in *Proceedings of the 17th international conference on World Wide Web*. ACM, 2008, pp. 895–904.

[11] A. Frank and A. Asuncion, "Uci machine learning repository," 2010.

[12] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.

Migration Dynamics in Artificial Agent Societies

Harjot Kaur
Dept of Comp. Sci. and Engg.
GNDU Regional Campus, Gurdaspur
Punjab, INDIA
Email: harjotkaursohal@rediffmail.com

Karanjeet Singh Kahlon
Dept of Comp. Sci. and Engg.
Guru Nanak Dev University, Amritsar
Punjab, INDIA
Email: karanvkahlon@yahoo.com

Rajinder Singh Virk
Dept of Comp. Sci. and Engg.
Guru Nanak Dev University, Amritsar
Punjab, INDIA
Email: tovirk@yahoo.com

***Abstract*—An Artificial Agent Society can be defined as a collection of agents interacting with each other for some purpose and/or inhabiting a specific locality, possibly in accordance to some common norms/rules. These societies are analogous to human and ecological societies, and are an expanding and emerging field in research about social systems. Social networks, electronic markets and disaster management organizations can be viewed as such artificial (open) agent societies and can be best understood as computational societies. Members of such artificial agent societies are heterogeneous intelligent software agents which are operating locally and cooperating and coordinating with each other in order to achieve goals of an agent society. These artificial agent societies have some kind of dynamics existing in them in terms of dynamics of Agent Migration, Role-Assignment, Norm-Emergence, Security and Agent-Interaction. In this paper, we have described the dynamics of agent migration process, starting from the various types of agent migration, causes or reasons for agent migration, consequences of agent migration, and an agent migration framework to model the its behavior for migration of agents between societies.**

I. Introduction

An *Artificial Agent Society* [8] can be defined as a collection of agents interacting with each other for some purpose and/or inhabiting a specific locality, possibly in accordance to some common norms/rules.These societies are analogous to human and ecological societies, and are an expanding and emerging field in research about social systems. Social networks, electronic markets and disaster management organizations can be viewed as such artificial (open) agent societies and can be best understood as computational societies. The members of such artificial agent societies are heterogeneous intelligent software agents, which are operating locally and cooperating and coordinating with each other in order to achieve goals of an agent society. Artificial Agent Societies can also be viewed as *normative systems*, as in them, the member agents residing have to obey certain rules/norms which are created and abolished from time-to-time in a society and this process is dynamic in nature. Also, the agents residing in the society, while obeying the norms of the society and in order to achieve their individual as well as societal goals move in/between them, and this movement can be termed as *agent migration*.

A. What is Migration?

A **migration system** can be defined as a set of places linked by flows and counter-flows of people, goods, services and information [12], which tends to facilitate further exchange, including migration between the places. And, according to demographers, every act of migration involves an origin, destination, and an intervening set of obstacles [15]. The dictionary meaning of word **Migration** according to Oxford on-line dictionary is

> Movement of people to a new area or country in order to find work or better living conditions.

Considering its *broader demographic perspective*, the term *migration* [19] can also be defined as temporary or permanent move of individuals or groups of people from one geographic location to another for various reasons ranging from better employment possibilities to persecution. Talking in terms of agents (intelligent/software) in artificial agent societies as mentioned above, *agent migration* can be defined as movement of agents in and between societies to new locations for various reasons ranging from societal, economic, social or personal. Although, this movement is selective, but it forms an integral part of the broader process known as *development* of societies, which itself is a linear, universal process consisting of successive stages [12]. The study of migration is not only confined to social sciences, but it has got with a vast and significant importance in the areas of political sciences, economics, anthropology, biology, psychology and artificial intelligence (computer sciences).

In general, migration is considered as a complex dynamic process and a spatial phenomenon with multi-faceted nature and has its own sophisticated theory and dynamics associated with it. The concept of agent migration has been derived from human or animal migration, because like human or animal migration, it has also got associated with it reasons, circumstances, patterns as well as consequences [19] of/for migration. Therefore, various similarities [3] exist between agent, human and animal migration, which are described below.

1) Similarities between agent, human and animal migration: Firstly, both the agent and human/animal migration involve movement in order to migrate, be it in the form of migration in or between societies, and this movement involved in both the cases is linear in nature. Secondly, in the all these type of migrations, a special preparation is initially required for migration, which demands a special allocation of energy from the society. Thirdly, migrating species, whether they are agents, humans or animals, have to maintain a strong commitment to their migration mission or challenge, which is going to keep them undisturbed from all the side temptations. Fourthly, the journey of migration, whether its agent, human or animal migration, has to be continued at all costs.

In the case of human migration, migration can be categorized either as *internal (intra-continental or intraregional)* or *external (inter-continental)* migration [14] [15], talking in same terms, **Agent Migration** can be categorized [17] either as *internal Migration*, (i.e., Migration of agents between two sites in the same society) or as *external Migration*, (i.e., Migration of agents between two different societies). The human migration is normally referred as *immigration*, which is common term used in *demography*, whereas agent migration should be referred by term, *migration* only. In agent migration, the agent who migrates into a new society will be hence termed as *immigrant agent*, and while it is leaving its society of origin to migrate into another, will be termed as *em-migrant agent*.

In addition, to various similarities between agent, human and animal migration, there exist a number of dissimilarities also, which are described below.

2) Dissimilarities between agent, human and animal migration: As explained above, agent and human migration are different with respect to few points also, for instance, agent migration is organized and collective in nature whereas human immigration is disorganized and sporadic. In addition to this, human immigration is downright enigmatic, and this aspect is missing in agent migration in artificial agent societies.

B. Dynamics in Artificial Agent Societies

Artificial agent societies like any other working system have some kind of *dynamics* existing in them, which are in terms of dynamics of *Agent Migration*, *Role-Assignment*, *Norm-Emergence*, *Security* and *Agent-Interaction* [23]. These dynamics are basically changes that occur in the behavioral and emergent properties of the societies due to sensing of a stimulus from an environment as well as due to an external agent, who wants to migrate into the society. Also, these all types of dynamics are in one way or another related to each other as all of them are either directly or indirectly affected by one another.

1) Migration Dynamics in Artificial Agent Societies: The member agents residing in these societies keep on moving between these societies from time-to-time in the same manner as human-beings move between human-societies [20]. And, this movement of agents between artificial agent societies is called Agent Migration. And, all the *internal (related to agent)* and *external (societal)* issues related to agent migration are called agent migration dynamics. In this paper, we have described agent dynamics from their each and every aspect as they play a very pivotal role in the shaping of societies throughout their life time.

2) Why Migration Dynamics: An agent migration is considered as an integral part of development process [12] involved with the growth of artificial agent societies. Therefore, agent migration is one of very essential part of society dynamics, which has to be studied and worked in thoroughly and very less work has been done in this direction. Also, all the other types of society dynamics, (i.e., dynamics of role-assignment, agent interaction, norm emergence, trust and security) [23] are very closely related to it, as all of them are either directly or indirectly affected by migration of agents in/between societies.

Hence, we have focused our research on agent migration dynamics [27] in artificial agent societies in this paper. Our paper focuses on migration dynamics in artificial agent societies because of agent migration process, specifically external migration process, and we have tried to discuss almost everything which can be related to external migration dynamics, i.e., its types, reasons, framework, consequences or problems arising because of external agent migration in this article.

This paper is organized has follows. In section 2, we have discussed background work already performed in the area of agent migration. In section 3, we have covered various aspects of migration dynamics in agent societies. In section 4, we focus on the types of agent migration between agent societies. In section 5, reasons for agent migration between societies are stated. In section 6, consequences arising out of this agent migration are presented and in section 7, we have described an agent migration framework to facilitate migration of agents between agent societies. Then, conclusions and future work are covered in sections 8 and 9 respectively.

II. Background and Related Work

In this section, we have tried to summarize various works, which have been performed and are related to agent migration in artificial agent societies. Also, we have illustrated various possible extensions, which can be performed related to them.

The categorization of agent migration has also been performed by Costa et al [17], as *Internal Migration* and *External Migration* where external migration occurs if agent moves between societies and internal migration occurs if agent moves within a society from one site(group) to another. Also, various consequences related to agent migration are discussed by Glaser and Morignot in [9] and Dignum et al in [6], and they are in terms of *reorganization* of societies that takes place, i.e., society in which agent enters, has to modify (means modification of structure of its organization) itself in order to accommodate the agent in society's already existing roles, which are distributed amongst its member agents. Also, the agent which has joined the society needs to adapt itself so as to internalize in the organization of the society, it now lives in. The type of reorganization demonstrated by Glaser and Morginot is *static reorganization* and by Dignum et al is *dynamic reorganization* as in the case of latter, modifications in structure and behavior of an artificial agent society after addition, removal or substitution of an agent are done to the society while it is executing (i.e., without bringing down the system). This particular type of reorganization is in the form of *dynamic adaptation*.

The works presented by above three authors are one of the basic and initial works done in the area of agent migration dynamics, but all these can be expanded also, by adding more aspects and factors to the agent migration dynamics. For instance, the only types of migration presented by Costa et al [17] are internal and external migration, which are categorized taking into consideration only the factor of geographical distance spanned by an agent while moving from one place to another. But, there are many other factors, which can be considered (i.e., circumstances and adaptability) while agent is to move. Similarly, the only consequence of migration is listed by Dignum et al [6] and Glaser and Morginot [9], as reorganization (static and dynamic) of society, which is one of the societal consequences, where as there are many other societal and individual consequences, which can be explored upon, while working with the process of agent migration.

The basic motivation behind the reorganization of society is either increase in the *agent's utility* or the *society's utility*. Utility [9] of society means what society has gained out of reorganization after the new agent has joined the society and in order to evaluate the utility of convention of society, cost of reorganization and utility of existing structure once without the newly migrated agent and then with newly migrated agent is evaluated in the form of a utility function, where convention means distribution of roles between agents of the society. And, if the utility of the society increases with the above agent migration, reorganization and hence agent migration is considered beneficial for the society.

The utility of an agent depends on the roles it desires, the roles it has committed to and the confidence with which it is playing its roles. Normally, an agent chooses to integrate into that society, which increases its utility and in that case agent is beneficiary. And on the other hand, a society chooses to integrate with an agent, that increases the its utility. Initially, an agent joining an agent society will be consulting the institutional layer [5] of the society to commit itself to certain role, which it has to play in a society and in order to see, whether it satisfies the norms and rules enforced in the society by this layer.

The reorganization, which is an after effect of agent migration, (i.e., integration of a new agent with a society), demonstrated by Dignum et al [6] is dynamic in nature. They have classified it into types, i.e., *Behavioral* and *Structural*. *Behavioral Change* occurs when organizational structure remains the same but the agents enacting roles [4] [24], decide to use different protocols for the same abstract interaction described in the structure, in case an agent leaves the society and a new agent joins it. Therefore, it is only the interaction pattern, that has to be modified, but on the other hand in the case of *Structural Change*, a decision is made concerning the modifications of one of the structural elements, i.e., societies adapt to environmental changes due to addition, deletion or modification of its structural elements (i.e. agents, roles, norms, dependencies, ontologies and communication primitives). Therefore, behavioral changes temporary changes in an organization whereas structural changes can lead to permanent modification in the structure of an organization [18].

Two different *issues* related to agent integration to an existing agent society are described by Eijk et al in [7] and they are based upon open-ended nature of agent societies that allows for the dynamic integration of new agents into an existing open system. These issues are distinguished as **Agent Introduction** and **Agent Creation**. *Agent Introduction* is addition of a new agent which is existing outside the society into a society and *Agent Creation* is similar to object creation, i.e., a newly integrated agent constitutes a previously non-existent entity and is created in a society for agents with limited resources who may face an overloading of tasks to be performed by them. Although, this paper significantly contributes to the process of agent migration, by dynamically integrating new agents into an already existing society, by using an abstract framework in concurrent object oriented language called POOL and its various communication constructs, but still the abstract framework present in this paper for agent introduction and agent creation in a society can be expanded by considering various push-pull factors and the behaviors of the agents which are migrating in/between them, and that also, while the society is still in running state, i.e., without bringing it down.

TABLE I. SUMMARY OF RELATED WORK

Author	Work Direction	Key Points	Possible Extensions
Costa et al.	Migration Classification	Internal and External Migration	Various other categories of migration.
Glaser & Morginot	Migration Consequences	Reorganization (Static)	Various other societal and individual consequences.
Dignum et al	Migration Consequences	Reorganization (Dynamic)	Various other societal and individual consequences.
Eijk et al	Migration Issues	Agent Introduction and Creation	A complete migration framework for agent introduction.
Hafizoglu and Sen	Migration Patterns	Conservative, Moderate and Eager Migration	Various other patterns of migration.

When agents migrate between societies, some patterns emerge for migration and they are illustrated by Hafizoglu and Sen in [13] as *Conservative*, *Moderate* and *Eager* migration. These patterns emerge from opinions or choices of agents to migrate. For experimenting between choices and patterns they have used two-dimensional toroidal grid in which simulation proceeds in discrete time steps with agents having two types of opinions, i.e., binary or continuous. In addition to this, many other types of patterns emerging for agent migration process can be studied, like chain, return or step migration of agents amongst societies/groups.

All the above mentioned related works are summarized in table I, which explains various directions followed by various authors in the area of agent migration dynamics, various findings of these research directions as well as how their work can be extended in order to incorporate new aspects in migration dynamics.

Although, migration of agents between societies is discussed in all the above papers, which we have surveyed in this section from different perspectives, but none of them elaborates on the complete migration framework or protocols, which are required and can be used for migration of agents amongst agent societies, i.e., for organizational migration. Our paper focuses on the elaboration of a complete migration framework, which can be used by agents to migrate from one society to another. In addition to that, various dynamics (as mentioned in table I) and their extended forms, in the process of migration will be discussed one by one, in the coming sections.

III. MIGRATION DYNAMICS IN ARTIFICIAL AGENT SOCIETIES

Agent Migration occurs when an agent moves from one site to another site between a society or an agent moves from one society to another one, that also, physically as well as logically. Agent Migration or the process of entrance of new agent into the society has many issues related with respect to dynamics and structure of an open-agent society (as only in the case of open-agent societies, agents can leave or enter any time). In order to understand properly, all the aspects and behavior of agent migration subsystem [23], which can be viewed as an interdependent dynamic subsystem with its own dynamics, but is interlinked with other subsystems existing in various societies, feedback and adjustments coming from the migration process itself, a proper detailed study of agent migration dynamics is required and that is the main objective of this paper.

The agent migration can be physical as well organizational, because there is a possibility of agents migrating between platforms and that also may be on different computers or same as well, geographically. Our concern in this paper is, only organizational migration, as in the former case, everything is adjusted by network infrastructure for transportation of mobile agents (which are pieces of software), from one site to another. As, the agents in the case of physical migration (agents are mobile in nature), they find standard environments at every site they visit in addition to standard script interpreters for the execution of code and standard communication constructs. Therefore, they face no such problems during and after the physical migration, which are present in the case of organizational migration. Therefore, our main concentration in this paper is only on *organizational* migration and its dynamics.

Basically, dynamics [25] existing in every system can be classified as either *macro dynamics* (i.e., between various societies or inter-societal dynamics) or *micro dynamics* (i.e., inside a single society or intra-societal dynamics). The aspect of internal migration will be considered under micro dynamics and external migration under macro dynamics in agent societies. External Migration is only possible in open agent societies, i.e., society needs to be open internally as well as externally where as in the case of internal Migration society, it need not be a externally open society. This agent mobility actually facilitates efficient collaboration of an agent with other agents at intra-societal (micro dynamics) or inter-societal (macro-dynamics) level. As, we are concerned only with external migration, therefore, the dynamics studied by us will also be of the type of macro-dynamics only. Related to macro or external agent migration dynamics is the *reasons*, *types*, *consequences* and *patterns* adopted for migration. Our paper is dedicated to the illustration of these all dynamics related to agent migration in artificial agent societies, and all of them are covered one by one in the coming sections.

Both internal and external migrations are formally treated as the same kind of processes, since the type of procedures involved are the same. In the case of internal migration, migrating agents will be specialized, non-autonomous agents and in external migration, migrating agents will be fully autonomous agents with full capabilities of their own. The set of external migration dynamics [27] need a multi-dimensional theory and a formulation of proper migration framework. Hence, our main concentration in this paper will be on framework for external migration and study of migration dynamics, because this study have more relevance, when they are considered with respect to external migration. Therefore, all the dynamics related to agent migration, starting from the types, reasons and consequences have been discussed by us in the sections to come. Our next subsection is related to various types of migration, (i.e., migration categorization) in artificial agent societies.

IV. Types of Agent Migration

Agent migration, as explained above, is similar to human or animal migration because reasons, circumstances, patterns and finally consequences, associated with agent migration are almost same as human/animal migration. Also, Agent migration as inspired from human migration [21] [14] can be classified accordingly into various categories depending upon *three* factors, i.e., *geographical distance spanned, circumstances and adaptability*. These three factors can be stated as follows, Firstly, from geographical distance spanned, we mean that the actual distance which is covered by agent while migrating from one place to another. Secondly, from circumstances, we mean all the political, economic and environmental conditions which force the agent to move from one place to another in a same society or another society. Lastly, adaptability means ability of agent to cope up with new surroundings and new place, when agent moves there [22]. The society/site of origin from which, agent starts its migration can be termed as **migration source** or **sending society** and the destination society/site to which agent migrates is called **migration sink** or **receiving society**.

The first set of migration categories, which are classified according to *geographical distance* factor are *internal* and *external* migration[1]. Where,

- **Internal Migration** can be defined as migration of agents between two sites in the same society, and
- **External Migration** can be defined as migration of agents between two different societies.

The internal migration in agent societies is similar to intraregional migration in human societies and external migrations in agent societies are analogous to intra-continental and intercontinental (i.e., interregional) migrations in human societies. Therefore, consequences and conflicts arising in agent societies are almost same as consequences and conflicts arising in human societies because of similarities between external and internal migrations in agent societies with inter and intra-regional migrations in human societies.

The second set of migration categories, which are classified according to *circumstances* which lead to migration, they are *forced or involuntary*, *voluntary* or *imposed* migrations. Where,

- **Forced or Involuntary Migration** can be defined as migration in which agent is forced to migrate from source of migration to another society or site due to certain unfavourable circumstances at the migration source.
- **Voluntary Migration** can be defined as migration in which agent voluntarily migrates to another society or migration sink, due to some favourable circumstances there.
- **Imposed Migration** can be defined as migration in which agent is not forced to migrate to another society but due to persistent unfavourable circumstances at the migration source, agent leaves source.

In the case of forced or involuntary migration, agents are left with no choice, they have to migrate, i.e., leave the society. But, in the case of voluntary migration, which can also be termed as *choice migration*, agents by their own will and wish choose to migrate from one place to another. In the third case, i.e., in imposed migration, which can also be termed as *reluctant or impelled* migration, agents are not forced to migrate but it is due to persistent unfavourable circumstances, they decide to migrate from migration source. The concept of agent migration is also related to various reasons which are the root cause behind it and hence they become an essential part of agent migration dynamics, therefore, we have dedicated next section of our paper to various reasons leading to agent migration.

V. Reasons for Agent Migration

Agents are also assumed to be residing in communities like human beings that are connected in some known topological structure [13]. There are always certain *reasons* which are responsible for agent migration, for instance, there are always some social functions and resources needed by an agent which are present in some another society [16], different than its source society for performing that particular job. Agents can migrate because of various reasons, i.e., *social/societal, economical or personal/individual*. These reasons are further based on *push and pull factors* leading to migration of agents amongst agent societies [20] [15]. *Push factors are basically positive attributes perceived by agents, existing at the new location*, i.e., new society or new site to which agent is planning to migrate. Whereas, *pull factors are negative home, i.e., local site or society conditions that impel the agent's decision to migrate to a new society or site*. These push and pull factors are analogous to push and pull factors that result in human migration and can be categorized according to reasons for agents migration as *societal* push-pull factors, *economic* push-pull factors and *individual* push-pull factors.

The *societal* push-pull factors are combination of negative conditions arising in the atmosphere of source society and positive conditions arising in the destination society, as a consequence of which agents migrate between these source and destination agent societies. They are unequal/unfair role-assignment in the source society and comparatively fair/equal role-assignment in the destination society, ineffective security policies in the source society and effective security policies in the destination society. The unequal role-assignment, in a particular society is one of root causes of net overall migration.

Also, when in a particular society some particular role is required to played by an agent and the member agents of that society are incompetent to play that role. And, if that society comes to know, that, there is some particular agent in some other society, and it is fully competent to play the required role. In this case, the destination society, which requires an agent, can also request or pull an agent from its source society to play that particular role in the destination society. This pull factor can be considered as societal pull factor from the side of destination society.

The *economic* push-pull factors are combination of economic benefits received in the form of rewards for playing some specific role, which is existing in some destination society to which agent wants to migrate, and they pull an agent from its residential source society to migrate. Also, rewards obtained by agent for playing the same role in its source society will be lesser in amount as compared to rewards for the same role, which it can obtain in destination society. Therefore, this reward differential will be serving as push factor, which will be pushing an agent to migrate to some other society, where its economic conditions can be better.

The *individual* push-pull factors are combination of individual or personal benefits, which an agent can receive in the destination society to which it migrates. They can be in terms of increased individual agent's utility, autonomy, or role-playing opportunities in destination society as compared to the source society in which it currently resides. This factor is basically related to the favorable conditions, which an agent requires in order to function well in society and in search of these conditions, an agent migrates to that destination society, where these conditions are met.

These push and pull factors, i.e., societal, economic and individual, can also be categorized as *material* and *non-material* incentives related to agent migration, because some of the sub-factors from societal, economic and individual factors [12], result in material benefits/incentives, i.e., can be in terms of economic benefits, whereas others can result in non-material benefits/incentives, which is basically in terms of the chances of self-actualization, for the em-migrant/immigrant agents. After the migration is actualized by an agent or set of agents, *(we are considering here agent or set of agents only because migration is a selective process)*, and only those agents who are interested in undertaking this process, participate. The extent of migration amongst agent societies depends upon all these migration push-pull factors listed above.

VI. Consequences of Agent Migration

There are several *issues*, *problems* and *consequences* associated with the outcome of agent migration and broadly they can be categorized as issues related to the societies and to the migrating agents. Hence, consequences related to agent migration can be classified as *societal consequences* or *personal consequences*. Here, *Societal consequences* occurring because of agent migration, are *maintenance of functional integrity of a society*, i.e., placing the new immigrant agent in society's already existing social structure [11] whereas personal consequences occurring at the side of *immigrant agent*, are problems of *language and interaction protocols* while conversing with the member agents of the destination society, i.e., the society it has entered into, so that it is considered as a well-behaved agent and problems of *knowledge and performance*, which actually define the ways the agent will behave in order to properly perform the functions (to play the roles) that destination society wants it to perform.

Another set of problems related to an immigrant agent while acting as an em-migrant agent is of helping the society it is leaving, i.e., its source society, to preserve its functional integrity in its absence. An immigrant agent will only be integrating (entering) with a new destination society only if it is sure that it is capable of dealing with the problems that it will face when doing that, i.e., it is prepared to solve problems produced by its entrance or departure. This set of problems can be also considered to fall under personal or individual consequences of agent migration. This section of our paper will be describing all the societal as well individual consequences, related to agent migration.

A. Societal Consequences

Various societal consequences related to agent migration process are:

1) Reorganization: The societies of origin as well as destination always have to *reorganize* structurally as well as behaviorally, when the process of agent migration takes between them. These consequences can also be termed as social consequences arising out of agent migration. These consequences result in terms of constructs of *dynamic role-allocation* to immigrant agent in the destination society and dynamic *role- deallocation* and then *reallocation* in the source society in the place of em-migrant agent. This reorganization can be performed using dynamic role-allocation constructs available in the *role- assignment* module or subsystem for source and destination artificial agent societies.

2) Conflicts: As a result of agent migration process, conflicts will arise in the destination society, to which immigrant agent has entered, because of its integration. The conflicts arising will be *role-conflicts* and *norm-conflicts*. Role-conflicts arise, if an immigrant agent was allocated some role as as em-migrant agent in its society of origin. And, after migration to the destination society, as an immigrant agent, it still wants to play the same role in destination society and in the later, it is not possible. Norm-Conflicts arise, when em-migrant agent moves to some destination society, and there as an immigrant agent, does not wants to obey the norms of the destination society, and still wants to remain associated with the norms of its society of origin. These conflicts are only tolerated for immigrant agent, when its integration with the destination society increases the society's overall utility.

3) Norm Emergence: As a result of migration of agents between societies, the societies also come closer to each other because, when agents following an altogether different set of norms move between them, then *norm emergence* modules of both these societies will execute their respective *norm recognizer* sub-modules to recognize a new set of candidate norms, which later will be functioning as proper or actual norms in the source and destination societies.

B. Individual Consequences

Various individual consequences related to agent migration process are:

1) Interaction Language and Protocol Conflicts : These conflicts arise, when an immigrant agent enters the destination society, and the agent communication language and protocols used in the destination society for agent communication are different from what it was using in its society of origin, and agent is unable to participate in conversations with the member agents of the destination society. In order to resolve such conflicts, language and protocol adaptation at the level of immigrant agent is required in order to adapt already existing communication infrastructure in the destination society. Related to this, very significant amount of work has been performed by Bordini in his doctoral thesis [1] [2], which is dedicated to cultural adaptation being performed at the level of immigrant agents in destination society.

2) Knowledge and Performance Problems: These problems arise when immigrant agent in the destination society is unable to perform the role allocated to it in destination society in a proper manner, because of improper knowledge present with it to perform that role. And, if the problem of performance persists for the immigrant agent, the destination society may can force an immigrant agent to move back to its society of origin.

The outcome of agent's integration with a society is usually an establishment of a new convention between the members of a society and the newly entered agent, i.e., redistribution of roles, performed by role-assignment module which is assigning or distributing roles amongst various member agents in the society. The end result of migration process is a set of relatively stable exchanges, yielding an identifiable geographical structure [13] that persists across space and time. The details of the behavior of migration process are presented by us in the next section.

VII. An Agent Migration Framework

Agent Migration is very vital for the development of artificial agent societies, because it is basically agent migration, which will be making the set of existing societies dynamic and in-turn facilitating the movement of agents in/between societies. *An agent migration framework* is vital with respect to facilitation of process of migration of agents amongst agent societies. And, it can be considered to be comprised of a set of societies linked by flows and counter-flows of agents between societies, which tends to facilitate further exchange, including migration between the societies. An agent migration framework in order to facilitate migration in artificial agent societies, requires a migration model [26] which can be used to model the behavior of agents, when they migrate in/between agent societies.

The presentation of migration dynamics [25], is actually involved with modeling migration processes over time, which we have assumed to be comprising of discrete time intervals. An agent migration framework comprises of modeling an agent *migration model*, which in turn uses a *migration function* to facilitate migration of agents from source society or society of origin of agents to destination society. A *migration model* attempts to ascertain the relative importance of various determinants (push-pull factors) leading to migration and also addresses various personal/individual as well as societal consequences resulting after migration. Related to migration model is the term *migration interval*, which is the period of time over which a migratory move is taken by potential em-migrant agents from the source or sending society. This migration interval is also measured in discrete time steps.

The migration model, which we are going to present is an empirical migration model based on combination of gravity model [10] given by Greenwood and social- psychological model of immigration given by Andrew and Zara in [3], and we have given it a name **W5- SGMIM** for *W5 social-gravity migration model* for inter-societal or external migration between agent societies. The characters **W5** which are mentioned in the name of the SGMIM model, are basically **5 W's** *(Who, Whey, Where, What and When)* related to the process of migration and they stand for the following information related to migration model for modeling the behavior of migration process amongst agent societies.

1) *W for Who?* - *Who* amongst the set of member agents residing in the society wants to migrate? This *W* is related to the selection of em-migrant agents in the source society, who are to move to destination society.
2) *W for Why?* - *Why* a particular agent or set of agents want to migrate? This *W* is related to various factors or determinants in source and destination societies, that are leading to agent migration from the source society to destination society.
3) *W for Where?* - *Where* do agent(s) want or plan to migrate? This *W* is related to selection of the destination society to which em-migrant agents in source society, want or plan to migrate and act as immigrant agents there.
4) *W for What?* - *What* will be the consequences of this migration? This *W* is related to various societal and personal consequences, which will be resulting out of this migration in both source and destination societies.
5) *W for When?* - *When* do agent(s) want to migrate? This *W* is related to the description of beginning of *migration interval* of em-migrant agents from the source society to move to destination society.

This model is called a social-gravity model because the migration function used in the migration process is based on gravity model [10] of human migration, which is based on modified version of *Newton's Law of Gravitation*, i.e., the attractive force between two bodies is directly related to their size and inversely related to the distance between them. And, the keyword *social* is added, because, it is based on input of a set of social-psychological factors (push-pull factors), which are the basic determinants of the migration process.

A. A Migration Function

Definition VII.1. (The Agents World)

The Agents World is a set of societies existing in the world of agents, and is defined as $\mathscr{S} = \{\mathscr{S}_1, \mathscr{S}_2, \ldots, \mathscr{S}_n\}$ with a finite number of **n** societies.

Definition VII.2. (Population)

A Population is a set of member agents of any society $\mathscr{S}_i$, and is defined $\mathscr{P}_i = \{\mathscr{A}_{i1}, \mathscr{A}_{i2}, \ldots, \mathscr{P}_{im}\}$, with **m** being finite number of residing member agents in a society $\mathscr{S}_i$.

Definition VII.3. (Roles)

Roles are place- holders assigned to various member agents of the society according to their capability and role-assignment/allocation protocols of the society. They are defined using a finite set $\mathscr{R} = \{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_k\}$, where **k**, is the number of roles in the role set $\mathscr{R}$.

Definition VII.4. (Agent)

An Agent $\mathbf{a} \in \mathscr{P}_i$ is a tuple $\mathbf{a} = \langle \mathbf{r_{ia}}, \mathbf{u_a}, \mathscr{B}_\mathbf{a} \rangle$, where $\mathbf{r_{ia}}$ is the role allocated to an agent $\mathbf{a}$, $\mathbf{u_a}$ is utility function of an agent $\mathbf{a}$, and $\mathscr{B}_\mathbf{a}$ is behavior of an agent $\mathbf{a}$.

The behavior $\mathscr{B}_\mathbf{a}$ of an agent $\mathbf{a} \in \mathscr{P}_i$, where $\mathscr{P}_i$ is the population of society $\mathscr{S}_i$, describes the behavior of an agent from external point of view. This behavior function is a combination of many external (push-pull) factors, which are existing in the source and destination societies. This behavior

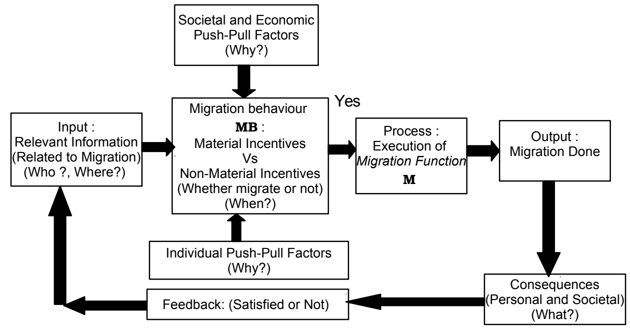

Fig. 1. The Social- Gravity Migration Model

function, acts as very important decision making factor in the migration process in the form of migration behavior function, in our migration model.

The *migration function* leading to the process of migration can be described as:

Definition VII.5. (A Migration Function)

A Migration function $\mathscr{M}_{i,j}$ for agent $\mathbf{a} \in \mathscr{P}_i$ in source society $\mathscr{S}_i$ for migrating to destination society $\mathscr{S}_j$ at any time $\mathbf{t} \in \mathscr{T}$, where $\mathscr{P}_i$ and $\mathscr{P}_j$ are populations of source and destination society respectively, and $\mathscr{D}_{i,j}$, is distance between them, leading to the process of agent migration amongst them can be defined as

$$\mathscr{M}_{i,j} : (\mathbf{a} \times \mathscr{T}) \rightarrow \mathscr{S}_j. \quad (1)$$

Here, $\mathscr{T}$ is a time set comprised of discrete time steps, $\{\mathbf{t}, \mathbf{t+1}, \ldots, \mathbf{t+n}\}$.

Also, The Migration Function $\mathscr{M}_{i,j}$ can de written as:

$$\mathscr{M}_{i,j} = \frac{G\mathscr{P}_i\mathscr{P}_j}{\mathscr{D}_{i,j}^2} \quad (2)$$

Where G is constant. and $\mathscr{D}_{i,j}$ is distance between the two societies $\mathscr{S}_i$ and $\mathscr{S}_j$ and it is measured in terms of number of hops agent has to take to move from one society $\mathscr{S}_i$ to $\mathscr{S}_j$.

The above relationship states that *migration function* is directly related to the populations of source and destination societies' population sizes and inversely related to square of distance between them.

B. A Migration Model

The *social- gravity model*, which has been framed by us for the modeling the behavior of migration process occurring between societies, is a multi-dimensional model based on the concepts of system modules of "input, process and output" as described in Figure 1.

1) `Input` : comprises of relevant information related to the agent or set of agents who want to migrate from their society of origin to some other society of destination. This information also contains the name

and location of the destination society or societies to which agents wish to migrate. Therefore, all the information related to the **"Who?"** and **"Where?"** part of migration process is provided in this module of migration model. This input information launches the individual agent's decision making process, which is known as *migration behavior*, which is going to help agent decide, whether to migrate or not, hence serving a stimulus to migration process.

2) `Migration Behavior` : comprises of the decision - making part and behavior of agent(s), who want to migrate from their society of origin into some other destination society, and this behavior is affected by societal and personal push- pull factors existing in the source or origin society. In this module, all the *material as well as not-material benefits* related to migration process, are compared with costs which will be incurred for migration, and it is decided by agent, that **"When?"**, it is going to initiate the process of migration. If everything goes well, and benefits over-weigh the costs incurred for migration, then agents initiate the migration process, other wise no migration will be performed and the process of migration will stop altogether.
3) `Process` : is the module in which migration process is actually realized, and the agent from its society of origin is relocated to the destination society, it has wished for. The theory for the migration function has already been described in the previous subsection by using equations 1 and 2.
4) `Societal-Economic Factors` : comprise of all the societal and economic push-pull factors existing in source and destination societies, which become very genuine causes and reasons for agents to migrate between them and become very vital **"Why?"** part of migration process. They have already been explained in the section dedicated to them. They help agents to evaluate the *material benefits*, which can be realized from agent migration process.
5) `Individual Factors` : comprise of personal push-pull factors existing in source and destination societies, which lead to agent migration. These factors also comprise **"Why?"** part of agent migration process and help agents evaluate all the *non-material benefits*, which can be realized from agent migration process.
6) `Output` : is the outcome of the process of migration, in which agent from its society of origin gets relocated to the destination society, it has chosen as migration sink.
7) `Consequences` : are various after effects related to the process of migration, which can be categorized at both societal and personal levels of source and destination societies and the agent(s), who have migrated between them. They form **"What?"** part of the agent migration process. All the consequences related to agent migration process have been already illustrated in the previous section, which was dedicated to their description.
8) `Feedback` : consists of the individual agent's level of satisfaction or dissonance, after it has migrated to the destination society and have become aware of the consequences resulted from the migration process. This is fed to the member agents of source society, in order to formulate their decision to migrate from the same to any destination society.

The *migration behavior* $\mathbf{MB}$ can be stated as

$$\mathbf{MB} = f\,(\mathbf{RI} + \mathbf{SE}\text{ factors} + \mathbf{Individual}\text{ factors}) \qquad (3)$$

where $\mathbf{MB}$ = migration behavior,

f = is a function of (i.e., result of certain variables and factors),

$\mathbf{RI}$ = Relevant Information or input variables,

$\mathbf{SE}$ = Societal and Economic Factors,

and $\mathbf{I}$ = Individual Factors respectively.

The migration model, which has been described by us, is based on migration behavior, which is derived from behavior of an agent, and the migration function, which is facilitating the migration process. There are three parties involved in migration process, i.e., an agent, the society of origin, and the destination society. All three of them help stimulate the process of migration and perpetuate it.

VIII. Conclusions

In this paper, we have presented, almost all the dynamics related to agent migration process (specifically organizational and external migration process), in artificial agent societies, which have not yet been covered under the dynamics of artificial agent societies. As, mentioned in the related work section, many authors have presented their work on agent migration, but none of them is providing deep insight into the dynamics of agent migration process. All the dynamic aspects which are discussed by us for agent migration are inspired by the dynamic aspects of human migration, as both of them have quite a number of similarities. We have elaborated on various types of agent migration, various causes which lead to agent migration between artificial agent societies and various after effects or consequences of agent migration also. Also, we have formulated in this paper, an empirical agent migration model, i.e., W5 Social-Gravity Model, to model the complex behavior of multi-dimensional, dynamic migration process, which is responsible for moving agents from their society of origin to any other destination society. This migration model uses a two functions, i.e., migration function and migration behavior function for modeling the spatial phenomenon of agent migration process. Also, all the five (5) W's (Who?, Where?, Why?, What?, When?), related to the migration process have described in this model by using its various modules which are based on system's concept of operation.

IX. Future Work

In this paper, we have covered various types, causes and consequences related to agent migration process and have formulated a 5W Social-Gravity model for agent migration process also, which is used to model the behavior of complex agent migration process for the migration of agents between societies. Although, we tried to cover many aspects of dynamics of agent migration process in our presented literature, but still many aspects such as discussion of migration metrics (which will illustrate the number of agent migrating into and out of the source and sink (destination) agent societies and hence, measuring the net and gross migration of the society can be taken up as future research directions. And, migration protocols, which will be governing the migration process in agent societies can be framed. The discussion of various types of patterns of migration, the concept of re-migration process and emergence of migration norms, can also be considered as few of the future research challenges in agent migration process.

X. Acknowledgements

The authors are really grateful to the reviewers for their valuable suggestions and comments which they have given for this manuscript. They also would like to express thanks to the editor of IJARAI for providing very timely help for publishing this manuscript.

References

[1] R. H. Bordini, *Contributions to an Anthropological Approach to the Cultural Adaptation of Migrant Agents*, Department of Computer Science, University College London, London, U.K., 1999.

[2] R. H. Bordini, *Linguistic Support for Agent Migration*, Universidade Federal do Rio Grands do Sul (UFRGS), Instituto de Informatica, Porto Alegre, RS -BRAZIL, 1995.

[3] D. Z. S. Andrew and M. Zara, *Immigration Bahviour : Towards Social Pscychologial Model for Research*, In Proceedings of ASBBS Annual Conference, ASBBS Annual Conference: Las Vegas, Vol. 20(1), February 2013. ou

[4] M. Dastani and V. Dignum and F. Dignum, *Role-Assignment in Open Agent Societies*, In Proceedings of AAMAS'03: Second International Joint Conference on Autonomous Agents and MultiAgent Systems, Melbourne, Australia, ACM, 2003.

[5] V. Dignum and F. Dignum, *Modeling Agent Societies: Coordination Frameworks and Institutions*, In Proceedings of 10th Portuguese Conference on Progress in AI, Knowledge Extraction, MAS Logic Programming and Constraint Solving, 2001.

[6] V. Dignum and F. Dignum and L. Sonen Berg, *Towards Dynamic Reorganization of Agent Societies*,In Proceedings of CEAS: Workshop on Coordination in Emergent Agent Societies at ECAI2004, Valencia, Spain, September 22-27, 2004.

[7] R. M. Eijk and F. S. de van Boer and W. van der Hoek and Ch. J. J. Meyer, *Open Multi-Agent Systems: Agent Communication and Integration*, In Proceedings of Intelligent Agents VI ATAL 1999, pp. 218-222, 1999.

[8] N. Gilbert and R. Conte, *Artificial Societies: The Computer Simulation of Social Life*, UCL Press, London, 1995.

[9] N. Glasser and P. Morignot, *The Reorganization of Societies of Autonomous Agents*, In Proceedings of 8th European Workshop on Modeling Autonomous Agents in MultiAgent World, Ronneby, Sweden, LNCS 1237, Springer, May 13-16, 1997.

[10] M. J. Greenwood, *Modeling Migration*, Encyclopedia of Social Measurement, Elsevier, Vol. 2, 2005.

[11] H. S. B. Filho and F. B. de Lima Neto and W. Fusco, *Migration and social networks — An explanatory multi-evolutionary agent-based model*, 2011 IEEE Symposium on Intelligent Agent (IA), IEEE, Paris, 11-15 April, 2011.

[12] H. de Haas, *Migration and Development : A theoretical perspective*, International Migration Review, Vol. 44(1), pp. 1-38, 2010.

[13] F. M. Hafizoglu and S. Sen, *Patterns of Migration and Adoption of Choices By Agents in Communities*, Eds. Conitzer, Winikiff, Padgham and van der hock, In Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012), International Foundation for Autonomous Agents and MultiAgent Systems, Valencia, Spain, 5-8 June, 2012.

[14] R. King, *Theories and Typologies of Migration*, International Migration and Ethnic Relations, Willy Brandt Series of Working Papers, 3/12, Malmo Institute for Studies of Migration, Diversity and Welfare (MIM), Malmo University, Sweden, 2012.

[15] E. S. Lee, *A Theory of Migration*, Demography, Vol. 3(1), pp. 47-57, 1966.

[16] G. Lekeas and K. Stathis, *Agents acquiring Resources through Social Positions: An Activity Based Approach*, In Proceedings of 1st International Workshop on Socio-Cognitive Grids: the Net as A Universal Human Resource, Santorini, Greece, 1-4 June, 2003.

[17] A.C. da. R. Costa and T.F. Hubener and R. H. Bordini, *On Entering an Open Society*, XI Brazilian Symposium on AI, Fortaleza, Brazalian Computing Society, pp. 535-546, October 1994.

[18] V. Dignum and F. Dignum and V. Furlado A. Melo, *Towards a Simulation Tool for Evaluating Dynamic Reorganization of Agent Societies*, Proceedings of WS. on Socially Inspired Computing, AISB Convention, 2005.

[19] J. R. Weeks, *The Migration Transition*, in *Population : An introduction to concepts and issues*, Ninth Edition, Wadsworth Learning, pp. 273-315, 2005.

[20] J. H. Zanker, *Why do people migrate ? A review of the theoretical literature*, Working Paper MGSoG/2008/WP002MPRA, Munich Personal RePEc Archive, Maastricht University, Maastricht Graduate School of Governance, Netherlands, January 2008.

[21] *Human Migration Guide, Xepeditions, National Geographic Society*, A tutorial.

[22] *Human Migration* , Wikipedia, The Free Encyclopedia

[23] H. Kaur and K.S. Kahlon and R. S. Virk, *A formal Dynamic model of Artificial Agent Societies*, International Journal on Information and Communication Technologies, Vol. 6(1-2), pp. 67-70, Januray-June 2013.

[24] M. Dastani and B. M. van Riemsdijk and J. Hulstijn and F. Dignum and Ch. J. J. Meyer, *Enacting and Deacting Roles in Agent Programming*, Eds. Odell J.J., Giorgini P., Muller J.P, AOSE, LNCS, Vol. 3382, pp. 189-204, Springer Heidelberg, 2004.

[25] R. Brown, *Group Processes: Dynamics Within and Between Groups*, John Wiley and Sons, 1991.

[26] H. S. Filho and F. B. de Lima Neto and W. Fusco, *Migration, Communication and Social Networks - An Agent-Based Social Simulation*, Complex Networks, Studies in Computational Intelligence, Springer Berlin Heidelberg, Vol. 424, pp. 67-74, 2013.

[27] H. de Haas, *The internal dynamics of migration processes: A theoretical inquiry.*, Journal of Ethnic and Migration Studies, Vol. 36(10), pp. 1587-1617, 2010.