T. N. Nagabhushan
V. N. Manjunath Aradhya
Prabhudev Jagadeesh
Seema Shukla
Chayadevi M. L. (Eds.)

Cognitive Computing and Information Processing

Third International Conference, CCIP 2017
Bengaluru, India, December 15–16, 2017
Revised Selected Papers

Springer

Communications in Computer and Information Science   801

Commenced Publication in 2007
Founding and Former Series Editors:
Alfredo Cuzzocrea, Xiaoyong Du, Orhun Kara, Ting Liu, Dominik Ślęzak,
and Xiaokang Yang

More information about this series at http://www.springer.com/series/7899

T. N. Nagabhushan · V. N. Manjunath Aradhya
Prabhudev Jagadeesh · Seema Shukla
Chayadevi M. L. (Eds.)

Cognitive Computing and Information Processing

Third International Conference, CCIP 2017
Bengaluru, India, December 15–16, 2017
Revised Selected Papers

Springer

Editors
T. N. Nagabhushan
Sri Jayachamarajendra College
  of Engineering
Mysuru, Karnataka
India

V. N. Manjunath Aradhya
Sri Jayachamarajendra College
  of Engineering
Mysuru, Karnataka
India

Prabhudev Jagadeesh
JSS Academy of Technical Education
Bengaluru, Karnataka
India

Seema Shukla
JSS Academy of Technical Education
Noida, Uttar Pradesh
India

Chayadevi M. L.
JSS Academy of Technical Education
Bengaluru, Karnataka
India

Preface

The editorial team is honored to announce the publication of the conference proceedings of the Third International Conference on Cognitive Computing and Information Processing (CCIP 2017). The conference, which was jointly organized by the Departments of Computer Science and Information Science and Engineering of JSS Academy of Technical Education, Bengaluru, was the third in the series.

The focus of the conference was to understand the recent advancements in the cognitive era of computing systems that emulate the capability of human mind. The conference provided a forum for researchers and practitioners to present and discuss new research results, practical applications, and also promote collaborative research activities in cognitive computing and information processing and related field.

There were 130 papers submitted to CCIP-2017 (General track and four special sessions, namely, Cognitive Computing in Video Analytics, Cognitive Computing for Applications on Smart City, Cognitive Computing in Medical Information Processing, and Women in Information Processing) and they all underwent a rigorous review process. Upholding the quality requirements, around 33% of the papers received were selected for presentation.

The conference also comprised a pre-conference tutorial from Dr. Manjunath S., Samsung Electro Mechanics, and Dr. Punitha, IBM and, as well as keynote talks from Prof. N. Sundararajan, NTU, Singapore, Prof. Michael Carl, Copenhagen Business School, Denmark, Dr. M. Pratama, NTU, Singapore, Prof. S. R. M. Prasanna, IIT Guwahati, Prof. Savitha Ramasamy, ASTAR, Singapore, and Dr. Vigneswaran, ASTAR Singapore.

We would like to gratefully acknowledge the support received from J.S.S. Mahavidyapeetha in organizing this conference. We are grateful to all the researchers, reviewers, speakers, and the organizing team for helping us attain the objectives of the conference. We are sure that this conference has opened up a new platform for the researchers to share their ideas for continuous improvement. We hope that the conference proceedings will be inspiring to readers to pursue further research in the area of cognitive computing and information processing.

March 2018

<div align="right">
T. N. Nagabhushan
V. N. Manjunath Aradhya
Prabhudev Jagadeesh
Seema Shukla
Chayadevi M. L.
</div>

Organization

Chief Patron

His Holiness

Jagadguru Sri Shivarathri Deshikendra Mahaswamigalavaru
Veerasimhasana Math, Sutturu Srikshethra, President, JSSMVP, Mysuru

Patrons

C. G. Betsurmath	JSSMVP, Mysuru
M. H. Dhananjaya	JSSMVP, Mysuru
C. Ranganathaiah	JSSMVP, Mysuru
B. G. Sangameshwara	JSSSTU, Mysuru
Karisiddappa	VTU, Belagavi
B. R. Umakanth	JSSMVP, Mysuru

Co-patrons

Mrityunjaya V. Latte	JSSATE, Bengaluru
T. N. Nagabhushan	SJCE, Mysuru

General Chairs

Prabhudev Jagadeesh	JSSATE, Bengaluru
Chayadevi M. L.	JSSATE, Bengaluru

Advisory Committee

H. N. Jagannatha Reddy	VTU, Belagavi, India
Satish Annigeri	VTU, Belagavi, India
D. K. Subramanian	IISc, Bengaluru, India
S. S. Iyengar	Florida International University, USA
Ajith Abraham	MIR Labs (Global Operations), USA
Buyya Rajkumar	University of Melbourne, Australia
Venkatesh Babu R.	IISc, Bengaluru, India
Erik Cambria	SCE, NTU, Singapore
K. Subramanian	NTU, Singapore
Haijun Rong	Xi'an Jiaotong University, China
V. Mani	IISc, Bengaluru, India
R. V. Babu	IISc, Bengaluru, India
F. Bremond	Inria-Sophia Antipolis, France

P. B. Sujit	IIIT, Delhi, India
S. Vasaily	Catholic University, Seoul, South Korea
G. S. Babu	SimTech, Singapore
S. Jagannathan	MUST, USA
Guang-Bin Huang	NTU, Singapore
B. Tripathi	HB Technology, India
Ravindran	IIT, Chennai, India
Shigeyoshi Tsutsui	Osaka Prefecture University, Japan
Kirill Krinkin	Russia
Pradeep Kumar	Amphisoft Technologies, Coimbatore, India
Harsha	Motorola, Bengaluru, India
Dinesh M. S.	Philips, Bengaluru, India
S. R. Mahadeva Prasanna	IIT, Guwahati, India
Suryakanth V. Gangashetty	IIIT, Hyderabad, India
S. Manjunath	Samsung, India

Steering Committee Chair

K. Chidananda Gowda	Kuvempu University, Shivamogga, India

Steering Committee

N. Sundararajan	EEE, NTU, Singapore
S. Suresh	SCE, NTU, Singapore
H. J. Kim	Korea University, Seoul, South Korea
Prabhu Shankar	UCDavis, USA
Oleg Medvedev	Moscow State University, Russia
Hanseok Ko	Korea University, Seoul, Korea
S. N. Omkar	IISc, Bengaluru, India
Jayadeva	IIT, Delhi, India
P. Nagabhushan	IIIT, Allahabad, India
Guru D. S.	University of Mysore, India
S. K. Padma	SJCE, Mysuru, India
Nagabhushana	JSSATE, Bengaluru, India
Swamy D. R.	JSSATE, Bengaluru, India

Organizing Chairs

N. C. Naveen	JSSATE, Bengaluru
Dayananda P.	JSSATE, Bengaluru

Technical Program Chair

V. N. Manjunath Aradhya	SJCE, Mysuru

Technical Program Co-chairs

D. V. Ashoka	JSSATE, Bengaluru
P. B. Mallikarjuna	JSSATE, Bengaluru

Special Session Chairs

B. S. Mahanand	SJCE, Mysuru
Dinesh R.	Samsung, India
Vinita Khemchandani	JSSATE, Noida
Anita Sahoo	JSSATE, Noida
M. P. Pushpalatha	SJCE, Mysuru
M. A. Anusuya	SJCE, Mysuru

Publication Chairs

Malini M. Patil	JSSATE, Bengaluru
Sneha Y. S.	JSSATE, Bengaluru
Seema Shukla	JSSATE, Noida

Tutorial Chair

Nagasundara K. B.	JSSATE, Bengaluru

Finance Chair

Snehalatha N.	JSSATE, Bengaluru

Publicity Chairs

Sharana Basavana Gowda	JSSATE, Bengaluru
Abhilash C. B.	JSSATE, Bengaluru
Anil B. C.	JSSATE, Bengaluru

Organizing Committee

Bhavani B. H.	Shwetha Kaddi	Rohitaksha
Rajeshwari K. S.	Savita S.	Rashmi B. N.
Shanthala K. V.	Pradeep H. K.	Vinutha H. D.
Pooja H.	Mahesh Kumar	Prasad M. R.
Manjunath Talawar	Sreenatha	Sangeetha H. S.
Niranjan K. C.	Renuka Rajendra	Sumathi H. R.

Mamatha G.   Nethravathi   Fathima Afroz
Apsara M. B.   Anitha P.   Nagashree S.
Nagamani N. Purohit   Sudha P. R.   Sahana V.
Rekha P. M.   Sowmya K. N.

Keynote Addresses

A Minimal Cognitive Model for Translating and Post-editing

Michael Carl

Centre for Research and Innovation in Translation and Translation Technology,
Copenhagen Business School, Frederiksberg, Denmark
`mc.ibc@cbs.dk`

This study investigates the coordination of reading (input) and writing (output) activities in from-scratch translation and machine translation post-editing. We segment logged eye movements and key logging data into minimal units of reading and writing activity and model the process of post-editing and from-scratch translation as a Markov model. We show that the time translators and post-editors spend on source or target text reading predicts with a high degree of accuracy how likely it is that they engage in successive typing. We further show that the typing probability is also conditioned by the degree to which source and target text share semantic and syntactic properties. The minimal Markov model describes very basic factors which are suited to model the cognitive processes occurring during translation and post-editing.

Human Cognition Inspired Learning Strategies for Particle Swarm Optimization Algorithm

Narasimhan Sundararajan

Nanyang Technological University, Singapore
ENSUNDRA@ntu.edu.sg

These days, the nature of global optimization problems especially for engineering systems have become extremely complex and it is difficult to locate the true optimum solutions. For these types of problems, finding the optimal/near-optimal solution in a quick and efficient way is very important and here only search based methods are found to be effective. Among the search-based methods, nature inspired optimization algorithms are providing much better solutions. A well-known nature inspired method, the Particle Swarm Optimization (PSO) algorithm has been mostly preferred due to its simplicity and ability to provide better solutions and it has been proven to be more effective for solving complex real-world problems. The limitations associated with PSO have been extensively researched and different modifications, variations and refinements to PSO have been proposed for enhancing the performance of the algorithm. These include parameter tuning, neighbourhood topologies and unique learning strategies. The PSO variants with unique learning strategies are found to be more effective in enhancing the convergence characteristics of the basic PSO algorithm. All these variants have utilized the behaviour of the swarm which limited the usage of intelligence and motivated us towards exploring human cognitive learning principles for PSO.

As discussed in learning psychology, human beings are known to be intelligent and have good social cognizance. Therefore, any optimization technique employing human-like learning strategies should prove to be more effective. In this talk, first the basic PSO and its variants will be presented to show the status of current work. Then the human cognition inspired learning strategies will be introduced to address the limitations of PSO and enhancing its convergence characteristics. By mimicking the human-like behaviour, the PSO algorithm has shown faster convergence and closer to the optima over diverse set of problems being a potential choice for complex real-world applications. Recent developments in this area undertaken by our group will be highlighted.

Speech Processing: Human Cognition vs Cognitive Computing

S. R. Mahadeva Prasanna

Indian Institute of Technology, Guwahati, India
`prasanna@iitg.ac.in`

Cognitive computing tries to address challenging problems. These include processing natural data like speech. Human cognition has remarkable ability to perform spoken communication. Most of the speech processing tasks are trivial for human cognition. Alternatively, they prove to be harder for cognitive computing. This talk will explain some speech processing talks to bring the challenges in the field of cognitive computing for speech processing.

Data Stream Analytics in Complex Environments

Mahardhika Pratama

Nanyang Technological University, Singapore
`mpratama@ntu.edu.sg`

The era of big data in highly complex environments calls for algorithmic development of advanced machine learning techniques and visualizations to transform massive amounts of information into useful references to help decision making process in real-time. This talk aims to discuss online real-time strategies for data stream analytics that provide concrete solutions to unsolved issues in data streams analytics, namely uncertainty in data distribution, uncertainty in data representation, uncertainty in data dimensions, uncertainty in data processing, and uncertainty in data visualization.

Understanding Brain Effects
of Neuro-Psychiatric Disorders via MRI

Vigneshwaran

Agency for Science Technology and Research (A*STAR), Singapore
vigneshwaran_subbaraju@sbic.a-star.edu.sg

Neuroimaging technologies, especially functional and structural MRI, are increasingly being looked upon as tools capable of revealing the complex mechanisms in the human brain that underlie several neuro-psychiatric conditions. Thus, they have the potential to influence not just psychiatric diagnosis but also therapy and treatment. The main stumbling block in translating these tools from the laboratory to the mainstream clinical environment is the dearth of sufficiently large datasets to derive generalizable and reproducible conclusions. This is being addressed by the research community, by collating and releasing disorder specific datasets for open research. Few examples are, the ABIDE, ADHD-200, ADNI and the SchizConnect initiatives. The release of these large datasets presents an unprecedented opportunity to perform large scale studies and uncover the neural mechanism underlying these conditions. The adoption of the analysis methods in the mainstream clinical process will depend on two factors: accuracy and interpretability. The emphasis on clinical interpretability is so significant that the analysis has so far been restricted to simple statistical tools. In this talk, I will discuss the application of spatial filtering to fMRI analysis that can help achieve high accuracy and also provide better interpretability. Based on the large datasets in ABIDE and ADHD-200, I will also show the how these methods can be used touncover potential compensatory effects within the brain as well.

Predictive Analysis for Personalized Clinical Decision Support

Savita Ramasamy

Institute for Infocomm Research, Data Analytics Department, Singapore
ramasamysa@i2r.a-star.edu.sg

Public healthcare providers are faced with the challenge of providing quality care at lesser cost. The availability of huge heterogenous data sets available in such healthcare settings provide a wide scope for the development of clinical decision support using predictive analytics and data mining tools. Existing clinical decision support tools do not address the individual needs of patients, and often do not translate to care decisions that clinicians can leverage on. This talk will focus on our recent work in the development of a predictive analytics solution for a personalized clinical decision support tool. The decision support tool comprises of two main components, namely, neural network predictor to predict the readmission risk for patients and the cost of care, and a multi-objective optimization that optimizes the risk and cost based on the model-based outcomes, constrained based on the desired clinical outcomes. Preliminary results from the study will also be discussed.

Contents

Cognitive Computing and Its Applications

Cognitive Computing in Video Analytics

Cognitive Computing in Medical Information Processing

Eyes Open and Eyes Close Activity Recognition Using EEG Signals

Barjinder Kaur[1]([✉]) [ID], Dinesh Singh[1] [ID], and Partha Pratim Roy[2] [ID]

[1] DCRUST, Sonepat, India
kaur.barjinder@gmail.com, dinesh.madhav@gmail.com
[2] IIT Roorkee, Roorkee, India
proy.fcs@iitr.ac.in

Abstract. So far Electroencephalography (EEG) has been analyzed by the research community for interaction with the computers. Studies regarding EEG signals has gained attention in the recent past as it gives an alternate way of communication for the persons suffering from partially or fully paralytic disability. Every second different activities are performed by millions of neurons. Decoding and detecting such complex activity of the brain while analyzing the EEG signals is a challenging task. In this paper, we have proposed an activity recognition system using EEG signals. The two activities, namely, eyes open (EO) and eyes close (EC) have been considered in this work. The recorded signals are then decomposed using Discrete Wavelet Transform (DWT) to analyze the impact of both the activities. The recognition of activities has been performed using Support Vector Machine (SVM) classifier. For experimentation, a publicly available dataset i.e. PhysioNet consisting data of 109 users while performing one minute EO and EC activity has been used. A notable activity recognition rate of 86.08% has been recorded using gamma band feature. The paper further proposes that the system can be used as a reference to detect different types of activities performed at different instance of time and for rehabilitation purposes also.

Keywords: Electroencephalography (EEG) · Eyes open (EO) · Eyes close (EC)
Discrete Wavelet Transform (DWT) · Support Vector Machine (SVM)

1 Introduction

Brain Computer Interface (BCI) is a mechanism of acquiring the neuro signals. These acquired signals store mirage of information that has been proved helpful in diagnosing various diseases, and providing help for motor-disabled people. Various methods like Electroencephalogram (EEG), Magnetoencephalogram (MEG), functional Magnetic Resonance Imaging (fMRI), Positron Emission Tomography (PET), etc. are available to capture the brain functionality. However, taking into account the cost and easy portable aspect, EEG has been widely used by researchers [5, 6, 14] as it provides a noninvasive monitoring method to analyze the brain recordings, and is frequently used to implement different BCI systems.

The brain oscillations are usually measured between 0.5–40 Hz frequency range.

Thus, based on these measurements EEG signals are segregated into five rhythms, namely, delta (0.5–<4 Hz), theta (4–7 Hz), alpha (8–15 Hz), beta (16–32 Hz) and gamma (32 Hz *above*) that are omnipresent in different parts of the brain [5]. EEG helps in detecting the various abnormalities of the brain like sleeping disorders, emotional variance, seizures detection, and in the recent past researchers have been widely using EEG signals for proposing user identification systems [5, 6], Neuromarketing [14], rating prediction systems [4] and envisioned speech recognition [7]. Although normal EEG patterns are easily recognizable, but to evaluate the reason of change in EEG signals is a cumbersome task. The human brain consists of about 86 million neurons which helps in controlling all the conscious and unconscious daily decisions [3]. Recognizing and detecting the activity that could have lead to abnormal behavior of the neuro signals has made the researchers to analyze the activity happened during the course of action. Saghafi et al. [11] used power spectrum, variance, maximum, and range features to discriminate between the eye state change. The authors have used Logistic Regression and Artificial Neural Network (ANN) classifier to evaluate the system performance whereas an auto regressive power feature has been used to discriminate between eyes close and eyes open state. High frequency power has been reported in frontal, parietal, and occipital areas in eyes close state.

Öner et al. [8] investigated the moments of the eyes using one-channel. The authors focused on the ocular movements of the eye. The features were selected using analysis of variance (ANOVA) function and a linear fit model has been used to report the higher presence of delta rhythm while performing the task. However, their study was conducted with EEG data of 2 subjects. In [13], the authors have analyzed left and right hemisphere of the brain using EEG signals recorded in eyes open state. Occipital lobe and left hemisphere were found more active by analyzing alpha band waves. Qidwai et al. [10] analyzed the data of six subjects to discriminate the two motor activities EO and EC. Different angle features of the eye states were used for feature extraction and fed directly into a Multilayer Perception (MLP) based Neural Network classifier for recognition purpose. The authors in [2] have proposed a methodology to detect the left and right fist motor movement activity using EEG signals of 6 subjects only. The signals were filtered with a notch filter and were further preprocessed using Automatic Artifact Removal (AAR) technique. They have used DWT based signal decomposition technique and three different features, namely, power, mean and energy were extracted for classification purpose. Similarly, the classification of imagined left and right fists movement was proposed in [1] using EEG signals of 20 users. Root Mean Square (RMS) and Mean Absolute Value (MAV) features were extracted and modeled using SVM and ANN classifiers where accuracies of 84.5% and 82.1% were recorded from 3 channels, respectively.

Accurate recognition of the performed activity by analyzing EEG signals not only proves helpful in healthcare but also contributes towards self-evaluation of changes in daily health parameters. Therefore, in this paper, we analyzed the neuro signals for activity recognition. For this, two activities, namely, eyes open (EO) and eyes close (EC) are considered. Figure 1 depicts the flow diagram of the proposed approach, where brain signals are recorded in one of the two states i.e. EO and EC. The recorded signals are then preprocessed and two features, namely, Standard Deviation (STD) and RMS are

extracted. Finally, the recognition of activities is performed using Support Vector Machine (SVM) classifier.
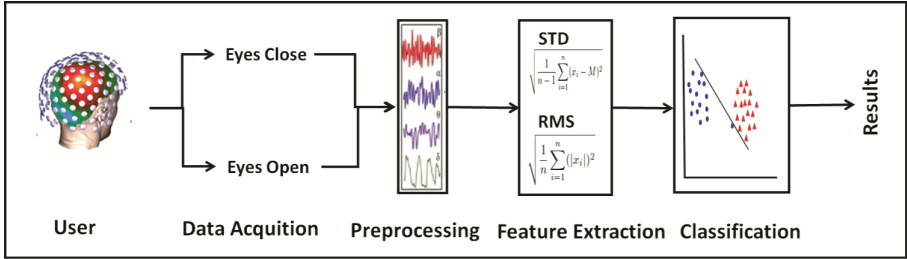
Fig. 1. Framework for eyes open and eyes close activity recognition using EEG Signals.

Rest of the paper is organized as follows. The description of signals preprocessing and feature extraction techniques are presented in Sect. 2. Section 3 presents the experimental results. Finally, we conclude the paper in Sect. 4 by highlighting some future possibilities of the present work.

2 Signal Preprocessing and Features

Here, we present the details of the preprocessing techniques used to filter the raw signal data and the features extracted before classification of the activities.

Preprocessing. It has been analyzed that while recording the data, the EEG signals are always contaminated with unwanted noise from the surroundings, hair, ocular movements etc. To get the enhanced performance of the system these artifacts need to be removed. Bio-signals being non-stationary in nature tends to get distorted due to sensitiveness of the device. Also, it has been researched that the human brain activities are limited upto 40 Hz, where higher value signifies noise. Therefore, to remove such noise and to restrict the frequency till 40 Hz, a low pass filter is used. For this, we have used a Butterworth low pass filter of 5th order to preprocess [9] the EEG signals. The filter can be defined using (1).

$$H(z) = \frac{q(1) + q(2)z^{-1} + \ldots + q(n+1)z^{-n}}{1 + p(2)z^{-1} + \ldots + q(n+1)z^{-1}} \quad (1)$$

where n denotes the filter order with normalized cutoff frequency Wn. It returns the filter coefficients of length $n + 1$ as row vectors q and p, along with the coefficients z in descending powers.

Wavelet Analysis and Statistical Features. As EEG signals are non-stationary, therefore, DWT provides a flexible way to analyze the signals in time-frequency domain [5, 14]. It is an analysis technique that is based on sub-band decomposition coding providing easy implementation and also helps in reducing the computation time. DWT is

widely used in applications including emotion detection, security, seizure detection and gaming, etc. In this work, we have used Daubechies-4 (DB-4) wavelet to decompose the EEG signals into five sub-bands, namely, gamma, beta, alpha, theta and delta. Next, two statistical features i.e. STD and RMS have been extracted from the sub-bands. More details about the DWT and features can be found in [5].

2.1   Activity Recognition Using SVM

SVM has been widely used by researchers for the classification of EEG signals. It separates the samples into two different classes using the hyperplane [5]. It transforms the original feature space into a high dimensional space with maximum margin. Different kernel functions are used to perform the mapping by performing calculations in the data space and returns it to feature space. For given training data, $(x_1, y_1), (x_2, y_2), (x_n, y_n)$, the decision function $f(x)$ is defined in (2).

$$f(x) = sgn\{(w \cdot x) + b\} = sgn\left(\sum_{i=1} y_i a_i k < x_i, x > + b \right) \qquad (2)$$

Here $(w \cdot x)$ represents the dot product between weight vector w and x data with binary class. The parameter b is defined as a scalar that is often referred to as the bias. The term α_i denotes the embedding coefficients and $k <x_i, x>$ is the kernel represented by the dot product i.e. $<., .>$. The experimental results are classified by varying kernel functions i.e. linear and Radial Basis Function (RBF).

3   Results

Here, the details of the dataset and the results of activity recognition are presented. The computation of the results has been performed using 2-fold cross validation. The details are as follows.

3.1   DataSet Description

For activity recognition, a freely online available EEG-based motor movement and imaginary dataset provided by PhysioNet BCI [12] has been used. EEG signals were recorded from all 64 channels for 1 min. Two baseline tasks, eyes open (EO) and eyes close (EC) resting state have been used to collect the data from 109 users and are considered in this work. In order to detect each activity accurately, the EEG data has been segmented into 10 s. Thus, a total of 1308 EEG files (i.e. 654 for EC and 654 for EO) have been created for analysis.

3.2   Activity Recognition Results Using SVM

The SVM classifier has been trained for recognizing the activity performed by the users using the extracted features. The accuracies have been computed for all 5 sub-bands.

The recognition results are depicted in Fig. 2, where the maximum accuracy of 86.08% is recorded on gamma band features.
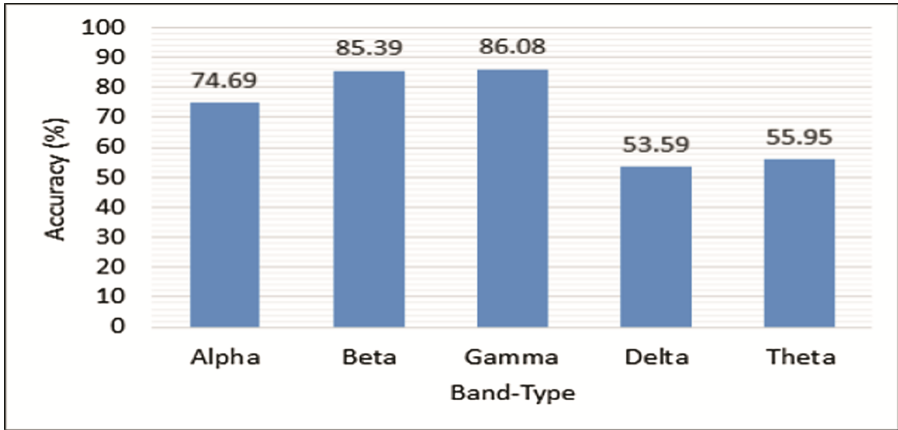
Fig. 2. Activity recognition results on different sub-bands.

Recognition results for EO and EC have also been evaluated for the gamma band waves as depicted in Fig. 3, where accuracies of 92.05% and 80.10% are recorded for both the activities, respectively.
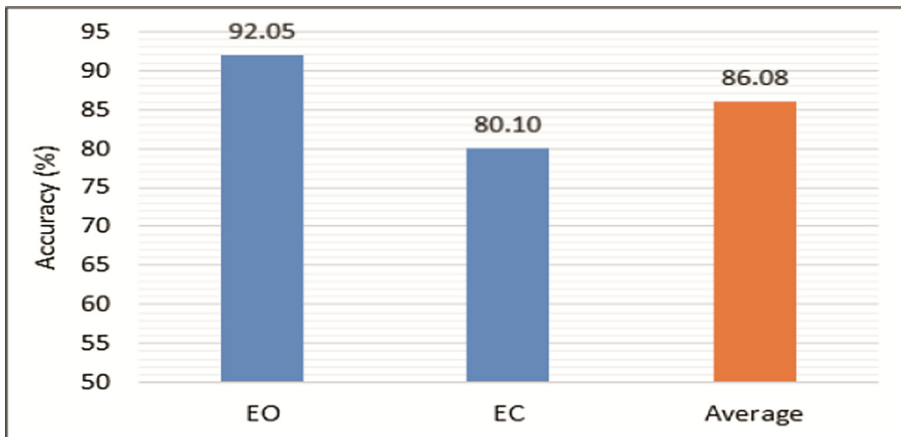
Fig. 3. Recognition rates of EO and EC activities.

A comparative analysis has also been performed by varying SVM kernels. For this, we have trained the SVM with Linear kernel using gamma band features and compared the performance with RBF kernel for activity recognition purpose. The comparison is depicted in Fig. 4, where the proposed RBF kernel based activity recognition outperform the other by a margin of 14.37%.
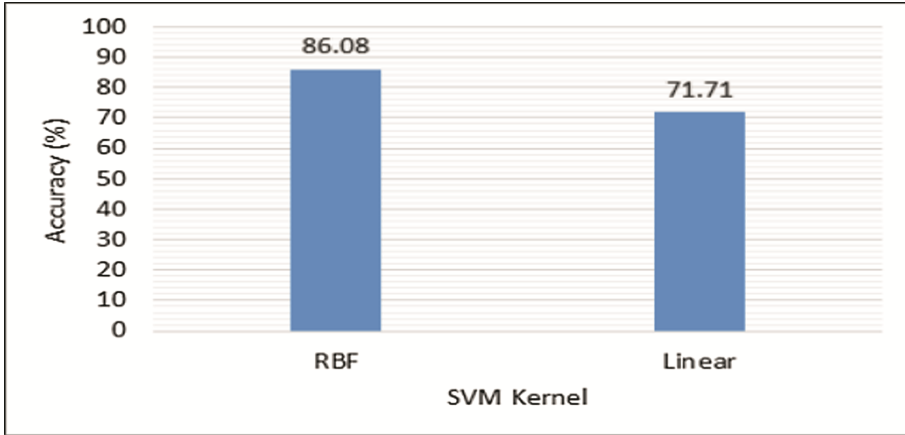
Fig. 4. Comparative performance analysis with linear kernel based SVM.

4   Conclusion

The human brain is a complex structure. Therefore, recognizing the activity performed by the user is a challenging task. As eyes open (EO) and eyes close (EC) are the most commonly used baseline activities that users perform while recording EEG signals. In this paper, we have proposed an activity recognition framework for the two activities using EEG signals. DWT based signals decomposition technique has been used to extract statistical features. An online freely available dataset Physionet has been used in this work that consists EEG signals data of 109 users. The recognition of activities has been performed using SVM classifier achieving an accuracy of 86.08%. In future, more activities can be recognized using the proposed framework that will prove helpful for various healthcare applications.

References

1. Alomari, M.H., Baniyounes, A.M., Awada, E.A.: EEG-based classification of imagined fists movements using machine learning and wavelet transform analysis. Int. J. Adv. Electron. Electr. Eng. 83–87 (2014)
2. Alomari, M.H., Samaha, A., AlKamha, K.: Automated classification of L/R hand movement EEG signals using advanced feature extraction and machine learning. Int. J. Adv. Comput. Sci. Appl. **4**, 207–212 (2013)
3. Azevedo, F.A., Carvalho, L.R., Grinberg, L.T., Farfel, J.M., Ferretti, R.E., Leite, R.E., Lent, R., Herculano-Houzel, S., et al.: Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. J. Comp. Neurol. **513**, 532–541 (2009)
4. Gauba, H., Kumar, P., Roy, P.P., Singh, P., Dogra, D.P., Raman, B.: Prediction of advertisement preference by fusing EEG response and sentiment analysis. Neural Networks **92**, 77–88 (2017)
5. Kaur, B., Singh, D., Roy, P.P.: A novel framework of EEG-based user identification by analyzing music-listening behavior. Multimed. Tools Appl. **76**, 1–22 (2017)

6. Kumar, P., Saini, R., Roy, P.P., Dogra, D.P.: A bio-signal based framework to secure mobile devices. J. Netw. Comput. Appl. **89**, 62–71 (2017)
7. Kumar, P., Saini, R., Roy, P.P., Sahu, P.K., Dogra, D.P.: Envisioned speech recognition using EEG sensors. Pers. Ubiquit. Comput. **22**(1–15), 2017 (2017)
8. Öner, M., Hu, G.: Analyzing one-channel EEG signals for detection of close and open eyes activities. In: International Conference on Advanced Applied Informatics, pp. 318–323 (2013)
9. Palaniappan, R., Ravi, K.: A new method to identify individuals using signals from the brain. In: International Conference on Information, Communications and Signal Processing, vol. 3, pp 1442–1445 (2003)
10. Qidwai, U., khazaal Shams, W.: A source-discrimination approach for detection of ASD using EEG data. Int. J. Biosci. Biochem. Bioinf. **3**, 492–496 (2013)
11. Saghafi, A., Tsokos, C.P., Goudarzi, M., Farhidzadeh, H.: Random eye state change detection in real-time using EEG signals. Expert Syst. Appl. **72**, 42–48 (2017)
12. Schalk, G., McFarland, D.J., Hinterberger, T., Birbaumer, N., Wolpaw, J.R.: BCI 2000: a general-purpose brain-computer interface (BCI) system. IEEE Trans. Biomed. Eng. **51**, 1034–1043 (2004)
13. Trivedi, P., Bhargava, N.: Comparing alpha wave activity of left and right hemisphere of brain recorded using EEGlab. Int. J. Sci., Eng. Technol. Res. **6**, 170–174 (2017)
14. Yadava, M., Kumar, P., Saini, R., Roy, P.P., Dogra, D.P.: Analysis of EEG signals and its application to neuromarketing. Multimed. Tools Appl. **76**(18), 19087–19111 (2017)

Analysis of Breast Thermal Images Using Anisotropic Diffusion Filter Based Modified Level Sets and Efficient Fractal Algorithm

S. Prabha[1(✉)] , S. S. Suganthi[2] , and C. M. Sujatha[3]

[1] Hindustan Institute of Technology and Science, Chennai, India
sprabha@hindustanuniv.ac.in
[2] Tata Elxsi, Chennai, India
suganthiss@gmail.com
[3] CEG Campus, Anna University, Chennai, India
sujathacm1@gmail.com

Abstract. Asymmetry analysis of thermal images plays a prominent role in automated diagnosis of normal and abnormal breast tissues. Accurate segmentation of breast tissues further enhances the analysis. In this work, asymmetry analysis of breast thermal images is carried out using efficient fractal algorithm. Prior to feature extraction, the breast tissues are delineated by integrating anisotropy diffusion filter into modified level set framework. In order to validate the segmented results against ground truth images, Overlap measures are estimated. Breast regions such as left and right are separated from the segmented tissues. A set of binary images are generated by subjecting the separated breasts to thresholded binary decomposition (TBD) method. The features such as fractal dimensions, mean gray level and pixel count are computed from the resultant binary images. The segmented results had shown high degree of correlation against the ground truth images. Accuracy of the segmented results is found to be 98%. Asymmetry analysis using TBD method shows that the features are capable of discriminating the change in structural information which is caused by varied metabolic conditions. Finally, mean gray level feature show 11% improvement in demarcating normal and abnormal breast tissues. Therefore, the proposed method appears to be efficient in computer diagnosis of early breast abnormalities using infrared images.

Keywords: Breast thermogram · Anisotropic diffusion filter · Level sets
Asymmetry analysis

1 Introduction

Breast cancer is originated by the unrestrained development of abnormal cells in breast tissues. A report indicates that the second leading cancer among women is breast cancer after lung disease [1]. The mortality rate is high due to late identification of the disease [2]. Therefore, proper screening and early identification is necessary to develop the

survival rate of breast cancer. Breast imaging systems generates breast images using different techniques for diagnosis and analysis of breast cancer. The current gold standard screening tool is mammography and it has several limitations such as high cost, radiation exposure, painful and uncomfortable imaging procedure [3]. Mammography had showed the result as 20% false positive for dense breast tissues [3].

Medical Thermography (MT) is most extensively established as an accessory tool for breast cancer diagnosis [4]. It has been estimated that its sensitivity and specificity are 90% [5]. This technique is used to quantify and visualizes the skin surface temperature using high sensitive IR camera. Accurate segmentation of breast tissues is a challenging one in the regions of lower breast boundaries and infra mammary folds owing to the limitations of IR imaging [6]. Extraction of lower breast boundaries using Hough transform and several edge techniques are adopted for further analysis. A review paper demonstrates detail report of segmentation and asymmetry analysis in breast thermograms [7].

The asymmetrical thermal patterns are observed in abnormal breast tissues whereas the healthy human body shows symmetrical pattern [8]. Therefore, asymmetry analysis plays the significant role in quantitative analyses of breast thermograms. Several thresholding techniques are employed for the extraction of desired regions in breast thermal images [9]. Multi-threshold Otsu method demonstrates the gray level regions of an image which is split into many classes, in an optimal manner by maximizing the between class variance [10]. Efficient extraction of textural patterns had proved to be appropriate in content based image retrieval (CBIR) and for classification of images by two-threshold binary decomposition method [11].

In this work, new edge map is integrated in to anisotropic diffusion filter based level set method instead of Gaussian filter. The segmented results are compared against ground truth images and validated by using overlap measures. Asymmetry analysis is performed using TBD method for extracting symmetrical and asymmetrical thermal patterns in breast tissues.

2   Materials and Methods

A breast thermal image to perform this experimentation is acquired from the project PROENG (http://visual.ic.uff.br/). Total 120 images are considered to validate the accuracy of segmentation procedure [12]. To perform asymmetry analysis, 20 images are considered which are diagnosed as having one among the three different abnormalities such as cysts, carcinoma and fibro adenoma either on left or right breasts.

2.1   Anisotropic Diffusion Filter Based Modified Level Sets

The regularization term in the conventional level set method without reinitialization is to maintain signed distance function property. When curve approaches the true boundary, regularization term becomes redundant during the curve evolution process. To address this problem, [13] modified level set method is introduced by integrating the

geometrical information into the regularization term and thus it is made forceful along with the curve progression. Thus the modified level set evolution equation is given as

$$\frac{\partial \varphi}{\partial t} = \tau \left\{ \mu g_{new} \left[\nabla^2 \varphi - \text{div} \left(\frac{\nabla \varphi}{|\nabla \varphi|} \right) \right] + v g_{new} \delta_\varepsilon(\varphi) \right\} \qquad (1)$$

Where φ is the initial contour, τ, μ and v are evolution constants. g_{new} is the edge map that is used for curve evolution [14]. In this work, anisotropic diffusion filter based edge map is integrated into the modified level set method instead of Gaussian filter.

2.2   Anisotropic Diffusion Filter Based Edge Map

In conventional level set formulation without re-initialization, the edge indicator function is obtained as inverse gradient magnitude of Gaussian smoothed input image I. This linear homogenous diffusion filter blurs the image edges as smoothing is performed in each location. Therefore, to avoid smoothing across boundaries, an anisotropic diffusion filter [15] is adopted, which performs smoothing within a region and preserves the sharp edge boundaries. It is specified as

$$I_t = \text{div}(c(x, y, t)\nabla I) = \frac{\partial}{\partial x}\left(c(x, y, t)I_x\right) + \frac{\partial}{\partial y}\left(c(x, y, t)I_y\right) \qquad (2)$$

Where conduction diffusion coefficient is assumed as a constant to performs smoothing operation in each location and is represented as c in the above equation. The gradient of brightness function (∇I) plays a vital role for performing region specific smoothing operation. The large gradient strength in inter-region edges is resulted as a low coefficient value whereas the low gradient strength in non-edge regions is resulted as high coefficient value. The possible function is

$$c(x, y, t) = \exp\left(-(\|\nabla I\|/K)^2\right) \qquad (3)$$

Edge strength threshold is indicated as K, a constant and must be tuned for a particular application. For larger values of K, the diffusion process will over smooth resulting in a blurred image whereas smaller values will stop the diffusion process in early iterations and hence, the smoothing operation. For the input image I, the anisotropic diffusion filtered image is defined as I_t and g_{new}, the edge map is defined as

$$g_{new} = 1/1 + |\nabla I_t|^2 \qquad (4)$$

This information g is exploiting as an edge indicator function for the LSF to evolve the initial contour towards the boundaries.

2.3 Validation of Segmented Results

The comparison is performed between the segmented results and four sets of ground truth images in order to understand the performance among inter person's variations. The regional statistics about the number of pixels in region of interest and non region of interest are calculated. The ratios of those statistics such as accuracy, sensitivity, specificity, positive predictive rate (PPR) and negative predictive rate (NPR) are calculated [16].

2.4 Asymmetry Analysis

Thresholded Binary Decomposition (TBD) method is employed to perform asymmetry analysis. The set of thresholds are computed from input image gray level distribution information by using this technique. Multi-level Otsu algorithm (MLOA) is adopted for this technique. The basic principle behind MLOA is for separating the image's pixels into many classes, and verifies the best threshold value which is obtained by maximizing the input image inter-class variance or minimizing the intra-class variance. Number of threshold value chosen for this method is four. TBD method decomposes the gray level input in to set of binary images.

The input grayscale image I is modeled by binary decomposition algorithm as a 2D function I(x, y), which is given by

$$I_b(x, y) = \begin{array}{l} 1 \ \ \text{if } I(x, y) \geq t \\ 0 \ \ \text{otherwise} \end{array} \qquad (5)$$

where I(x, y) is a grayscale value. Then, the fractal dimension is estimated from the resulted binary image. Then calculate the mean gray level and pixel counting for the corresponding region.

3 Results and Discussion

A classic breast thermogram is exposed in Fig. 1(a). The nature of thermal image has low contrast which results the breast boundaries close to lower breast regions and infra mammary folds are not discernible from surroundings tissues.

(a) Typical breast image   (b) Gaussian edge map   (c)Anisotropic Diffusion Filter Based Edge Map   (d) Segmented breast tissue   (e) Seperated left and right breast image
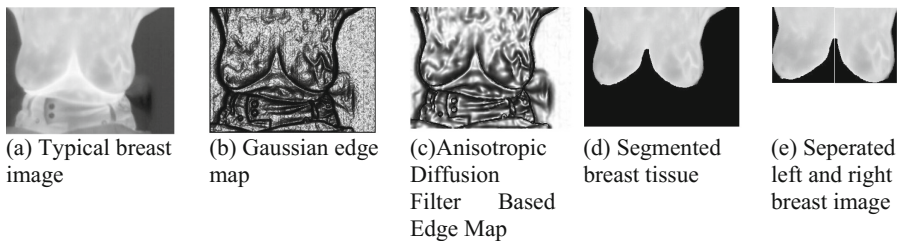
Fig. 1. (a) Typical breast image. (b) Gaussian edge map. (c) Anisotropic diffusion filter based edge map. (d) Segmented breast tissue. (e) Seperated left and right breast image

Figure 1(b) shows the edge map obtained using conventional Gaussian filter. Thick edges are observed which leads to false boundary during level set evolution. Figure 1(c) shows the edge map obtained using anisotropic diffusion filtered image. The edges are preserved, smoothed and sharpened by adopted filtering operation.

The raw image and generated binary mask are multiplied to segment the breast tissues as shown in Fig. 1(d). Separated breast regions such as left and right are shown in Fig. 1(e). The area of the segmented images and ground truth mask are calculated. A high degree of correlation (0.96) is observed and is shown in Fig. 2.
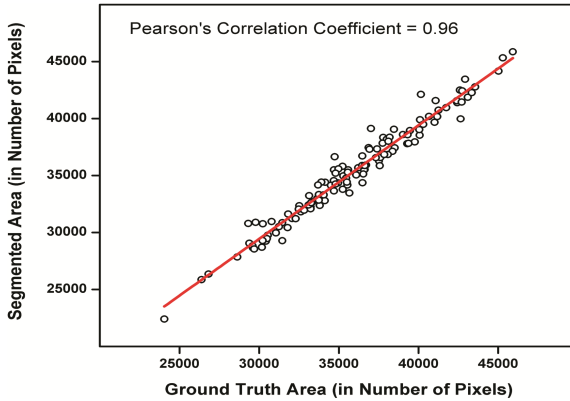
Fig. 2. Scatter plot between areas of ground truth masks and segmented images using proposed level set method for different threshold levels

In order to validate quantiatively, overlap measures are calculated and are shown in Fig. 3. Irrespective to the inter person's variations, the segmented results of the proposed method are observed to be more consistent with all overlap measures. An significant improvement in terms of accuracy and specificity is observed with the proposed segmentation method. Features such as Fractal dimension (FD), mean gray level (MGL)and pixel count (PC) are acquired from the separated breast regions. Binarization of seperated breast regions is performed through the application of different threshold levels.

The asymmetry in the thermal image is well differentiated using TBD features such as fractal dimension, mean gray level and pixel count values obtained for normal and abnormal conditions are represented in Table 1. Among four threshold values, third threshold level shows significant distinction among normal and abnormal breast tissues. This is originated by the employment of new blood vessels by abnormal tissues which establish the modification of vascular patterns and therefore, consequences in hot spots owing to asymmetry in temperature distributions between the breasts. The region of increased vascularity can be seen as hot spot. In breast thermal images, high intensity values are seen as hot spots. The third threshold level correctly extracts these hot spot area among all other thresholds.
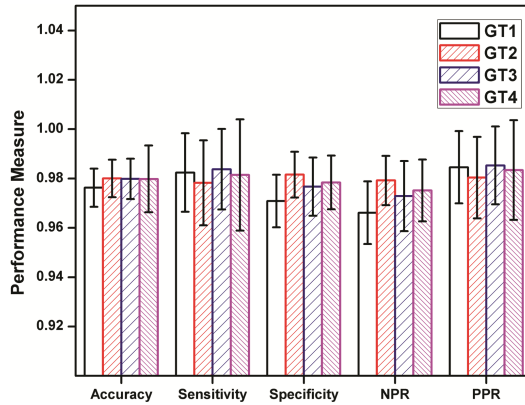
Fig. 3. Average variation of regional statistics based quantitative measures using basic levelset based segmentation framework

Table 1. TBD features values between normal and abnormal subjects for different threshold levels

	First threshold		Second threshold		Third threshold		Fourth threshold	
	N	A	N	A	N	A	N	A
FD	0.93	0.90	0.98	0.42	0.98	0.52	0.41	0.72
	0.55	0.24	0.57	0.23	0.75	0.32	0.53	0.70
	0.92	0.31	0.87	0.48	0.92	0.44	0.85	0.58
MGL	0.42	0.34	0.61	0.72	0.55	0.85	0.74	1.00
	0.23	0.32	0.36	0.37	0.01	0.07	0.08	0.03
	0.48	0.59	0.64	0.69	0.23	0.52	0.33	0.77
PC	0.29	0.32	0.33	0.24	0.44	0.37	0.06	0.16
	0.49	0.46	0.57	0.56	0.65	0.62	0.20	0.32
	0.29	0.31	0.29	0.33	0.40	0.43	0.25	0.23

Extraction of hot spot region is occurs only at third threshold level. Due to this property, the differentiation among normal and abnormal subjects is high in all three features such as fractal dimension, mean gray level and pixel counting at third threshold level. The fractal dimension value is high for normal subjects owing to the occurrence of homogenous patterns of breast tissues as compared to abnormal subjects. The normalized values of mean gray level is high for carcinoma subjects since thresholded blood vessel area (hot spot area) is high in abnormal regions compared to normal region. The mean value of pixel count is low for carcinoma subjects compared to normal subjects which may be owing to structural changes related in pathology condition.

Among these three features, mean gray level is able to significantly differentiate normal and different pathological conditions. The importance of asymmetry is explored by estimating the absolute difference between the features of normal and abnormal subjects are shown in Fig. 4. Among the extracted features, mean gray level at third threshold level are identified to be powerful for distinguish normal and carcinoma breast

tissues by 11%. Region for choosing four threshold levels here is maximum of 11% difference is examined between normal and abnormal breast tissues. The same operation is performed for lower and higher threshold values but it shows maximum of 9% differentiation among normal and abnormal subjects.
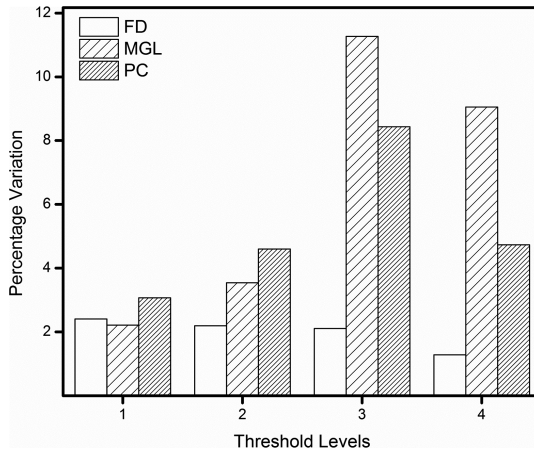
Fig. 4. Variation (%) in feature values between normal and abnormal subjects for different threshold levels

4   Conclusion

The breast tissues in thermal images are analyzed using a pipeline approach integrating new edge enhancement into the modified level set framework and asymmetry analysis using TBD method. The segmented results validated against ground truth images and are compared against the conventional level set, CSM and QBM methods. For asymmetry analyses, normal and all pathological conditions are considered. Features such as fractal dimension, mean gray level and pixel counting are extracted. The consequence shows the ability of the proposed method which results as proficient to delineate the breast tissues appropriately. On average, the segmented areas are found to be highly correlated (0.98) with the all ground truth images. It is observed that the TBD method shows the structural changes owing to diverse vascularized nature of the various tissues. Further, the mean gray level features shows 11% improvement in distinguishing normal and carcinoma tissues. Therefore, the proposed segmentation method and the extracted features emerge to be useful for improving the early identification of breast abnormalities.

References

1. NBCF: National Breast Cancer Foundation, Inc. (2010). http://www.nationalbreastcancer.org/about-breast-cancer/what-is-breast-cancer.aspx

2. Ng, E.Y.K.: A review of thermography as promising non-invasive detection modality for breast tumor. Int. J. Therm. Sci. **48**, 849–859 (2009). https://doi.org/10.1016/j.ijthermalsci.2008.06.015

3. Qi, H., Nicholas, A.: Infrared Imaging in Medicine (2007). ch. 1–10

4. Minikina, W., Dudzik, S.: Infrared Thermography, Error and Uncertainities, 1st edn. Wiley, Hoboken (2009). ch. 1–10

5. Keyserlingk, J.R., Ahlgren, P.D., Yu, E., Belliveau, N., Yassa, M.: Functional infrared imaging of the breast. IEEE Eng. Med. Biol. Mag. **19**, 30–41 (2000). https://doi.org/10.1109/51.844378

6. Zhou, Q., Li, Z., Aggarwal, J.K.: Boundary extraction in thermal images by edge map. In: Proceedings of the 2004 ACM Symposium on Applied computing, ACM, Cyprus (2004). https://doi.org/10.1145/967900.967956

7. Borchartt, T.B., Conci, A., Lima, R.C.F., Resmini, R., Sanchez, A.: Breast thermography from an image processing viewpoint: a survey. Signal Process. **93**, 2785–2803 (2013)

8. Diakides, N.A., Bronizino, J.D.: Medical Infrared Imaging. CRC Press Taylor and Francis group, Boca Raton (2007). ch. 11

9. Etehad Tavakol, M., Sadri, S., Ng, E.Y.K.: Application of K-and fuzzy c-means for color segmentation of thermal infrared breast images. J. Med. Syst. **34**(1), 35–42 (2010). https://doi.org/10.1007/s10916-008-9213-1

10. Huang, D.Y., Wang, C.H.: Optimal multi-level thresholding using a two-stage Otsu optimization approach. Pattern Recogn. Lett. **30**(3), 275–284 (2009). https://doi.org/10.1016/j.patrec.2004.10.003

11. Costa, A.F., Humpire-Mamani, G., Traina, A.J.M.: An efficient algorithm for fractal analysis of textures. In: IEEE Conference on Graphics, Patterns and Images (SIBGRAPI), pp. 39–46 (2012)

12. PROENG: image processing and image analyses applied to mastology (2012). http://visual.ic.uff.br/en/proeng/

13. Pisano, E.D., Zong, S., Hemminger, B.M., DeLuca, M., Johnston, R.E., Muller, K., Braeuning, M.P., Pizer, S.M.: Contrast limited adaptive histogram equalization image processing to improve the detection of simulated speculations in dense mammograms. J. Digit. Imaging **11**, 193–200 (1998). https://doi.org/10.1007/BF03178082

14. Kuo, H.-C., Giger, M.L., Reiser, I., Boone, J.M., Lindfors, K.K., Yang, K., Edwards, A.: Level set segmentation of breast masses in contrast-enhanced dedicated breast CT and evaluation of stopping criteria. J. Digit. Imaging **27**, 1–11 (2014). https://doi.org/10.1007/s10278-013-9652-1

15. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. **12**(7), 629–639 (1990). https://doi.org/10.1109/34.56205

16. Machado, D.A., Giraldi, G., Novotny, A.A., Marques, R.S., Conci, A.: Topological derivative applied to automatic segmentation of frontal breast thermograms. In: Proceedings of IX Workshop de VisaoComutational, FGV campus in Botafogo, Rio de Janeiro (2013)

An Empirical Evaluation of Savitzky-Golay (SG) Filter for Denoising ST Segment

C. K. Roopa[✉] and B. S. Harish

JSS Research Foundation, Sri Jayachamarajendra College of Engineering,
Mysore 570 006, Karnataka, India
{ckr,bsharish}@sjce.ac.in

Abstract. Electrocardiogram (ECG) is the graphical illustration of the heart's electrical activity and it is acquired through linking electrodes to body surface. It is useful in determining cardiac diseases non-invasively. The general concerns in ECG signal processing is the restrain of noise. This paper has been encouraged by the necessary to recognize the ability of Savitzky-Golay (SG) filter for denoising the ST segment of ECG signals. It essentially implicates the extraction of essential cardiac components by refusing the background noise with the aid of filtering technique. The simulation is carried out using MATLAB and the experiments are performed on physionet database. Peak Signal-to-Noise Ratio (PSNR), Signal-to-Noise Ratio (SNR), Mean Square Error (MSE) and Percentage Root mean square Difference (PRD) are used as performance evaluation metrics. This work uses SG filter to denoise the ECG signal and comparison is provided for SG filter with Median, Butterworth and SOS filters. The comparison results states the performance of the SG filter is better than that of the other filters.

Keywords: ST segment · Denoising · ECG signal · Filtering
Cardiac diseases

1 Introduction

Cardiovascular Disease (CVD) continued to be the major reasons of death worldwide. The early detection of CVD may considerably reduce the time of dealing in hospital. As a result, the patient condition will be improved with good examination output and maintains the costs of treatment below control [1]. According to the study conducted by Registrar General of India (RGI) and Indian Council of Medical Research (ICMR) in 2010, increased numbers of deaths are caused due to the heart diseases. In 2014, out of all deaths caused due to different types of diseases, 23 million were affected with cardiovascular diseases.

Nowadays most of the demise in worldwide transpires due to the unanticipated heart attack. To avoid this, the likelihood of attack should be predicted in the initial stage. This can be done by performing the electrocardiogram (ECG) test [2]. ECG signals are the representation of heartbeat values. It is a useful tool to study functional and structural status of the heart. In recent years, an automated method is used for the examination of ECG signals by means of concurrent processing. It is exceptionally necessary for the analysis of various cardiac ailments precisely. The entire ECG wave is an outline illustrated in Fig. 1 [3].
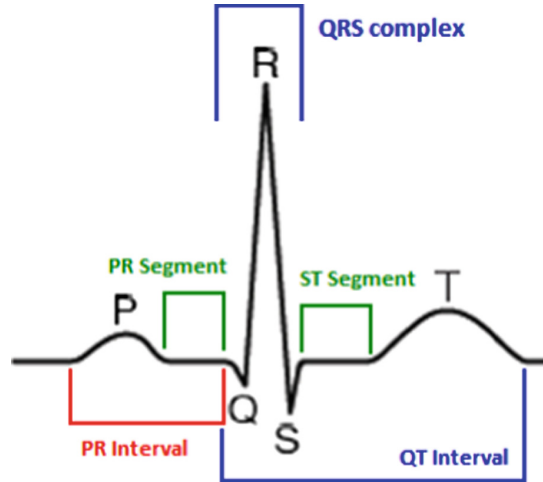
Fig. 1. Schematic representation of ECG

The signals are produced when heart muscles depolarize in reaction to electrical impulses produced by pacemaker unit. Leading to depolarization, the muscles contract and thrust the blood all through the body. The ECG exposes various things concerning the heart, including its beat, whether their electrical transmission paths are undamaged, whether certain chambers are distended, and even the estimated ischemic spot in the occasion of a heart attack (Myocardial Infarction). ECG is employed clinically in identifying a variety of irregularities and circumstances linked with the heart [4, 15]. Due to the high mortality rate of heart disease, it is difficult to accurately recognize myocardial infarction so as to support effective treatment. The ST segment Myocardial Infarction (STEMI) is the severe type and the eminent one. ST segment at the ECG information represents the massive quantity of coronary heart muscle disorder [5]. The ST segment is extracted and measured to recognize various cardiac diseases. Before recognizing, we need to preprocess the signal for the different types of noise exclusion. The recording of cardiac events which changes during cardiac cycle is P wave, QRS complex and T wave. Here, depolarization of the atria is depicted in P wave, depolarization of the ventricles is characterized in QRS complex and T wave is the repolarization of the ventricles [6]. ECG reports the unusual electrical potentials produced by the heart.

ECG machines utilize various machine learning procedures for diverse calculations and distinctive methodologies to examine exact outcomes. The detection of cardiac arrhythmias and myocardial infarction using ECG signals are supported with three main steps: preprocessing, feature extraction and classification. ECG signal consists of various kinds of noises such as baseline drift, electrode contact noise, muscle noise, frequency interference, electrosurgical noise, polarization noise, motion artifacts and internal amplifier noise [7]. These interferences are analyzed with various filters along with the PSNR, SNR, MSE and PRD values. A noisy ECG may hamper the doctor's accurate assessment on patients. In this means, denoising of ECG happen to be a

restrictive prerequisite [8]. To achieve a distortion less, precise and fault free signals, filtering method is employed by means of filters. The main aim of this filtering process is to eliminate the noise and do the further analysis. In this paper, we have used four filters for noise removal: Savitzky-Golay filter (SGF), Butterworth filter (BWF), Median filters (MF) and Second Order Section filter (SOS) [9].

The main objective of this paper is towards exploiting SG filter to pre-process the noise present in the ST segment dataset. Dataset is been used from physio-net data bank of ECG signal. SG filter is a digital filter, and it mainly increases the SNR excluding distortion in the signal. The make use of the SG filter is assessed and compared against various filters. The quantitative outcome specifies that the SG filter can achieve consistently better even when the signal is impure by numerous types of intervention. The remainder of the article is structured as follows: existing methods are explained in Sect. 2. Various filters and its explanation are given in Sect. 3. The results as well as the performance comparison of the various filtering techniques are presented in Sect. 4. Conclusion along with future work is drawn in Sect. 5.

2   Existing Methods

This section provides an overview of various filtering techniques for the process of de-noising the ECG signal. Several filters have been presented in the last decade, but there were considerable differences in each with respect to the obtained result and parameter values. The very recent works related to filtering of ECG Signals is presented below.

Sharma and Sunkaria in [10] discussed about the detection of QRS complex from ECG wave. The detection of QRS complex is a very important stage than identifying any additional signal element, beat fragments, or additional morphological aspects of ECG. Further, for the recognition of QRS complex, a method with least pre-processing was presented. The method employed two-stage filtering method for pre-processing by means of median filter plus SG smoothing filter. The assessment was supported using RMS value of the wave. Experiments were conducted on FTD, MIT-AD, MIT-NSD and BIDMC – CHFD datasets [10] for the course of identifying cardiac health prognosis. Arif et al. in [11] discussed about the automatic identification and localization of myocardial infarction (MI). The major step in processing the ECG signal is pre-processing. Before pre-processing the signal detection and delineation of QRS complex was performed and after that wavelet transform was applied to de-noise the signal. Further, the existence or nonexistence of myocardial infarction was predictable by the mining of Q wave, T wave and ST level elevation from 12 lead ECG signal. Finally, K-NN classifier was used for the course of classification. Here the experiments were carried out on PTB database available on Physio-bank.

Ventricular Late Potentials (VLP) classification technique was proposed by Nasario-Junior *et al.* [12]. The proposed classification technique was based on Principal Components Analysis (PCA). VLP consists of low amplitude and high frequency transients recorded on high resolution electrocardiogram (HRECG) and they were the markers of life threatening ventricular tachyarrhythmia. In the experiment section, the database of 36 patients signal was divided into two groups namely 18 healthy controls

and 18 patients with induced sustained monomorphic ventricular tachycardia was considered. Mahalanobis distance measure was used for the classification by logistic regression which is determined by optimal separation threshold between groups for each matrix. HRECG signals were also analyzed using many classical approaches.

A dynamic single-lead ECG demarcation system supported with wavelet transform (WT) was proposed by Martínez *et al.* [13]. Primarily, QRS complexes were identified. Later on, every QRS was demarcated by identifying and recognizing the peaks of every wave, in addition the complex onset and end. They were P, Q, R, S, R' and T signals, and also the P, QRS and T signal borders. These waves were identified by the employment of single lead investigation phase known as the DWT of the ECG. Finally, the purpose of P and T signal peaks, onsets and ends were achieved. To corroborate the efficacy, the algorithm was assessed on numerous manually interpreted databases. Cuomo *et al.* in [14] illustrated an improved method for ECG signal denoising supported by Recursive Filtering (RF) process. This process can be applied in the context of concurrent health monitoring, where signal dispensation was decisively insisted. A suitable set of kernel functions were used to remove the rarity in ECG, including noise frequencies in the Fourier domain. These techniques do not order high computational necessities and this feature offered the opportunity of realization of the technique explicitly on portable instruments. In this work the experiments were conducted using the facts from Physio-net LT – ST Database. Therefore, pre-processing stage was the major stage for the elimination of the extra artifacts on the ECGs. As a result, this scheme used local denoising technique for ECG signals.

All in all, many filtering techniques were proposed in the literature towards denoising the ST segment of ECG signal. However, to the best of our understanding no work has been reported in the literature towards exploiting SG filter on ST segment for localization of Myocardial Infarction. Hence in this paper, we are making an initial attempt of employing SG filter to preprocess ST segment to denoise the signal. Further, the proposed denoising technique will also helps in improving the rate of detection and classification of different heart diseases.

3   Methodology

Normally, the existence of noise corrupts the signal and make the feature extraction and classification process less accurate. Thus, to overcome this process we are employing preprocessing techniques to remove noise from input ECG signal. In this paper, we explored the usage of various filters for denoising the ECG signal. Filters such as Savitzky-Golay filter, Butterworth filter, Median filter and Second Order Section filter are used in this work. The working of individual filter with respect to ST segment of ECG signal is presented below.

3.1   Savitzky-Golay (SG) Filter

ECG voltages measured across the body are on the order of hundreds of micro volts up to one milli volt. In practice during ECG measurement some noises gets add up along with the ECG signal and affect its lucidity. Some of the source of noise in ECG is

Power line intervention, Electrode contact noise, baseline drift and motion artifacts, EMG from the chest wall, Instrumentation noise and electrosurgical noise.

In a Savitzky Golay filter (SGF) three basic stages are used: Moving Average Method, Convolution and Least Square method. The output is the replacement of sequence that is rounded similar to simple moving average method. Convolution method is used to obtain a weighted moving average by the way of weighting specified as a polynomial of definite quantity of the signal and Least Square method minimize the squared error on all points, which determine the gap between the original points and the smoothened points. The weighted coordinates while related to the ECG signal attain a polynomial least-squares fit included by the filter window. The polynomial is projected to preserve better moments within the facts and to reduce the partiality instigated by the filter. The formula of the vital Least-Square Convolution for time-series smoothing is as follows:

$$A_i = \frac{1}{N} \sum_{i=-m}^{i=m} C_i Y_{j+i} \tag{1}$$

Where, A_i is the resulting filtered ECG signal, C_i is the coordinates for the ith value of the smoothing window, and N is the quantity of convoluting numerals and is equal to the smoothing window size (2m + 1). The key j is the consecutive key of the new ordinate facts.

The SGF impulse response can be extracted directly as the lowest row of the inverse matrix; hence a conclusion is made to fit a fourth-order polynomial which is shown in Eq. (2) with five samples.

$$y = ax^4 + bx^3 + cx^2 + dx + e \tag{2}$$

The general form of impulse response is given by,

$$y = Ma \tag{3}$$

Where a is the vector of polynomial coefficients, and M is the matrix form of the ECG signal.

The general form of Eq. (1) can be rewritten according to the specification as follows in Eq. (4),

$$
\begin{bmatrix} y_1 \\ . \\ . \\ . \\ y_5 \end{bmatrix} =
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 \\
32 & 16 & 8 & 4 & 2 & 1 \\
243 & 81 & 27 & 9 & 3 & 1 \\
1024 & 256 & 64 & 16 & 4 & 1 \\
3125 & 625 & 125 & 25 & 5 & 1
\end{bmatrix}
\begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix}
\tag{4}
$$

Here the five possible equations with six unknown variables are represented in matrix form, where the numerical values in the matrix M represent the least squares of

the constraining points. To find the polynomial coefficients, Eq. (3) can be written in the form as in Eq. (5),

$$a = M^{-1}y \tag{5}$$

From Eq. (3) the inverse matrix is given as,

$$
\begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} =
\begin{bmatrix}
-0.0013 & 0.0035 & -0.0013 & -0.0029 & 0.0000 \\
0.0361 & -0.0943 & 0.0288 & 0.0747 & 0.0052 \\
-0.3875 & 0.9372 & -0.2191 & -0.7067 & -0.1049 \\
1.9592 & -4.2181 & 0.6064 & 2.9346 & 0.6789 \\
-4.6159 & 8.1636 & -0.3111 & -5.0311 & -1.5443 \\
4.0000 & -4.7500 & -0.1667 & 2.7500 & 1.0000
\end{bmatrix}
\begin{bmatrix} y_1 \\ . \\ . \\ . \\ y_5 \end{bmatrix} \tag{6}
$$

By solving the above Eq. (6) the exact assessment of a, b, c, d and e can be attained. Once when the estimation of the polynomial coefficients is made, the actual impulse response can be computed and the same process is continued for the whole ECG signal. Thus the signal is de-noised for increasing the system classification accuracy.

3.2 Butterworth Filter

To sort the exact ECG signal, the identification circuit has to exist competent of lessening the distorted signal as well as out-of-band interfered signals. For that reason, the power of the filter with the cut-off occurrence has to be carefully selected to guarantee maximum reduction of the undesired interference. The inevitability for ECG identification system outcome in escalating the filter constraint such as elevated order, small total harmonic distortion (<-50 dB), small contribution referred noise concentration (<400 µVrms/Hz), dynamic range (>60 dB) and low power consumption (<50 µW). Therefore, a second order Butterworth filter through a highest even response by ladder topology is used for this. A simple butterworth low pass filter is designed using,

$$[b, a] = butter(n, w_n) \tag{7}$$

Where, n is the order of filter, w_n is the cutoff frequency and both b and a are the filter coefficients of length $n + 1$. In our experiment, 2nd order butterworth is employed and the equation for 2nd order butterworth filter is rewritten as follows.

$$[b, a] = butter\left(2, \frac{fc}{fs/2}\right) \tag{8}$$

Where fc is the cutoff frequency and fs is the sampled data.

3.3   Median Filter

The median filter is a better technique to differentiate out-of-range isolated noise from legitimate signal. It is a non-linear filtering method, normally applied to eradicate distortion from a signal. Reduction in distortion is a general pre-preparing scheme to improve the penalty of afterward analysis. The main idea of median filter is to substitute each wave entry with the median of neighboring entry. The occurrence of neighboring is called the "window", which moves, item wise, all through the entire signal. In support of signals, the most obvious window is just the first few former and following entries are used. If the window has an odd variety of entries, then the median is straightforward to define. After calculating median all the entries in the window are sorted numerically.

The yielded trial of linear distinct filter contains linear mixture of preceding yielded trial $y(n - i)$, i represent numerical series from 1 to N, and input trial $x(n - k)$, k represent numerical series from 0 to M. The abovementioned linear combination is in case of nonlinear Discrete Dynamic Systems (DDS) substituted by means of nonlinear function with appropriate count of variables.

$$y(n) = F\{x(n), x(n - 1), \ldots x(n - M), y(n - 1), y(n - 2), \ldots y(n - N)\} \quad (9)$$

Median filter symbolize nonlinear dynamic method obtained from vector of values $x(n) = (x(n), x(n - 1), \ldots x(n - M))^T$. The yielded trial of median filter y(n) is distinct as center trial from classified compilation of constituents of vector x(n).

$$y(n) = med\{x(i)\} \quad (10)$$

Where, $i = n, n - 1, \ldots n - M$.

A quantified trial of $M + 1$ median filter is regularly termed length D of it; generally it is articulated through odd number. Length of median filter determines suppression of impulse errors with various widths.

3.4   Second Order Section (SOS) Filter

SOS represents the second-order section digital filter $H(z)$ and the equation for SOS is derived and it is given below.

$$H(z) = \prod_{k=1}^{L} H_k(z) = \prod_{k=1}^{L} \frac{b_{0k} + b_{1k}z^{-1} + b_{2k}z^{-2}}{1 + a_{1k}z^{-1} + a_{2k}z^{-2}} \quad (11)$$

Where z is a composite variable, k is the number of segments, a is the set of repeal coefficients, and b is the set of onward coefficients. The coefficients are frequently regularized to facilitate $a_0 = 1$.

During the process of ECG signal, a digital biquad filter is a 2nd order repeated linear filter, which comprises of 2 poles and 2 zeros. It affirms the element to facilitate in the region; its shift function is the quantity of 2 quadratic functions:

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{a_0 + a_1 z^{-1} + a_2 z^{-2}} \qquad (12)$$

High-order IIR filters be able to enormously vulnerable to quantization of their coefficients as well as can simply turn out to be unbalanced. This is significantly fewer of a matter with 1st and 2nd order filters; as a result, superior order filters will be normally realized as successively surged biquad areas.

4   Results

The evaluation of the presented technique is assessed on ST-segment data samples taken from the Physio-net data bank. The dataset consists of huge number of ECG recordings. It was selected to display a range of measures of ST segment modification, including ischemic ST episodes, axis-related non-ischemic ST episodes, episodes of slow ST level drift, and events together with mixture of these occurrences. The signals in the ST-segment Physio-net database comprises number of ECG signals with known pathology. To estimate the de-noising methods, experiments were conducted on two category of noisy ECG signal. First, ECG signals were randomly selected from Physio-net MIT-BIH Database which contains internal noise and secondly on the original ECG signal with artificial noise produced by adding White Gaussian Noise (AWGN). These distorted ECG waves will be de-noised mainly through SG filter and three other filtering methods.

4.1   Performance Comparison

The performance of various filters such as SG, Butterworth, Median and SOS filters on both categories (original signal and additive noise signal) of data sample were compared in terms of PSNR, SNR, MSE and PRD values. All the four performance evaluation metrics are defined as below:

Mean Square Error (MSE)
The value of MSE for any system must always be low. The equation for MSE is given below

$$MSE = \frac{\left(\sum (f - f_p)^2 \right)}{N} \qquad (13)$$

Where, f is the new signal, f_p is the restructured signal and N is the total signal length.

Peak Signal-to-Noise Ratio (PSNR)
The value of PSNR for any system must be always high. The equation for PSNR is given by

$$PSNR = 10 \log_{10}\left(\frac{M^2}{MSE}\right) \tag{14}$$

Where, M is the maximum value from original ECG signal.

Root mean square Difference (PRD)
The value of PRD for any system must always be low. The equation of PRD is given by

$$PRD = \sqrt{\frac{MSE}{\sum f^2}} \times 100 \tag{15}$$

Where, MSE is mean square error of the signal.

Signal-to-Noise Ratio (SNR)
The value of SNR for any system must be always high. The equation of SNR is given by

$$SNR = \frac{RMS(signal)}{RMS(noise)} \tag{16}$$

Where, RMS stands for Root Mean Square.

The above selected parameters gives the least MSE and the least PRD for the selected dataset. For contrast, we have evaluated SG filter with the median filter, SOS filter, butterworth filter. The above mentioned filters are evaluated using 4 parameter values (MSE, PSNR, SNR and PRD). The behaviors of the diverse filtering methods on different ST segments are given in Table 1. Note that the SG filter outperforms other filters for all the four performance measures. The biomedical waves include a many investigative characteristics which contain to be visually examined through the cardiologists. Therefore, it is of major significance to conserve those characteristics. The significant investigative characteristics of an ECG signals are P-waves, QRS complexes, T-waves, U-waves, and their intervals. During the exclusion of noise, these characteristics typically experience smoothing, consequently disturbing the investigative correctness.

Table 1. Comparison of MSE, PSNR, PRD and SNR values of various filters

Filters	Metrics							
	MSE		PSNR		PRD		SNR	
	Original ECG signal	ECG + Noise	Original ECG signal	ECG + Noise	Original ECG signal	ECG + Noise	Original ECG signal	ECG + Noise
SG	3.180	6.517	24.639	27.553	0.047	0.038	0.059	0.052
Butterworth	6.086	12.360	21.859	24.735	0.064	0.053	0.057	0.052
SOS	12.508	12.959	22.382	22.606	0.061	0.077	0.102	0.096
Median	12.920	25.123	18.779	21.466	0.092	0.078	−2.463	−2.568

The Fig. 2(a) and (b) demonstrates the original ECG signal and a signal added through additive white noise respectively. The results obtained after preprocessing the signal by using SG filter is portrayed in Fig. 3(a) and (b).
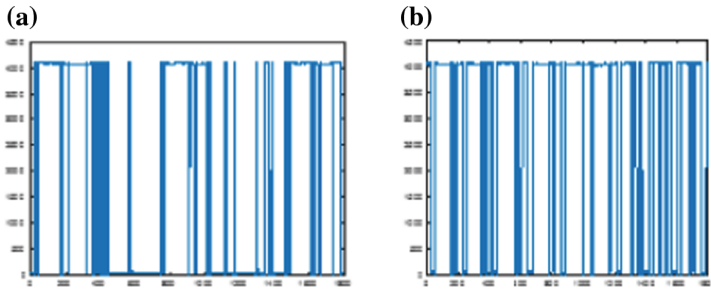
Fig. 2. (a) Original ECG signal. (b) Signal added with white noise
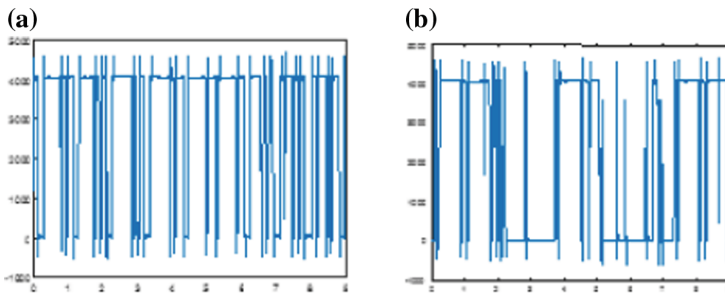
Fig. 3. (a) and (b) Denoised signal using SG filter

5   Conclusion

This paper elucidated the effect of SG filter for pre-processing of ECG signal. The main contribution of this paper is to use the four different filters viz: SG, Butterworth, SOS and median filter to preprocess the ECG signal. Further, all four filters are evaluated using four different evaluation metrics. The obtained results on both original signal and additive noise signal are compared for its evaluation. From the obtained results it was found that SG filter obtained efficient result when compared with other filtering techniques. It has some desirable characteristics such as high SNR and PSNR and low MSE and PRD values that are absent in other filters. The implementation was performed with physio-net database. The performance assessment of proposed filter on original signal resulted with low MSE value of 3.180, low PRD value of 0.047, high PSNR value of 24.639 and high SNR value of 0.059. For additive noise signal the proposed system resulted in low MSE value of 6.517, low PRD value of 0.038, high PSNR value of 27.553 and high SNR value of 0.052. These results proved that the use of SG filter in

preprocessing ST segment of ECG signal is efficient. In future, we have indented to work on the preprocessed ST segment to extract good set of features for accurate detection and classification of different Cardiovascular Disease.

References

1. Ripoll, V.J., Wojdel, A., Romero, E., Ramos, P., Brugada, J.: ECG assessment based on neural networks with pretraining. Appl. Soft Comput. **49**, 399–406 (2016)
2. Salas-Boni, R., Bai, Y., Harris, P.R., Drew, B.J., Hu, X.: False ventricular tachycardia alarm suppression in the ICU based on the discrete wavelet transform in the ECG signal. J. Electrocardiol. **47**, 775–780 (2014)
3. Varanini, M., Tartarisco, G., Balocchi, R., Macerata, A., Pioggia, G., Billeci, L.: A new method for QRS complex detection in multichannel ECG: application to self-monitoring of fetal health. Comput. Biol. Med. **85**, 125–134 (2017)
4. Zidelmal, Z., Amirou, A., Ould-Abdeslam, D., Moukadem, A., Dieterlen, A.: QRS detection using S-Transform and Shannon energy. Comput. Methods Programs Biomed. **116**, 1–9 (2014)
5. Talbi, M.L., Ravier, P.: Detection of PVC in ECG signals using fractional linear prediction. Biomed. Signal Process. Control **23**, 42–51 (2016)
6. Mjahad, A., Rosado-Muñoz, A., Bataller-Mompeán, M., Francés-Víllora, J.V., Guerrero-Martínez, J.F.: Ventricular fibrillation and tachycardia detection from surface ECG using time-frequency representation images as input dataset for machine learning. Comput. Methods Programs Biomed. **41**, 119–127 (2017)
7. Al Rahhal, M.M., Bazi, Y., AlHichri, H., Alajlan, N., Melgani, F., Yager, R.R.: Deep learning approach for active classification of electrocardiogram signals. Inf. Sci. **345**, 340–354 (2016)
8. Awal, M.A., Mostafa, S.S., Ahmad, M., Rashid, M.A.: An adaptive level dependent wavelet thresholding for ECG denoising. Biocybern. Biomed. Eng. **34**, 238–249 (2014)
9. Zidelmal, Z., Amirou, A., Ould-Abdeslam, D., Merckle, J.: ECG beat classification using a cost sensitive classifier. Comput. Methods Programs Biomed. **111**, 570–577 (2013)
10. Sharma, L.D., Sunkaria, R.K.: A robust QRS detection using novel pre-processing techniques and kurtosis based enhanced efficiency. Measurement **87**, 194–204 (2016)
11. Arif, M., Malagore, I.A., Afsar, F.A.: Detection and localization of myocardial infarction using K-nearest neighbor classifier. J. Med. Syst. **36**, 279–289 (2012)
12. Nasario-Junior, O., Benchimol-Barbosa, P.R., Nadal, J.: Principal component analysis in high resolution electrocardiogram for risk stratification of sustained monomorphic ventricular tachycardia. Biomed. Signal Process. Control **10**, 275–280 (2014)
13. Martínez, J.P., Almeida, R., Olmos, S., Rocha, A.P., Laguna, P.: A wavelet-based ECG delineator evaluation on standard databases. IEEE Trans. Biomed. Eng. **51**, 570–581 (2004)
14. Cuomo, S., De Pietro, G., Farina, R., Galletti, A., Sannino, G.: A revised scheme for real time ECG signal denoising based on recursive filtering. Biomed. Signal Process. Control **27**, 134–144 (2016)
15. Roopa, C.K., Harish, B.S.: A survey on various machine learning approaches for ECG analysis. Int. J. Comput. Appl. **163**, 25–33 (2017)

Detection of Exudates Through Local Binary Pattern in Diabetic Retinopathy

R. Suma[(✉)] ⓘ, Deepashree Devaraj ⓘ, and S. C. Prasanna Kumar ⓘ

R.V. College of Engineering, Mysuru Road, Bengaluru 560059, India
sumastl4@gmail.com,
{deepashree,prasannakumar}@rvce.edu.in

Abstract. The long term diabetes leads to the retinal vascular disease called diabetic retinopathy (DR). As DR is a progressive disease, it should be diagnosed and treated as soon as possible to prevent the patient from blindness. The lesions like microaneurysms, exudates, hemorrhages and abnormal growth of blood vessels etc. will appear in DR. In this paper, the segmentation and classification of exudates is proposed. Exudates are segmented using modified morphological operation which deals with the intervention of optic disc before the detection of exudates. The feature extraction is done using local binary pattern (LBP) and classified using support vector machine (SVM). The proposed method is tested on DIARETDB0 and DIARETDB1, which are freely available datasets. The method is also evaluated on the hospital images. The average sensitivity, specificity and accuracy of 0.92, 0.82, and 0.95 respectively are obtained.

Keywords: Diabetic retinopathy · Microaneurysms · Exudates
LBP · SVM · DIARETDB0 · DIARETDB1

1 Introduction

Diabetes affects the retinal blood vessels and causes the eye disease called DR. So the diabetic patients need to attend regular eye screening tests [1]. According to WHO, the diabetic cases will rise to 366 million and DR cases to 285 million by 2030 [2]. The main symptoms are floaters, flashes, blurred vision and loss of sight. Therefore, detection of the disease by making use of automated system is of great importance to prevent the vision loss. This also helps the ophthalmologists to treat the disease effectively and thereby reduce the number of cases going undiagnosed [1]. The view of normal vision and DR vision is shown in Fig. 1.

Blood vessels, optic disc (OD) and macula are the main components of healthy retina. The changes to these components of retina indicate the sign of the disease [1]. Basically DR has two stages namely non-proliferative DR and proliferative DR. Non-proliferative DR is the early stage of the disease whereas proliferative DR is the advanced stage of the disease [3]. The image of healthy and abnormal retina is shown in Fig. 2 and image of NPDR and PDR is shown in Fig. 3.

(a)   **(b)**

Fig. 1. (a) Normal, (b) Diabetic retinopathy vision [4]

(a)   **(b)**
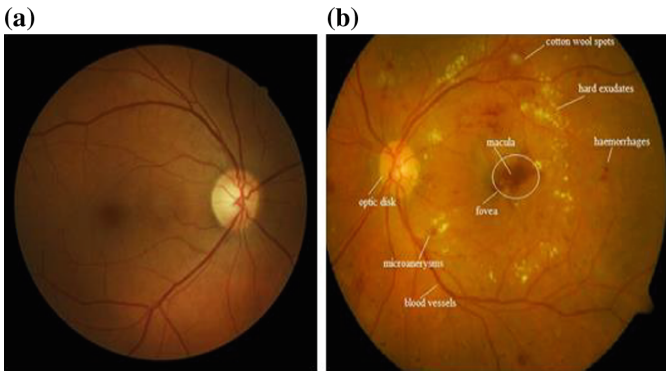
Fig. 2. Retinal image (a) Normal, (b) DR [4] (Color figure online)

The retinal lesions are microaneurysms, exudates and hemorrhages. Exudates appear as a bright yellowish patches on retina caused due to exuding of fluid from the tissue because of its injured capillaries [5]. These are the bright lesions of retina. The exudates and optic disc are of the same intensity. So, segmenting only the exudates region excluding optic disc is a challenging part.

Saha et al. [3] proposed, segmentation of both the dark and bright lesion using Fuzzy C means clustering. The bright lesions are classified by Naïve Bayes classifier. The accuracy of bright lesions was 88.88%, however it could be improved by using uniformly illuminated fundus images. Harangi and Hajdu [5] proposed detection of exudates using candidate extraction, precise contour segmentation and classification automatically. Candidate extraction is done using grayscale morphology, precise boundary segmentation is done using active contour method and classification using region wise classifier.
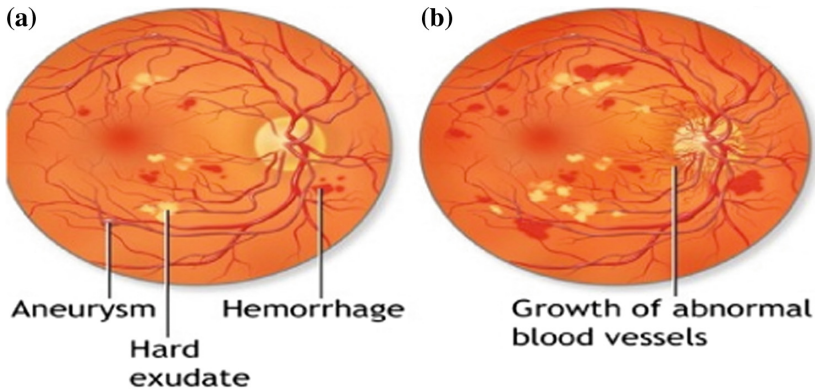
Fig. 3. (a) Non proliferative DR, (b) Proliferative DR [8]

Lachure et al. [6] proposed detection of both microaneurysms and exudates by morphological operators, features are extracted using GLCM and classified using SVM and KNN classifier. Specificity is 100% and sensitivity is more than 90% for SVM. Sánchez et al. [7] proposed automatic exudates detection based on mixture model to dynamic thresholds. Edge detection was used to distinguish exudates from other lesions. But this method neglects to identify non exudates.

Osareh et al. [9] proposed preprocessing using color normalization and contrast enhancement and segmented using fuzzy c means. To rank the features, genetic algorithm is used and classified using multilayer neural system. de la Calleja et al. [10] proposed a technique in which features are extracted using LBP and classified using ANN, random forest and SVM classifiers.

Krishnan et al. [11] have proposed the identification of normal and DR images by extracting the features from LBP, entropy and invariant moments, to calculate the diabetic retinopathy index, which helps to screen the disease. Since, this index is highly discriminative, it can reduce the burden of the ophthalmologists in identifying normal retinal images.

The method for detection of exudates, which occurs in non-proliferative DR stage through LBP, which is a texture descriptor is proposed in this paper. The features are classified using SVM classifier which provides an efficient result when compared to the other methods.

2 The Proposed Method

DR is major cause for blindness in diabetics. Screening diabetic patients for early diagnosis of DR can reduce the vision loss. Exudates are the early visible signs of DR. So, automatic and early detection of exudates are helpful to detect and treat the disease in early stage.

In this paper, exudates detection using LBP is proposed. The block diagram is shown in Fig. 4. It includes five main steps namely, the acquisition of retinal images, pre-processing operations, segmentation techniques, feature extraction and classification.

The retinal images used in the proposed method are acquired from standard retinal databases like DIARETDB0 and DIARETDB1. DIARETDB0 has 20 normal and 110 DR images and total of 130 color fundus images. DIARETDB1 has 89 images with 5 normal images and 84 DR images out of which 47 have signs of exudates. The field of vision of 50° was used to capture the image, it has a resolution of 1500 × 1152 and ZIESS FF 450plus retinal camera with Nikon F5 digital camera is used, which gives the image in PNG format [1]. Hospital images uses Optomed camera with 1920 × 1440 with 40° field of view, all the images are of JPEG format.
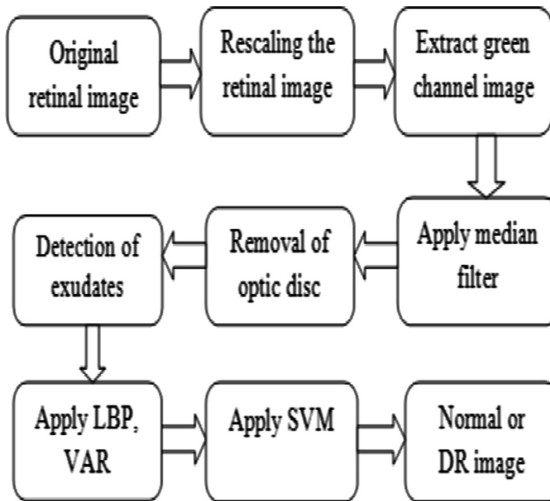
Fig. 4. Block diagram of the proposed method.

2.1   Preprocessing

The first step in preprocessing is rescaling the original image. As the images used in the work are obtained from different databases, which are of different sizes so, bicubic interpolation is used for resizing of the image [12]. The images are rescaled to 512 × 512. After rescaling, for noise removal, median filtering of 3 × 3 neighborhood is done. Later the green channel of image are extracted, as it has good contrast with the background and it contains most of the useful information when compared to other channel components of image. The input retinal image and green channel image are shown in Fig. 5.

(a) **(b)**
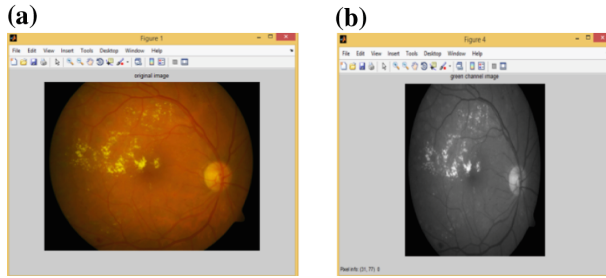
Fig. 5. (a) Input retinal image and (b) Green channel image

2.2 Segmentation of Exudates

As already mentioned, the exudates and optic disc are of the same intensity, their removal in the retinal image is a crucial step, so the techniques used to do so are the masking, thresholding, connected component analysis and morphological opening operation. Thus, optic disc is removed and exudates are detected. The image of optic disc with exudates and optic disc removed image are shown in Fig. 6.

2.3 Feature Extraction

The features extracted provide the necessary information required for the classification from segmented image. With the help of image texture, the abnormal lesions in the retina can be easily identified. The textural properties are smoothness, roughness and regularity of any structure etc. [12].

Local Binary Pattern
The texture of the retina is characterized by making use of LBP and VAR operators. The LBP is used for feature extraction because of its simplicity, less computational complexity and it is also a grayscale invariant. The green channel component of the retinal image is analyzed. The LBP and VAR values for sampling point P = 8 and radius of the circle R = {3, 5} are calculated. The resulting values provide the best description of the texture of the image [12].

(a) **(b)**
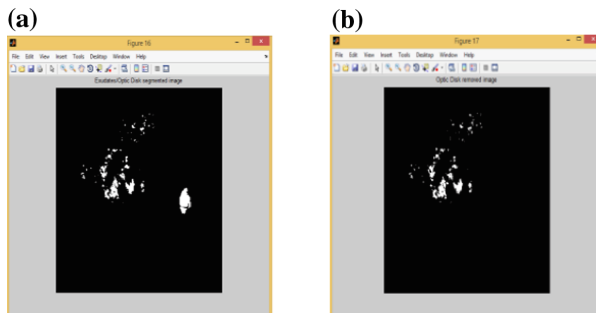
Fig. 6. (a) Exudates with optic disc image, (b) Exudates without optic disc image

The neighboring pixels of the centre pixels are thresholded to get the binary pattern. The LBP value is obtained by adding the binary pattern with the powers of two as shown in Eqs. 1 and 2.

$$\text{LBP}_{P,R} = \sum_{P=0}^{P-1} s\left(g_p - g_c\right)2^P \tag{1}$$

$$s(x) = \begin{array}{l} 1 \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 \end{array} \tag{2}$$

where, g_c and g_p are the gray levels of central and neighbourhood pixels respectively. The gray values for the circular neighbourhood of $P = 8$ and $R = 1$ and thresholding the pixel values is shown in the Fig. 7. The rotational invariant local variance is given in Eqs. 3 and 4.

$$\text{VAR}_{P,R} = \frac{1}{P}\sum_{p=0}^{P-1}\left(g_p - \mu\right)^2 \tag{3}$$

$$\mu = \frac{1}{P}\sum_{p=0}^{P-1} g_p \tag{4}$$

where, μ is mean value of neighboring pixel.
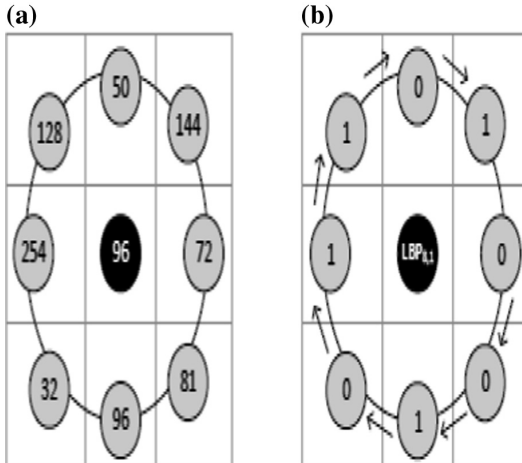
Fig. 7. (a) Gray values for the circular neighbourhood of $P = 8$ and $R = 1$, (b) Thresholding of neighbouring pixels based on centre pixel [12]

The statistical values calculated are mean, median, standard deviation, area and entropy. Mean, median and standard deviation values are calculated from the green channel pixels of the lesion region. Area is total number of pixels occupied by the required lesion. Entropy feature gives a statistical measure of randomness of the retinal image [13]. The LBP features extracted image and VAR image of the exudates is shown in the Fig. 8.

(a)   **(b)**
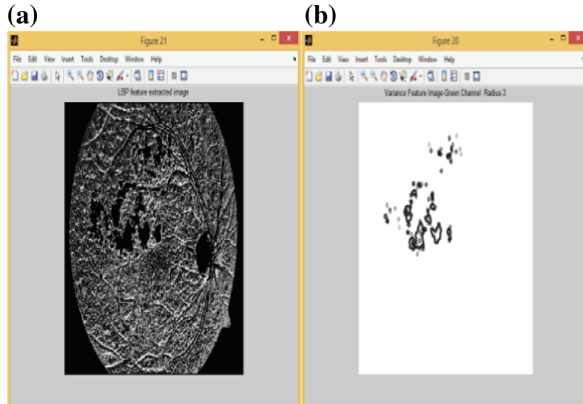
Fig. 8. (a) LBP features extracted image, (b) VAR image

2.4 Classification

Compared to other classifiers, SVM classifiers have established outstanding performance in variety of pattern recognition problems in terms of accuracy. The concept and performance of linear SVM is explained in our previous paper [14]. SVM is a supervised learning process used to analyze the training data and to classify into normal or DR images. SVM constructs hyperplane to separate the given data into classes linearly. The retinal images are classified based on the features obtained from the feature extraction technique [15].

3 Results

The proposed method was evaluated on the standard diabetic retinopathy databases and hospital images. Matlab R2013a was used to implement the proposed algorithm. 50 images were trained and more than 100 images were tested with the algorithm. Sensitivity, specificity and accuracy were measured for performance comparison. The percentage of true positives is given by sensitivity, Specificity gives measure of negative cases that were classified correctly and Accuracy is success rate [3]. The sensitivity, specificity and accuracy are computed as,

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{5}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{6}$$

$$\text{Accuracy} = \frac{TP}{TP + TN + FP + FN} \tag{7}$$

where, TP (true positive) is exudates pixels correctly detected, TN (true negative) is non-exudates pixels that are correctly detected as non- exudates, FP (false positive) is non-exudates pixels that are wrongly detected as exudates and FN (false negative) is exudates pixel that are not detected as exudates [16]. The sensitivity, specificity and accuracy obtained are 0.92, 0.82, and 0.95 respectively. The comparison of the proposed method with the other exudates segmentation algorithms are shown in Fig. 9. From Fig. 9, it can be justified that the proposed method shows good performance when compared with the other methods. The time complexity and accuracy of the algorithm for different databases and hospital images are shown in Table 1 and the segmented, LBP and VAR results of few retinal images are shown in Table 2.
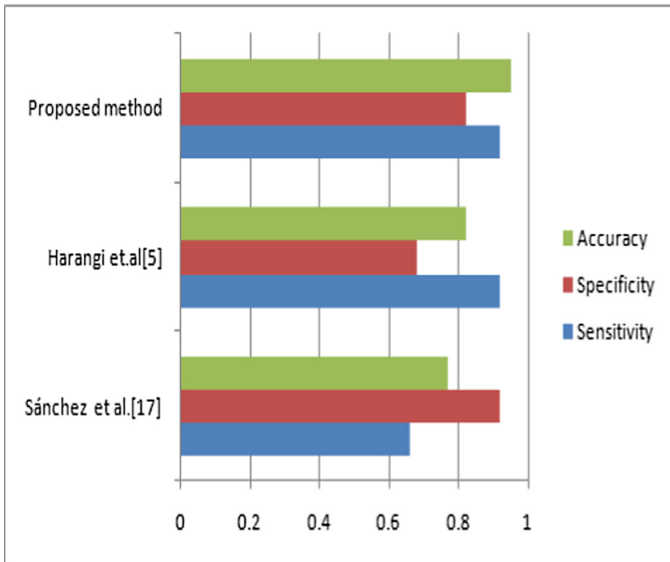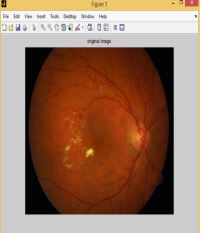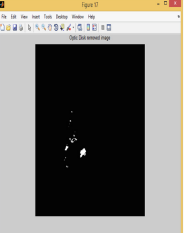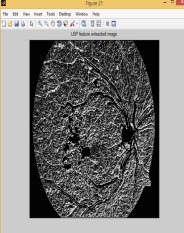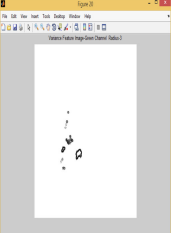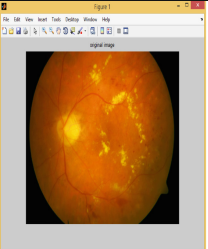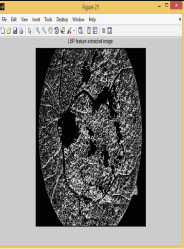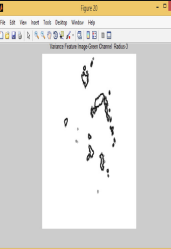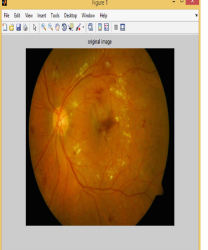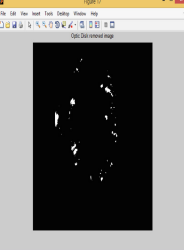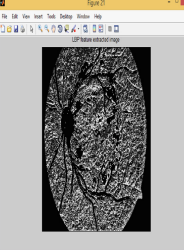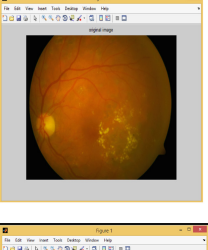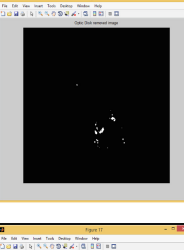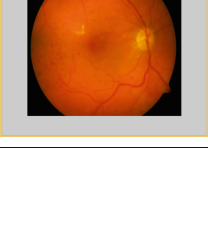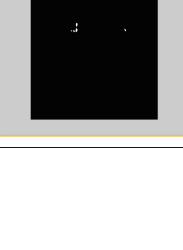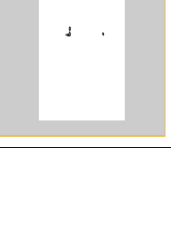
Fig. 9. Comparison of exudates segmentation algorithms

Table 1. Time complexity and accuracy of the algorithm for different databases and hospital images

Databases	Accuracy of classifier	Time taken
DIARETDB0	0.95	8 min 18 s
DIARETDB1	0.95	5 min 42 s
Hospital images	0.96	3 min 13 s

Table 2. Results of the algorithm for few retinal images

4   Conclusion

The detection of exudates with the help of local binary pattern, which is a texture descriptor, is proposed in this paper. The features extracted through LBP helps to classify the fundus image into normal and DR image. The algorithm is tried on DIARETDB0, DIARETDB1 and hospital images. The accuracy, sensitivity and specificity obtained are 0.951, 0.916 and 0.82 respectively. The obtained results are validated by ophthalmologist of Vijaya Nethralaya Hospital, Bengaluru. The method provides better performance when compared with the other exudates segmentation methods. In future, we have plans to continue the approach for detection of the dark lesions like microaneurysms and hemorrhages in the retinal images. By making use of multiclass SVM, the stages of DR can be identified.

References

1. Akram, M.U., Khalid, S., Khan, S.A.: Identification and classification of microaneurysms for early detection of diabetic retinopathy. Pattern Recognit. **46**, 107–116 (2013)
2. Tamilarasi, M., Duraiswamy, K.: Automatic detection of microaneurysms using microstructure and wavelet methods. Sadhana **40**, 1185–1203 (2015). Indian Academy of Sciences
3. Saha, R., Chowdhury, A.R., Banerjee, S.: Diabetic retinopathy related lesions detection and classification using machine learning technology. In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) ICAISC 2016. LNCS (LNAI), vol. 9693, pp. 734–745. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39384-1_65
4. Sreejini, K.S., Govindan, V.K.: A review of computer aided detection of anatomical structures and lesions of DR from color retina images. Int. J. Image Graph. Signal Process. **7**, 55–69 (2015)
5. Harangi, B., Hajdu, A.: Automatic exudates detection by fusing multiple active contours and regionwise classification. Comput. Biol. Med. **54**, 156–171 (2014)
6. Lachure, J., Deorankar, A.V., Lachure, S., Gupta, S., Jadhav, R.: Diabetic retinopathy using morphological operations and machine learning. IEEE (2015)
7. Sánchez, C.I., García, M., Mayo, A., López, M.I., Hornero, R.: Retinal image analysis based on mixture models to detect hard exudates. Med. Image Anal. **13**, 650–658 (2009)
8. Amin, J., Sharif, M., Yasmin, M.: A review on recent developments for detection of diabetic retinopathy. Scientifica, Article ID 6838976 (2016)
9. Osareh, A., Shadgar, B., Markham, R.: A computational-intelligence-based approach for detection of exudates in diabetic retinopathy images. IEEE Trans. Inf. Technol. Biomed. **13**, 535–545 (2009)
10. de la Calleja, J., Tecuapetla, L., Auxilio Medina, M., Bárcenas, E., Urbina Nájera, A.B.: LBP and machine learning for diabetic retinopathy detection. In: Corchado, E., Lozano, J.A., Quintián, H., Yin, H. (eds.) IDEAL 2014. LNCS, vol. 8669, pp. 110–117. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10840-7_14
11. Krishnan, M.M.R., Laude, A.: An integrated diabetic retinopathy index for the diagnosis of retinopathy using digital fundus image features. J. Med. Imag. Health Inf. **3**, 306–313 (2013)
12. Morales, S., Engan, K., Naranjo, V., Colomer, A.: Retinal disease screening through local binary patterns. IEEE J. Biomed. Health Inf. **21**, 184–192 (2015)

13. Devaraj, D., Prasanna Kumar, S.C.: Robust detection of hard exudates for diagnosis of non-proliferative diabetic retinopathy using integrated approach. In: Abraham, A., Cherukuri, A.K., Madureira, A.M., Muda, A.K. (eds.) SoCPaR 2016. AISC, vol. 614, pp. 525–534. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-60618-7_52
14. Mamatha, B.V., Srilatha, L.R., Devaraj, D., Prasanna Kumar, S.C.: A survey on different classifiers for medical diagnosis and grading: application to diabetic retinopathy. Int. J. Healthc. Sci. **2**, 210–216 (2015)
15. Devaraj, D., Prasanna Kumar, S.C.: Automated diagnosis of diabetic retinopathy using SVM and multiclass SVM. Int. J. Adv. Sci. Tech. Res. **3**, 316–323 (2014)
16. Ramya, J., Soundarya, S., Nagoormeeral, A., Revathi, E.: Detection of exudates in color fundus image. Int. J. Innov. Res. Sci. Eng. Technol. **3**, 10659–10665 (2014)
17. Sánchez, C.I., Hornero, R., López, M.I., Aboy, M., Poza, J., Abásolo, D.: A novel automatic image processing algorithm for detection of hard exudates based on retinal image analysis. Med. Eng. Phys. **30**, 350–357 (2008)

Automated Lung Parenchyma Segmentation in the Presence of High Attenuation Patterns Using Modified Robust Spatial Kernel FCM

Shyla Raj[1(✉)], D. S. Vinod[1], and Nagaraj Murthy[2]

[1] Department of Information Science and Engineering,
Sri Jayachamarajendra College of Engineering, Mysuru, Karnataka, India
shylanataraj@gmail.com, dsvinod@daad-alumni.de
[2] Department of Radiology, JSS Medical College, Mysuru, Karnataka, India
nagnishu@gmail.com

Abstract. Employing an accurate lung segmentation procedure is an essential stride in any Computer Aided Diagnosis (CAD) system designed for lung pathology evaluation, which significantly affects the performance of the system. In this paper a fully automated Modified Robust Spatial Kernel Fuzzy C Means (MRSKFCM) algorithm is used for extracting the lung parenchyma affected with Diffuse Lung Disease (DLD). The algorithm comprises of two steps, the Robust Spatial Kernel Fuzzy C Means (RSKFCM) is applied to acquire the coarse lung segmentation in the first step and in the second step the results of coarse segmentation is improved using Convex Hull (CH) algorithm and morphological operations. The algorithm used demonstrates high segmentation results with Dice index of 0.9057 ± 0.025, Jaccard Index of 0.8561 ± 0.029 and Accuracy 0.9864 ± 0.002. The execution time for each slice is approximately 2.5 s. Based on segmentation results and the evaluation time, the algorithm used can be considered as one of the desirable method for the extracting lung parenchyma in the CAD system.

Keywords: Diffuse lung diseases · High attenuation patterns
Lung segmentation · MRSKFCM · Convex hull · TALISMAN

1 Introduction

The DLDs are a group of chronic lung disorders, often collectively referred to as the Interstatial Lung Diseases (ILDs) [6] that affects the tissues and spaces around the alveoli in the lungs, by making them scarred and thickened. This gradually cause lung stiffness, reducing the capability of the alveoli to transfer oxygen into the bloodstream and ultimately leads to irreversible loss in the capability to breath. The cases of DLDs has significantly increased annually by range of 39 cases per every 1,00,000 cases globally [11]. DLDs are caused by inhaling

hazardous substances, infections, autoimmune diseases, genetic abnormalities and exposures to certain drugs. Since DLDs exhibit similar clinical, physiological and/or pathological features, the radiologist use High Resolution Computed Tomography (HRCT) scans for accurate diagnosis of DLDs. HRCT of thorax is generally considered as the most appropriate protocol for DLD quantification as it illustrates significant radiographic differences between pathologies and provides invaluable information that improves the accuracy and overall reproducibility in the interpretation of pathologies [25]. Interpreting HRCT images is a challenging and time captivating job even for the experienced radiologists and lung specialists, due to lack of strict clinical guidelines, high resemblances in the visual pattern of DLD and huge volume of data produced by HRCT. So researches are carried on developing the CAD system that can assist radiologist in the interpretation of HRCT scans [18].

CAD refers to the diagnosis made by the radiologist taking into regard the outcome of an automated computer analysis of the radiographic images. This helps in improving the overall accuracy in the perception of radio-graphs when the analysis made by computers is used as the second opinion [9,15]. In the CAD system developed for DLD evaluation prior to the interpretation of DLD pattern, the lung parenchyma is segmented from the HRCT scan.

The normal lungs filled with air appear darker when imaged in HRCT while the surrounding tissues and bones appear lighter. The work based on thresholding technique [10,20] depends on the above observation to segment the lung parenchyma from HRCT image. However this observation becomes ineffective in recognizing the lung borders when the high attenuation abnormalities affect the lung borders, because the intensity of abnormalities closely matches with the intensities of the surrounding muscle or bones. Inaccurate lung border identification leads to the exclusion of such abnormalities and would results in fallacious assessment of pathology and incorrect estimation of lung volume.

Hence MRSKFCM algorithm has been used for automatic lung segmentation without the need of any manual intervention for extracting the lung parenchyma.

2 Literature Survey

The conventional methods based on thresholding technique [10,20] requires manual selection of threshold value and yields good results only for healthy lungs and lungs with low attenuation patterns. Studies have shown that the local texture analysis methods is applied to analyse the lung patterns at its periphery [13,14,23], but these methods are time consuming and require manual intervention and in addition they were evaluated on less number of slices.

The shape based techniques in [21,24] uses the lung shape information for improving the lung segmentation accuracy. However the main limitation of shape models is the requirement of large dataset for training the models to learn shape information. The segmentation by registration based approach [17,22] requires development of an atlas model which serves as the guide to segment lungs. This method is efficient in most of the difficult cases but it is time consuming and requires large databases of annotated images.

The clustering techniques divides the pixels into constituent groups based on some similarity measures such as distance, connectivity and intensity. The fuzzy clustering techniques is widely used in medical image segmentation for its ability to introduce fuzziness in classifying the pixels. So it can very well handle limited spatial resolution, non uniform intensity variations, poor contrast and noise problems faced in the medical image segmentation. Ahmad et al. [1] suggested a fully automated, unsupervised segmentation approach using FCM clustering. Since traditional FCM does not consider spatial co-relation and uses only euclidean distance, this method fails in case of high attenuation patterns. Shi et al. [16] introduced a Gaussian kernel-induced distance metric along with spatial information for automatic lung segmentation. Aruna Kumar and Harish [2] suggested novel Robust Spatial Kernel FCM (RSKFCM) algorithm incorporating both kernel induced distance and spatial function for segmenting MRI Brain images. Though the methods in [2,16] overcomes the disadvantages of FCM, it fails to segment when high attenuation patterns affect the lung periphery. Dash et al. [7] compared the performance of Gaussian Mixture Model, Mean Shift and Hidden Markov Random Field for automatic lung segmentation and found Mean Shift to be most appropriate among them. This work used TALISMAN dataset for the experimentation.

The method to be used in clinical evaluation, must produce accurate segmentation results in short time period. The shape based techniques and segmentation by registration methods are too slow and non-viable to be used in clinical practice. So this work is aimed at evaluating DLD cases in as short time possible.

3   Dataset Used

Publically available standard ILD database, TALISMAN built at the University Hospitals of Geneva (HUG) is used for the work [8]. The database contains of 108 DICOM image series collected from 128 patients influenced with anyone of the 13 histological recognized ILDs. The size of each image matrix is 512×512 with 1–2 mm slice thickness and 10–15 mm spacing between the slices. The database is provided with 3D annotated regions of pathologies along with clinical parameters delineated by 2 experienced radiologists. For the work 92 images series from TALISMAN dataset is considered, the details are as follows:

6 images series of Consolidation, 4 of Emphysema, 28 Fibrosis cases, 19 Ground Glass cases, 7 Healthy patterns, 3 Macronodules patterns, 18 cases of Micronodules and 7 cases of Reticulation patterns. All the slices in a scan are considered for the evaluation.

4   Methodology Used

The MRSKFCM algorithm comprises of two steps, the coarse lung segmentation is obtained using RSKFCM algorithm [2] in the first step and in the second step the results of coarse segmentation is improved using CH algorithm and morphological operations to obtain final lung segmentation. The MRSKFCM is explained in the Algorithm 1.

Data: The HRCT lung image series $l = \{l_1, l_2, \ldots, l_n\}$
Result: Segmented lung parenchyma $s = \{s_1, s_2, \ldots, s_n\}$
repeat

> Read a HRCT slice from the image series.
> Obtain the coarse lung segmentation by segmenting the image
> into two clusters c_1 c_2, one for representing lung region
> another for non-lung region using RSKFCM algorithm.
> Label the pixels in the image and generate the segmentation results.
> Determine the cluster containing lung region c_2.
> Apply CH algorithm to the results of RSKFCM (calls Algorithm 2)
> CH = CH(c_2)
> Perform morphological operations on the results of CH (calls Algorithm 3)
> to obtain final segmentation results
> Final Result = Morphological Operations(CH)

until *For all slices in the image series $l = \{l_1, l_2 \ldots, l_n\}$*;

<div align="center">

Algorithm 1. MRSKFCM algorithm

</div>

4.1 Coarse Lung Segmentation

In the work RSKFCM segmentation algorithm is employed to extract the lung parenchyma from the CT scans. The HRCT lung image is segmented into two clusters one for representing lung region and another for representing the non-lung region. The RSKFCM overcomes the disadvantages of traditional FCM [4] by considering the spatial co-relation between the neighbouring pixels and by using the kernel induced metric for distance calculation.

To exploit the spatial co-relation between the pixels the spatial function S_{ij} is calculated using Eq. (1)

$$S_{ij} = \sum_{k \epsilon NW(x_i)} U_{jk} \tag{1}$$

where x_i denote the data point, v_j denote the cluster center, NW(x_i) represents the square window centred around x_i. The Spatial function S_{ij} exhibits the likelihood of pixel x_i belonging to j^{th} cluster.

The Spatial function S_{ij} is integrated into the Membership Function as shown in Eq. (2)

$$U_{ij}^{'} = \frac{U_{ij}^a S_{ij}^b}{\sum_{k=1}^{C} U_{ki}^a S_{ki}^b} \tag{2}$$

The parameters a and b regulate the respective significance of the two functions.

To kernel substitution is used to convert the non-linear data into linear data by transforming them into higher dimensional space [5]. The kernel substitution is shown in Eq. (3)

$$\kappa(x, y) = (\phi(x), \phi(y)) \tag{3}$$

The Euclidean distance in FCM $\|x_i - v_j\|^2$ is replaced by kernel function $\|\phi(x_i) - \phi(v_j)\|^2$, where the distance between the data point x_i and the cluster center v_j is

computed as the sum of their self-similarity $\kappa(x_i, x_i)$, $\kappa(v_j, v_j)$ and the difference of their cross similarity $\kappa(x_i, v_j)$ as shown in Eq. (4)

$$\|\phi(x_i) - \phi(v_j)\|^2 = \kappa(x_i, x_i) + \kappa(v_j, v_j) - 2\kappa(x_i, v_j) \tag{4}$$

Since the Gaussian Radial Basis Function (GRBF) is used in the work, the self similarity $\kappa(x_i, x_i)$ and $\kappa(v_j, v_j)$ equates to 1 and the kernel distance in Eq. (4) is reduced as follows:

$$1 + 1 - 2(\kappa(x_i v_j)) \tag{5}$$

Simplification of Eq. (5) gives the kernel distance measure as shown in Eq. (6)

$$\|\phi(x_i) - \phi(v_j)\|^2 = 2(1 - \kappa(x_i v_j)) \tag{6}$$

The algorithm aims to find a local optima by minimizing the function given by the Eq. (7)

$$J = 2 \sum_{i=1}^{N} \sum_{j=1}^{C} U'_{ij}(1 - \kappa(x_i v_j)) \tag{7}$$

The weights are updated using the Eq. (8)

$$\frac{\sum_{k=1}^{N} U'_{ij}\kappa(x_i v_j)x_i}{\sum_{k=1}^{N} U'_{ij}\kappa(x_i v_j)} \tag{8}$$

4.2   Final Lung Segmentation

The RSKFCM segmentation algorithm fails to recognize the lung border when high attenuation patterns appear at the periphery, because the attenuation of the pattern at the lung borders closely matches to that of their surrounding bones or tissues. To identify the abnormalities present at the lung border, CH algorithm is applied on the results of the RSKFCM results.

CH is one of the simplest shape proximation technique used for approximating complex shapes [12]. The CH algorithm is used to refine the lung borders by approximating the structure of the lungs. Given a finite set of points in a plane, CH of the set is the smallest convex polygon, which comprises of all points. The work uses Quick hull algorithm, based on divide and conquer approach to compute the CH [3]. The Quick hull algorithm is explained in Algorithm 2.

The results of CH algorithm and morphological operations are as shown in Fig. 1. The results obtained after applying CH to the coarse segmentation includes some non-lung regions. To eliminate these non-lung region and to obtain the final lung segmentation, the morphological operations are performed. The Algorithm 3 explains the morphological operations performed on the results of CH.

Data: The set of points 'n' in plane S = $\{p_1, p_2, \ldots, p_n\}$
Result: Convex Hull
Identify the points with minimum and maximum x co-ordinates in S by sorting.
The p_1 and p_n form the vertices of hull.
The line formed by p_1 p_n divides the set
into upper and lower hull S_1 and S_2 which will be solved recursively.
repeat

> From the S_1 find the point with maximum perpendicular distance
> P_{max} from line p_1 p_n.
> The points p_1 p_n and p_{max} form a triangle.
> The line p_1 p_{max} divides the S_1 into left subset S_{11} and right subset S_{22}.
> The points lying inside the triangle cannot be part of the convex hull
> and it is not considered for the succeeding steps.
> Similarly find the points from p_1 S_{11} p_{max} and p_{max} S_{12} p_n.

until *Until no more points are left*;
The same procedure is carried for the lower hull.
The union of upper and lower hull constitute the convex hull containing all
points in the plane.

Algorithm 2. Quick hull

Fig. 1. Results of convex hull and morphological operations (a) coarse lung segmentation by RSKFCM (b) convex hull on RSKFCM (c) difference in the results of RSKFCM and CH (d) flaws in segmentation (e) final lung segmentation by MRSKFCM

5 Experimental Set-Up and Evaluation

5.1 Parameter Configuration

The square window of size 3×3 is used to obtain the spatial co-relation between the neighbouring pixels and kernel width (σ) in the GRBF is set to 775. The value of a is set as 0 and b is set as 2 which makes RSKFCM fully spatial.

5.2 Methodology Evaluation

The evaluation of the MRSKFCM segmentation algorithm is done by correlating the segmentation results with the available ground truth. The efficiency of methodology used is estimated based on following metric categories: overlap based metrics and volume similarity metrics. Since medical images are inherently fuzzy in nature, overlap based metrices are used to estimate the quality of segmentation. The overlap based metrices are estimated based on the cardinalities of the confusion matrix. Dice Similarity Co-efficient (DSC), Jaccard Index

Data: The results of CH c = $\{c_1, c_2, \ldots, c_n\}$ and
Coarse segmentation result from RSKFCM $s' = \{s'_1, s'_2, \ldots, s'_n\}$
Result: The final lung segmentation results s = $\{s_1, s_2, \ldots, s_n\}$
repeat
> To identify the non-lung regions,the results of initial RSKFCM segmentation are subtracted from the results of the CH.
> $nl_1 = c_1 - s'_1$
> The undesired non-lung regions is removed by using a morphological erosion operation and the small responses in the image are eliminated based on their size.
> $nl'_1 = \text{morphological erosion}(nl_1)$
> $nl''_1 = \text{area}(nl'_1)$
> The region left after the erosion operation and removing responses indicate the flaws in the segmentation.
> $fs_1 = nl''_1$
> The final segmentation result is obtained by subtracting the flaws in segmentation from initial results of the CH.
> $s_1 = c_1 - fs_1$

until *For all the results of CH c=$\{c_1, c_2, \ldots, c_n\}$;*

Algorithm 3. Mathematical morphological operations

(JI), Accuracy (Acc), True Negative Rate (TNR), True Positive Rate (TPR), Global Consistency Error (GCE) are used as overlap based metrics. In the volume similarity metric the volumes of the test segment and reference segment are considered to assess the quality of segmentation. For volume similarity metric Volume Similarity (VS) index is used [19]. The good segmentation results are depicted by the higher values DCS, JI, Acc, TPR, TNR, VS and lower values of GCE.

6   Results and Discussion

The results of RSKFCM is considered as coarse segmentation result and the results of MRSKFCM is considered as the final segmentation result. It is evident from the results of Fig. 2 that the coarse lung segmentation obtained from RSK-FCM algorithm performs effective segmentation for high attenuation patterns

(a)   **(b)**   **(c)**   **(d)**
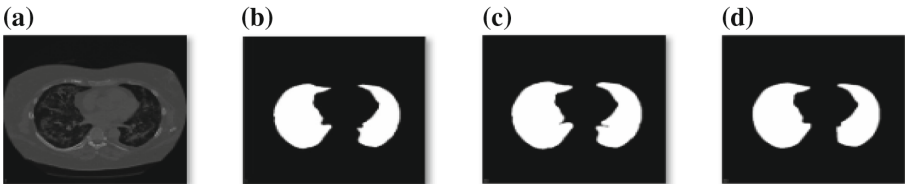
Fig. 2. Segmentation results for high attenuation consolidation pattern distributed within the lungs (a) original image (b) ground truth (c) coarse lung segmentation (d) final lung segmentation
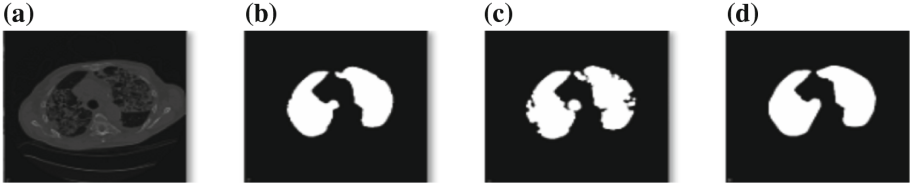
Fig. 3. Segmentation results for high attenuation fibrosis pattern at the periphery (a) original image (b) ground truth (c) coarse lung segmentation (d) final lung segmentation

Table 1. Disease pattern wise segmentation performance measures

Type	Segmentation result	DSC	JI	Acc	TPR	TNR	GCE	VS
Consolidation	Coarse	0.8905	0.8182	0.9778	0.9822	0.9755	0.0409	0.9041
	Final	0.9368	0.8925	0.9884	0.9832	0.9879	0.0217	0.9476
Emphysema	Coarse	0.8804	0.8089	0.9781	0.9902	0.9739	0.0398	0.8818
	Final	0.9295	0.8856	0.9879	0.9850	0.9858	0.0225	0.9348
Fibrosis	Coarse	0.8187	0.7448	0.9755	0.9211	0.9762	0.0403	0.8434
	Final	0.8565	0.8059	0.9803	0.9201	0.9816	0.0302	0.8818
Ground Glass	Coarse	0.8549	0.7738	0.9758	0.9861	0.9724	0.0432	0.8593
	Final	0.9006	0.8465	0.9857	0.9857	0.9837	0.0260	0.9050
Healthy	Coarse	0.8711	0.8012	0.9805	0.9921	0.9776	0.0353	0.8754
	Final	0.9149	0.8692	0.9890	0.9896	0.9877	0.0201	0.9197
Macronodules	Coarse	0.8344	0.7452	0.9767	0.9413	0.9790	0.0427	0.8770
	Final	0.8836	0.8248	0.9865	0.9544	0.9877	0.0253	0.9136
Micronodules	Coarse	0.8804	0.8106	0.9767	0.9906	0.9726	0.0422	0.8839
	Final	0.9169	0.8709	0.9868	0.9884	0.9851	0.0241	0.9213
Reticulation	Coarse	0.8654	0.7897	0.9778	0.9811	0.9756	0.0402	0.8774
	Final	0.9065	0.8538	0.9862	0.9752	0.9861	0.0252	0.9221

distributed within the lungs, as RSKFCM takes into consideration the spatial co-relation between the neighbouring pixels and uses kernel induced metric for distance calculation. The results for high attenuation patterns at the periphery is as shown in Fig. 3. It can be inferred from the coarse segmentation results that the RSKFCM fails to identify the high attenuation patterns at the periphery due to close match in intensities between the patterns and the surrounding bones and tissues. However after applying CH and performing morphological operations on the coarse segmentation results, the final segmentation results are improved and better segmentation is achieved. The Table 1 displays the disease pattern wise segmentation performance for coarse segmentation and final segmentation. It was anticipated from the results of Fig. 2 that RSKFCM algorithm

Table 2. Comparison of the work with the work in literature

Method	DSC	JI	Acc
Mean shift [7]	0.5555 ± 0.07	0.4133 ± 0.06	0.8036 ± 0.08
FCM [1]	0.8494 ± 0.038	0.7735 ± 0.038	0.9742 ± 0.020
RSKFCM [2]	0.8620 ± 0.025	0.7866 ± 0.029	0.9774 ± 0.002
Current work	0.9057 ± 0.025	0.8561 ± 0.029	0.9864 ± 0.002

alone would work well for low attenuation patterns, healthy lungs and also for high attenuation patterns distributed within the lungs. The CH and morphological operations on RSKFCM results would further improve the performance of lung segmentation for high attenuation patterns at the periphery. However after examining the results of Table 1 it is inferred that the results of final segmentation are better than the coarse segmentation for all types of DLD patterns, as the CH helps in refining the lung borders by making it smooth.

The Table 2 presents the comparison of the current work with the work in literature performed on same dataset TALISMAN. The DSC, JI and Acc metrices are used to compare the results. It can be inferred from the experimental results that the MRSKFCM algorithm performs better segmentation than the FCM [1], Mean shift [7] and RSKFCM [2] methods in terms of DSC, JI and Acc. This proves the efficiency of MRSKFCM in handling the medical images in comparison to the clustering work.

It can be proved from the results of Fig. 3 that the MRSKFCM algorithm works efficiently for difficult patterns present at the periphery with some limitations in the base and apex of the lungs.

7   Conclusion

The supervised techniques employed for lung segmentation requires large dataset and time for training the models. The segmentation by registration approach and shape based approach are too slow to be used in clinical practice. To overcome these shortcomings, an unsupervised automatic segmentation technique is introduced in this paper without the need of training and testing procedures. The experiments is conducted on different types of DLD cases with varying severities. The results of the work proves its efficiency in segmenting the lungs with various types of DLDs including the high attenuation patterns at the lung periphery.

References

1. Ahmad, W.S.H.M.W., Zaki, W.M.D.W., Fauzi, M.F.A.: Lung segmentation on standard and mobile chest radiographs using oriented Gaussian derivatives filter. Biomed. Eng. Online **14**(1), 20 (2015)
2. Aruna Kumar, S.V., Harish, B.S.: Segmenting medical images using computational intelligence technique. Int. J. Inf. Process. **9**(1), 48–56 (2015)

3. Barber, C.B., Dobkin, D.P., Huhdanpaa, H.: The quickhull algorithm for convex hulls. ACM Trans. Math. Softw. (TOMS) **22**(4), 469–483 (1996)

4. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Springer Science & Business Media, Heidelberg (2013). https://doi.org/10.1007/978-1-4757-0450-1

5. Cover, T.M.: Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. IEEE Trans. Electron. Comput. **3**, 326–334 (1965)

6. Dalpiaz, G., Maffessanti, M.: Diffuse lung diseases. In: Guglielmi, G., Peh, W., Guermazi, A. (eds.) Geriatric Imaging, pp. 365–388. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-35579-0_16

7. Dash, J.K., Madhavi, V., Mukhopadhyay, S., Khandelwal, N., Kumar, P.: Segmentation of interstitial lung disease patterns in HRCT images. In: SPIE Medical Imaging, p. 94142R. International Society for Optics and Photonics (2015)

8. Depeursinge, A., Vargas, A., Platon, A., Geissbuhler, A., Poletti, P.A., Müller, H.: Building a reference multimedia database for interstitial lung diseases. Comput. Med. Imaging Graph. **36**(3), 227–238 (2012)

9. Doi, K.: Current status and future potential of computer-aided diagnosis in medical imaging. Br. J. Radiol. **78**(Suppl. 1), s3–s19 (2005)

10. Haider, C., Bartholmai, B.J., Holmes, D., Camp, J., Robb, R.A.: Quantitative characterization of lung disease. Comput. Med. Imaging Graph. **29**(7), 555–563 (2005)

11. Hutchinson, J., Fogarty, A., Hubbard, R., McKeever, T.: Global incidence and mortality of idiopathic pulmonary fibrosis: a systematic review. Eur. Respir. J. **46**(3), 795–806 (2015)

12. Jayaram, M., Fleyeh, H.: Convex hulls in image processing: a scoping review. Am. J. Intell. Syst. **6**(2), 48–58 (2016)

13. Korfiatis, P., Kalogeropoulou, C., Karahaliou, A., Kazantzi, A., Skiadopoulos, S., Costaridou, L.: Texture classification-based segmentation of lung affected by interstitial pneumonia in high-resolution CT. Med. Phys. **35**(12), 5290–5302 (2008)

14. Korfiatis, P., Skiadopoulos, S., Sakellaropoulos, P., Kalogeropoulou, C., Costaridou, L.: Automated 3d segmentation of lung fields in thin slice CT exploiting wavelet preprocessing. In: Kropatsch, W.G., Kampel, M., Hanbury, A. (eds.) CAIP 2007. LNCS, vol. 4673, pp. 237–244. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74272-2_30

15. Rogers, W., Ryack, B., Moeller, G.: Computer-aided medical diagnosis: literature review. Int. J. Biomed. Comput. **10**(4), 267–289 (1979)

16. Shi, Z., Zhou, P., He, L., Nakamura, T., Yao, Q., Itoh, H.: Lung segmentation in chest radiographs by means of Gaussian kernel-based FCM with spatial constraints. In: Sixth International Conference on Fuzzy Systems and Knowledge Discovery. FSKD 2009, vol. 3, pp. 428–432. IEEE (2009)

17. Sluimer, I., Prokop, M., Van Ginneken, B.: Toward automated segmentation of the pathological lung in CT. IEEE Trans. Med. Imaging **24**(8), 1025–1038 (2005)

18. Sluimer, I., Schilham, A., Prokop, M., Van Ginneken, B.: Computer analysis of computed tomography scans of the lung: a survey. IEEE Trans. Med. Imaging **25**(4), 385–405 (2006)

19. Taha, A.A., Hanbury, A.: Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. BMC Med. Imaging **15**(1), 29 (2015)

20. Uchiyama, Y., Katsuragawa, S., Abe, H., Shiraishi, J., Li, F., Li, Q., Zhang, C.T., Suzuki, K., et al.: Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography. Med. Phys. **30**(9), 2440–2454 (2003)

21. Van Ginneken, B., Frangi, A.F., Staal, J.J., ter Haar Romeny, B.M., Viergever, M.A.: Active shape model segmentation with optimal features. IEEE Trans. Med. Imaging **21**(8), 924–933 (2002)
22. Van Rikxoort, E.V., Van Ginneken, B.: Automatic segmentation of the lungs and lobes from thoracic CT scans. In: Proceedings of the 4th international Workshop Pulmonary Image Anal, pp. 261–268 (2011)
23. Wang, J., Li, F., Li, Q.: Automated segmentation of lungs with severe interstitial lung disease in CT. Med. Phys. **36**(10), 4592–4599 (2009)
24. Xu, T., Mandal, M., Long, R., Cheng, I., Basu, A.: An edge-region force guided active shape approach for automatic lung field detection in chest radiographs. Comput. Med. Imaging Graph. **36**(6), 452–463 (2012)
25. Zwirewich, C.V., Mayo, J.R., Müller, N.: Low-dose high-resolution CT of lung parenchyma. Radiol. **180**(2), 413–417 (1991)

A Heuristic Approach to Automatically Segment Signal from Background in DNA Microarray Images

S. S. Manjunath[1] , Priya Nandihal[1(✉)] ,
and Lalitha Rangarajan[2]

[1] Dayananda Sagar Academy of Technology and Management, Bangalore, India
mnj_ss2002@yahoo.co.in, talk2priya.nandihal@gmail.com
[2] University of Mysore, Mysore, India
lali85arun@yahoo.co.in

Abstract. Microarray is an efficacious tool used to detect, analyze and describe local features of the genome in the form of an image. A microarray is glass slide which consists of different nucleic acid inquest added chemically. Microarray provides enough guidance on infection pathology, progression, resistance to treatment, and response to cellular microenvironments and ultimately may lead to improved and innovative approaches for disease diagnosis. Study of microarray includes these steps: initially spot addressing followed by spot separation as background and foreground and finally intensity extraction of separated spots. In the spot separation stage, extended k-means and Fast EM methods are proposed. The proposed method is a heuristic approach where Coefficient-Variation of the input image is used as a metric to switch the image for segmentation to the appropriate existing method as well as to the proposed method to efficiently segment the foreground signals from the background in existence of undetermined or unwanted pixel values, and inadequately articulated spots in DNA microarray images. Experimental results and analysis of proposed methods shows promising results when compared with existing methods in the literature.

Keywords: Microarray experiment · Expectation maximization
K-means · Segmentation

1   Introduction

Microarray is an important tool and powerful technique which are helpful to investigate the expression of c-DNA in organisms for huge degree of gene sequences and expressions. Microarray represents a scientific intersection between life science and electronic machine that makes analysis of gene information in human on a genome large scale. Microarray technology allows particularly similar, maximum throughput formulation of cDNA expression in experimentation [1]. cDNA microarray technology enables complete genome analysis of gene expression levels. Possible applications of microarray have created much anticipation and some uncertainty, with in the bioinformatics community [2]. A few research scholars have recommended that within ten

years, microarray will be frequently used not only for the selection, assessment, and quality control of the most excellent drugs for pharmaceutical advancement, but also for sickness identification and for monitoring preferred and adverse outcomes of beneficial interventions. Microarray technology classifies tumors by molecular characteristics, rather than by cell features. Hence is capable of identifying the occurrence of cancer much easier than compared to that of existing diagnostic systems. Apart from this microarray have their application in the field of drug development and manufacturing, tumor detection and classification, gene mutation, detection functionalities of mutated genes etc. Microarray enables the instantaneous study of DNA nucleotide sequences on a one microscopic physical substrate. A microarray is the corporeal platform to which ordered genetic material data are attached. In case of two channel microarray experiments, the mRNA is extracted from a test sample, reverse transcribed and labeled with a fluorescent dye. Accordingly same procedure is repeated for another sample employing with a different fluorescent dye subsequently, both samples are exposed to a Microarray where the labeled target molecules will hybridize to the printed probes. By scanning the array at different wavelengths two images are produced one for each dye. Green stain shows that the test substance has lesser action than the indication material (up-regulated), red strains indicate that the test substance has more active than the indication material (down-regulated) and yellow spots imply that there is no change in the activity level between the two populations of test and reference material. The amount of fluorescent at each DNA spot corresponds to the transcript level and/or other stages of gene regulation at the mRNA level of gene printed in a particular spot. The prime objective of microarray is the discovery of biological knowledge like metabolic pathways. The dealing out of microarray images [3] gives the key for some more investigation of the collected microarray samples [4].

Microarray image examinations are classified into three categories namely spot identification, spot fragmentation, and quantification. Addressing or spot identification is the method of identifying position of the spots. Segmentation or spot fragmentation allows the classification of foreground area and background area of spots. Quantification includes estimating, red and green intensity value for each spot.

2   Review of Literature

Many commercial packages and research ideas have been discussed for the separation of foreground and background of microarray images [5–9]. Separating the foreground signal from the background methods can be g classified into 4 categories. Namely, Fixed circle separation, Adaptive circle and shape separation and Histogram segmentation [10].

ScanAlyze [6], flat circle methods has been described with the provision that spots are assumed to be around shaped with a definite, known radius. In GenePix [5], a flexible ring separation approach is used. With this approach, the radius of every mark is dynamic but regulated for each spot individually. However, a circle shaped spot mask gives a bad fit to asymmetrical spots or donut fashioned marks with interior holes because the methodology constraints to the morphology of the spots to be circles.

The most frequently applied approaches in flexible morphological separation technique are the watershed [11] and seeded region growing [12, 13] SPOT Software is one such tool which makes use of both methodologies [12]. Flexible segmentation based on morphology is the prominent method in microarray image analysis. Any prerequisites like size and the morphology of the spot are taken into consideration by described method. Both methods described fails in the actual specification of the preliminary points or the seeds. The seeded region growing procedure is more adaptive and it can applied spots with different shapes and sizes. But it is ineffective in presence of noise.

Histogram-based techniques [10] estimate a fixed value. In case, intensity of the picture element is lower than the estimated fixed value then picture elements are classified as background picture elements, else pixels with greater intensity than estimated fixed value then pixels are classified as foreground pixels. ImaGene [14], UCSF also uses histogram information. In [7], Scan array express a circular mask is used and calculation of fixed value is based on Mann-Whitney test. When bigger target mask is initialized for variation it affects to the quantification and stage, this is the major drawback of discussed method. Larger diameter consequences in consideration of pixels from adjacent spots and also spots of dynamic intensity surrounded by a spot region are not likely in the range of above the fixed value and hence those picture elements are classified as background. Unwanted pixels of type artifact with soaring intensity will be grouped as part of foreground.

Dapple [9] uses Laplacian of Gaussian border identification within a grating frame. The limitation here is that the ramp approach does not make assumption of composition of the image. So, for microarray images these methods often separate areas within one spot and boundaries not have connectivity and also these approaches do not make a distinction well between clamor of high intensity and real signal.

Methods discussed in [15], describes a primary curve positioning on the spot and a curve bend is outlined for the purpose of characterising the spot boundaries. But, the curve bend used is sensitive to noise.

Mixture model methods [16, 17] have been considered to segmenting microarray images. However, these approaches will not consider the dependencies among nearby pixels. In the analysis of model using mixtures [18], the most important negative aspect is the detection of spots with strength near enough to background.

In MRF models [19, 29] the use of appropriate and structure information has been projected for the task of fragmenting microarray images. The flaw in MRF based methodology is that, it has more computational cost and relies on prior postulation of division markering of all pixels.

In multi resolution approaches [20], the employ of DWT for the assignment of sculpting the tags at a variety of levels for segmentation is proposed. On the other hand for the duration of DWT disintegration process, the image dimension is condensed to partly and none of the additional information has been used. Nowadays, clustering is the most common technique used for the segmentation of the microarray images [21–26].

In clustering methods, such as K-means they consider some local features, like image intensity which is sensitive to noise. Also, the segmented foreground and background region need not be connected in this approach, whereas the actual foreground and background are connected regions. The foremost negative aspects of this technique are that they do not acclimatize well to asymmetrical based clusters and do not make the most of all the accessible former knowledge about the sample.

In [27], comparisons of most of the discussed segmentation methodologies proven so as to the most successful methods are those based on clustering by method of K-Means.

The main feature that influences the segmentation process is the undesirable contribution of noise, which directs to erroneous assessment of *first order momentum* intensities of spots and reduces the replication of the genetic expression levels, derived from microarray images. Also the noise signal may, sometimes be transient from the foreground region to the background region, arriving at misclassification of foreground/background pixels.

The reason behind using the concept of heuristic approach is that some images are noisy, weakly expressed and irregular shaped. Segmentation methods discussed in the literature would not be accurate method in classifying the foreground pixels from background pixels for the above discussed spot images, so the metric Coefficient of Variation is incorporated, based on which the existing or the proposed method is selected for segmentation.

The proposed method which is described in the section that follows is based on the assumption that gridding is already done and the input to the segmentation stage is a spot image. In Sect. 3, the proposed work, the metrics used for quantification of the proposed method are discussed. In Sect. 4, a detailed discussion on the experimental results and analysis is provided.

3 Proposed Segmentation Method

Segmentation is one of the crucial steps in microarray image analysis which separates foreground (signal) pixels and background pixels from the microarray spot. Since each spot is determined by gridding, then extended k-means from k-means and Fast EM proposed from EM (Expectation-Maximization) approaches are engaged to faction all the pixels into two clusters. Extended k-means is proposed to overcome the limitations of k-means in the segmentation of microarray spot. The major drawback in k-means is in the initialization step, which is resolved in the extended version of k-means. The membership for each data point xi, belonging to cj cluster of K-mean is calculated which depends on the minimum Euclidean distance as:

$$EKM(x, c) = \sum_{i=1}^{n} min|x_i - y_i|^2 \tag{1}$$

Thus combines the related clusters based on the EKM value, to form one whole connected cluster.

Similarly Fast EM is proposed to overcome the limitations in EM. Major drawback of EM is due to its assumption of sources of noise. Fast EM is obtained by filtering of Gaussian noise with low noise variance, say σ^2, which was not considered in EM. The sources are estimated separately and the sources are implicit to be whitened and of incorporated template form. By denoting one of the foundation signals in the best possible result as s_{opt}, optimality is achieved. The term optimality means that the standard EM-algorithm, which eventually converges to:

$$a_{opt} = X_{S_{opt}} \qquad (2)$$

the value of s_{opt} as

$$s_{opt} = E[s|a_{opt}, X, \sigma^2] \qquad (3)$$

For the optimal vector a_{opt}, we have

$$s_0 = \alpha s_{opt} + \beta s_G \qquad (4)$$

Where $\alpha^2 + \beta^2 = 1$. The clatter $\mathbf{s}_G$ is mostly due to the other sources and to a small extent the Gaussian noise in the data. The E-step (Expectation) filter reduces the unwanted pixels by creation of the prior distribution of s. This result in single sign on path to the sluggish junction: in a little clamor case the majority of the not needed pixels s_G is due to additional basis and a sluggish junction outcome. We can speed up the junction if we can sieve away that part of s_G. Thus by speeding up the convergence step the EM algorithm speeds up, resulting into fast expectation-maximization. The proposed framework is based on heuristic approach using coefficient of variation as a metric since the spot images are noisy, irregular shaped and weakly expressed, all the existing segmentation methods in the literature would not lead to a better result for all such type of spot images, so a metric which is the important factor in the segmentation of the microarray spot is proposed and is depicted below in Fig. 1.

The output of gridding process is taken as input to proposed methods (Extended k-means, Fast EM) for spot segmentation. Coefficient of variation (CV) is one of the potent metrics used to estimates the noise which is of great importance in microarray image segmentation. From the experimental analysis it has been found that for low CV value, extended K-mean yields accurate segmentation results, for high values of CV, Fast EM yields better results and for intermediate value of CV, existing k-mean would yield better result. Based on the above values of CV the input spot will be switched among the proposed methods and existing K-mean for segmentation. Thus, if coefficient of variation is least then it switches to Extended k-means, if coefficient of variation is highest then it switches to Fast EM otherwise it switches to existing K-mean thus making the whole process a heuristic approach. The heuristic approach also helps us to determine which method among the proposed and existing methods works efficiently in the occurrence of clamor and artifacts.
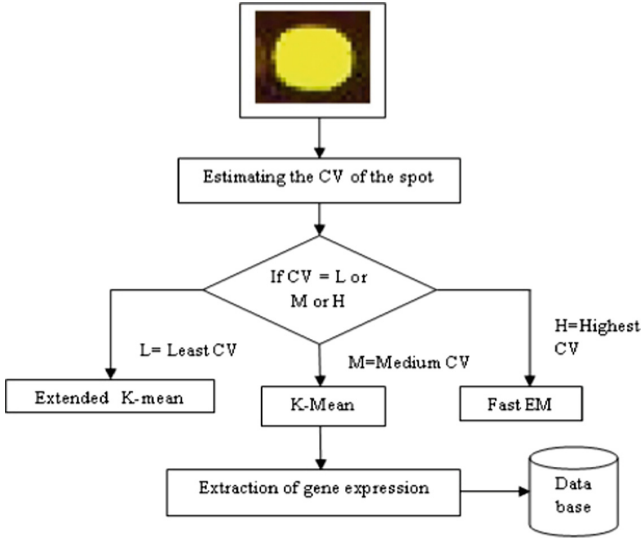
Fig. 1. Proposed heuristic approach

3.1 Materials and Methods

To assess the clustering ability of the anticipated method versus the existing method, a replicated cDNA image was formed as discussed in Sect. 2 [29]. In turn, to produce spots with pragmatic distinctiveness, the subsequent course of action was applied. An image with 6400 spots was employed as a sample, and its twofold edition was created by making use of a vergeing technique [29]. In the replicated image, not only the position but also the address of spot was already known. The first order moment concentration assessment of every spot was known prior, ranging between 0 and 2^{8-1} the red and green feeds [29]. Spot intensities were formed using an exponential allocation with first order moment and the pre-defined mean intensity. Background pixel values were pinched from a particular exponential allocation with first order momentum value concluded from the true cDNA image's mean intensity background [29]. To explore the algorithms' performance in presence of noise, the imitation image was corrupted with additive white Gaussian noise [17] of five different Signals to Noise Ratio (SNR) levels, 1, 3, 5, 7 and 9 dB.

Following metrics were calculated from the simulated images to assess the segmentation process:

1. The fluorescence ratio r [17] indicates the expression value of the corresponding gene according to the normalizing logarithmic ratio

$$r = \log_2\left(\frac{\mu_{FR} - \mu_{BR}}{\mu_{FG} - \mu_{BG}}\right) \qquad (5)$$

Where μ_{FR} and μ_{BR} are foreground and background mean of read channel respectively, and μ_{FG} and μ_{BG} are foreground and background mean of green channel respectively.

2. The coefficient of determination r^2 [27, 28] indicates the strength of the linear association between simulated and calculated cells, as well as, it gives the proportion of the variance (fluctuation) of the calculated data.

$$r^2 = \frac{\sum_{i=1}^{AllSpots}(Isegmented(i) - Iactual)^2}{\sum_{i=1}^{AllSpots}(Iactual - \bar{I}actual)^2} \tag{6}$$

Where *Isegmented* and *Iactual* are the first order momentum intensity values of the considered and created spots, respectively, individual cell images ranges from 1….624 and $\bar{I}actual$ is the whole first order momentum of the spot intensity values of the created image. If r^2 value is close to the unity, then the method has the best performance.

3. The concordance correlation P_c [27, 28] measures the agreement between simulated and calculated data and is used to evaluate the reproducibility of the proposed segmentation algorithms.

$$P_c = \frac{2S_A S_B r}{S_A^2 + S_B^2 + (\bar{A} - \bar{B})^2} \tag{7}$$

Where A and B are two samples, $\bar{A}$ and $\bar{B}$ are the mean values, S_A and S_B are the standard deviation of the samples. The higher the P_c value, the better the performance of the algorithm.

4. Absolute Mean Error (MAE) [28] that signifies the equivalence of the spots' intensities

$$MAE_{spot} = \sum_{i=1}^{n} |I_i - \bar{I}| \tag{8}$$

Lower the MAE value better is the reproducibility of the method.

5. Coefficient of Variation (CV), the metric used to heuristically select the segmentation method, defined as [29]

$$CV_{spot} = \frac{\sigma}{\mu} \tag{9}$$

Where σ is the standard deviation of the spot and μ is its mean.

Due to the replication, each spot should have the same intensity throughout the replicated experiments, and therefore the coefficient of variation between replicated experiments should be minimal. This also empirically proves the superiority of the proposed method.

4   Experimental Results and Analysis

To assess the performance of the segmentation stage, we compared the results obtained by proposed Fast EM, Extended K-Mean with existing K-Mean and SPOT software.

Simulated images were produced for assessment of the segmentation performance. Results obtained on the simulated data, the proposed Fast EM and Extended K-means achieved highest r^2 and P_c scores shown in Tables 2, 3 and Fig. 2 compared to conventional K-means and SPOT software techniques at all SNR levels.

In this experiment for one particular SNR, the obtained r^2 values for proposed extended K-means and Fast EM were 0.89 and 0.91 respectively, where as for extended K-means and SPOT software were 0.86 and 0.84 respectively.

Furthermore, the proposed extended K-means and Fast EM and k-means, SOPT software tools were used to segment real microarray images. The algorithms performance was evaluated by the reproducibility of the end results on the green channel and they were quantitatively assessed using two metrics, Absolute Mean Error, and CV (Coefficient of Variation).
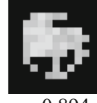
Extracted intensities employing the proposed methods were comparatively evaluated in terms of CV against the intensities obtained by SPOT software and K-Means as shown in Fig. 3. The mode of CV for the two distributions is at 0.110, 0.118 for extended K-Means and Fast EM, where as 0.96 and 0.90 for K-Means and SPOT software.

Results concerning MAE and CV for all 6400 spots are illustrated in Table 4.

Comparative results of segmentation on 3 noisy spots chosen randomly from real images (SMD) are shown in Table 1. The experimental results and analysis, using the performance measures like r^2, P_c, empirically proves the superiority of the proposed techniques against the conventional K-means and SPOT software techniques.

Tables 2 and 3 Summarizes the segmentation performance results of the simulated data for proposed Fast EM, Extended-K-Mean, existing K-mean and SPOT software methods. The coefficient of determination, the concordance correlation achieved for the simulated spots, is presented with respect to different SNR levels as shown in (Fig. 2 (a)–(b)).

Table 1. Comparative results for three different microarray spots, showing estimated fluorescence ratio for each method.

Spot Type	Original Spot	K-Mean	Extended K-Mean	SPOT	Fast EM
Typical Spot		r=0.302	r=0.293	r=0.342	r=0.296
Donut Spot		r=0.341	r=0.302	r=0.427	r=0.266
Spot with Artifacts		r=0.741	r=0.757	r=0.894	r=0.736

Table 2. r^2 calculation in simulated data

SNR (dB)	SPOT	K-mean	Extended K-mean	Fast EM
r^2				
1	0.84	0.86	0.89	0.91
3	0.87	0.89	0.91	0.96
5	0.92	0.93	0.95	0.97
7	0.94	0.95	0.96	0.98
9	0.96	0.97	0.98	0.99

Table 3. P_C calculation in simulated data

SNR (dB)	SPOT	K-mean	Extended K-mean	Fast EM
p_c				
1	0.81	0.84	0.85	0.86
3	0.82	0.85	0.87	0.88
5	0.84	0.85	0.88	0.89
7	0.85	0.86	0.89	0.91
9	0.85	0.87	0.90	0.93

Table 4. Absolute mean error and variation of coefficient results on real microarray images

Methods	MAE	CV
SPOT	600	1.12
K-mean	1300	1.10
Extended K-means	1213	0.96
Fast EM	521	0.90
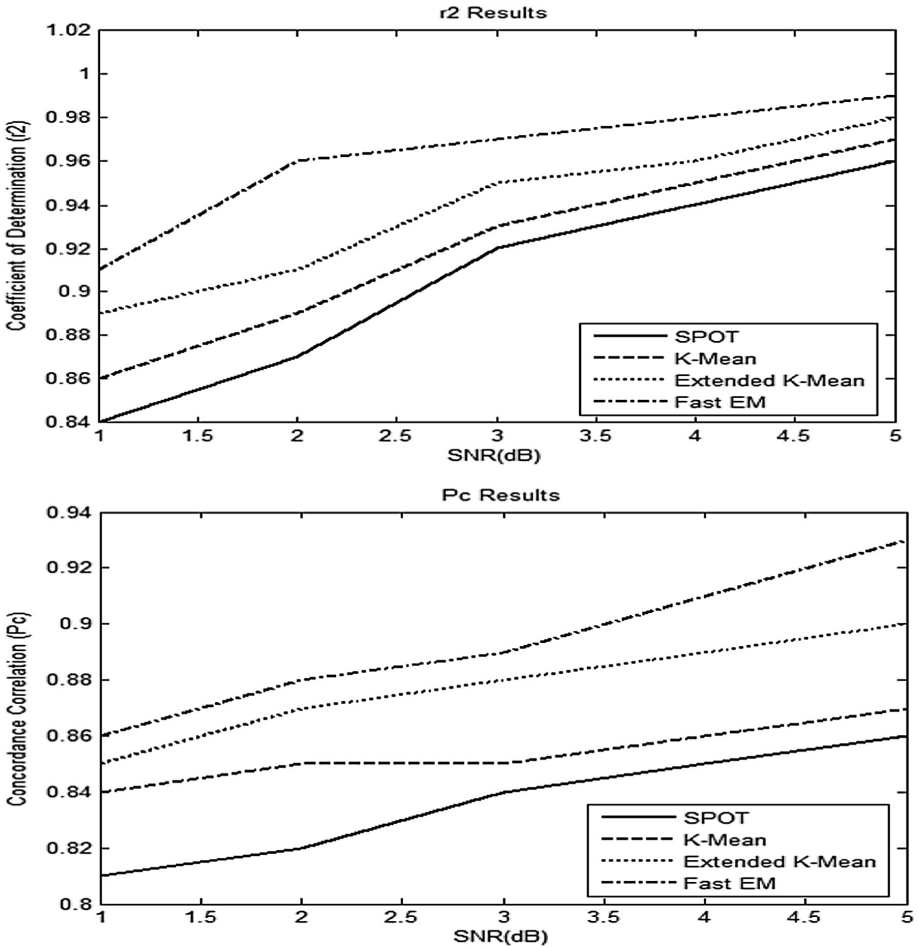
Fig. 2. (a) Coefficient of determination, (b) concordance correlation calculated by using SPOT, K-Mean, Extended K-Mean and Fast EM methods in respect to addictive white Gaussian noise with different SNR.
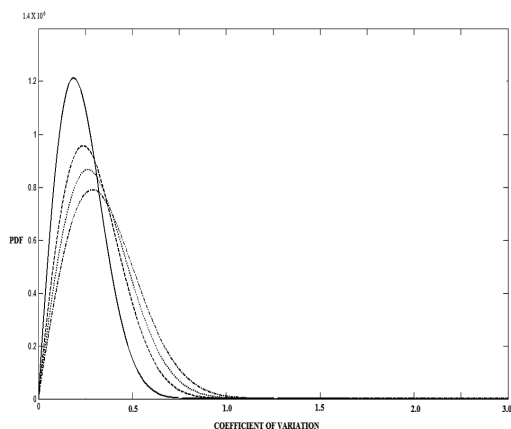
Fig. 3. Probablity density functions (PDF's) of the coefficient of variation for all the spots as calculated from the seven replications of the common reference channel. Dark line corresponds to the results obtained using the Fast EM method. Dashed line, dotted line and the last line corresponds to extended K-Means, K-Means, and SPOT software respectively.

5   Conclusion

In the recent years DNA microarray has attracted with tremendous interest among biologists and has emerged as a powerful tool for genomic research. The proposed segmentation methods Extended K-means, is much reliable in extracting the foreground signal from the microarray spot than existing K-means, SPOT and Fast EM is much faster and noise tolerant than the existing EM. Proposed methods provide accurate result compared to the existing methods. A Heuristic framework was proposed as a novel technique which is based on coefficient of variation as a metric to automatically segment the foreground signal from background in microarray images. The experimental analysis provides promising results compared to the existing ones.

References

1. Introduction to microarray www.cebitac.uni-bielefeld.de/groups/brf/software/emma
2. Lobenhofer, E.K., Bushel, P.R., Afshari, C.A., Hamadeh, H.K.: Progress in the application of DNA microarrays. Environ. Health Perspect. **109**, 881–891 (2001)
3. Yang, Y.H., Buckley, M.I., Dudoit, S., Speed, T.P.: Comparison of methods for image analysis on cDNA microarray data. J. Comput. Graph. Stat. (2002)
4. Ahmed, A.A., Vias, M., Iyer, N.G., Caldas, C., Brenton, J.D.: Microarray segmentation methods significantly influence data precision. Nucleic Acids Res.
5. Axon Instruments: GenePix A User's Guide (1999)
6. Eisen, M.B.: ScanAlyze http://rana.Stanford.EDU/software/ for software and documentation
7. Automatic techniques for gridding cDNA microarray images
8. Smyth, G.K., Yang, Y.H., Speed, T.: Statistical issues in cDNA microarray data analysis functional genomics: methods and protocols. In: Brownstein, M.J., Khodursky, A.B. (eds.) Methods in Molecular Biology Series Totowa. Humana Press, New York City (2002)

9. Buhler, J., Ideker, T., Haynor, D.: Dapple: improved techniques for finding spots on DNA microarrays. Washington UWTR Department of Computer Science and Engineering, University of Washington (2000)
10. GSI Lumonics: QuantArray analysis software, operator's manual (1999)
11. Siddiqui, K.I., Hero, A., Siddiqui, M.: Mathematical morphology applied to spot segmentation and quantification of gene microarray images In: Proceedings of Asilomar Conference on Signals and Systems (2002)
12. Buckley, M.J.: Spot User's Guide. CSIRO Mathematical and Information Sciences, Sydney (2002)
13. Wang, X., Ghosh, S., Guo, S.W.: Quantitative quality control in microarray image processing and data acquisition. Nucleic Acids Res. **29**(15), e75 (2001)
14. ImaGene, ImaGene 6.1 User Manual (2008). http://www.biodiscovery.com/index/papps-webfiles-action
15. Srinark, T., Kambhamettu, C.: A microarray image analysis system based on multiple snakes. J. Biol. Syst. Spec. **12** (2004)
16. Blekas, K., Galatsanos, N.P., Georgiou, I.: An unsupervised artifact correction approach for the analysis of DNA microarray images. In: Proceedings of the IEEE International Conference on Image Processing (ICIP), vol. 2 (2003)
17. Blekas, K., Galatsanos, N., Likas, A., Lagaris, I.E.: Mixture model analysis of DNA microarray images. IEEE Trans. Med. Imaging **24**(7) (2005)
18. Ceccarelli, M., Antoniol, G.: A deformable grid-matching approach for microarray images. IEEE Trans. Image Process. (2006)
19. Demirkaya, O., Asyali, M.H., Shoukri, M.M., Abu-Khabar, K.S.: Segmentation of microarray cDNA spots using MRF-based method. In Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 1 (2003)
20. Noba, H., Shirazi, M.N., Kawaguchi, E.: MRF-based texture segmentation using wavelet decomposed images. Pattern Recogn. **35** (2002)
21. Ergüt, E., Yardimci, Y., Mumcuoglu, E., Konu, O.: Analysis of microarray images using FCM and K-means clustering algorithm. In: Proceedings of International Conference on Signal Processing (2003)
22. Wu, S., Yan, H.: Microarray image processing based on clustering and morphological analysis. In: Proceedings of 1st Asia Pacific Bioinformatics Conference, vol. 2 (2003)
23. Bozinov, D., Rahnenführer, J.: Unsupervised technique for robust target separation and analysis of DNA microarray spots through adaptive pixel clustering. Bioinformatics (2002)
24. Rahnenführer, J., Bozinov, D.: Hybrid clustering for microarray image analysis combining intensity and shape features. BMC Bioinf. (2004)
25. Abbaspour, M., Abugharbieh, R., Podder, M., Tripp, B.W., Tebbutt, S.J.: Hybrid spot segmentation in four-channel microarray genotyping image data. In: IEEE International Symposium on Signal Processing and Information Technology (2006)
26. Dempster, A., Laird, N., Rubin. D.: Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Stat. Soc. **39**(1) (2006)
27. Lehmussola, A., Ruusuvuori, P., Yli-Harja, O.: Evaluating the performance of microarray segmentation algorithms. Bioinformatics (2006)
28. Athanasiadis, E.I., Cavouras, D.A., Spyridonos, P.P., Glotsos, D.T., Kalatzis, I.K., Nikiforidis, G.C.: Complementary DNA microarray image processing based on the Fuzzy Gaussian mixture model. IEEE Trans. Inf Technol. Biomed. **13**(4), 419–425 (2009)
29. Demirkaya, O., Asyali, M.H., Shoukri, M.M.: Segmentation of cDNA microarray spots using markov random field modeling. Bioinformatics **21**(13), 2994–3000 (2005)

Filter Based Approach for Automated Detection of Candidate Lung Nodules in 3D Computed Tomography Images

K. Bhavanishankar[(⊠)] and M. V. Sudhamani

RNS Institute of Technology, Bengaluru, India
bsharsh@gmail.com, mvsudha_raj@hotmail.com

Abstract. Cancer is considered to be one of the threat causing disease to human health. Amongst the various types of cancers, the one which originates in lung is most fatal. Lung cancer appears in the form of nodules and is caused by anomalous growth of cells in the lung organ. Though detecting and diagnosing lung cancer at the early stages can enhance the survival rate, there are still numerous challenges. Computer Aided Diagnosis (CAD) systems which makes use of various imaging modalities like Computed Tomography (CT) and several image processing techniques, assists the domain experts in the process of prognosis and treatment. A novel technique to detect the candidate lung nodule in 3D CT imagery is presented in this paper which employs minimum area filters and layer based filters designed to eliminate the irregularities, blood vessels and other nodules like lung artifacts. Finally the spherical and spiculated structures are retained by considering the diameter of components under observation. The algorithms are tested against 15 cases each having an average of 250 slices, obtained from LIDC.

Keywords: Computed Tomography · Thresholding · Lung nodule
Computed Aided Diagnosis (CAD)

1 Introduction

Lung cancer is considered to be one of the most harmful diseases, being the dominant cause of fatality among all other types of cancers. As evidence, the American Cancer Society's assessment reveals the shocking fact that in 2016 there was 2,24,390 new cases of lung cancer being reported and 1,58,080 deaths. Lung cancer appears as a Solitary Pulmonary Nodule (SPN) in its primary stages with almost spherical shape having a diameter not more than 30 mm in the lung organ. The nodule in the organ can be benign or malignant. As the size of nodule increases so does the degree of malignancy. All though unfortunate, yet it is a fact that about 80% of the patients are diagnosed correctly only when the cancer has surpassed the curable stages which is too late for treatment and the patient will succumb to the disease.

Analyzing the gravity of the disease, the early detection and diagnosis of this lung ailment can enhance the survival rate. But there are factors that make the early detection of the cancer difficult such as (i) The position of the nodule (ii) The nodules

may be masked by the other lung anatomical structures (iii) The difference in size/shape of the nodule among slices (iv) Inferior quality of the image (v) Skewed and fluctuating decision criteria used by the radiologists.

To detect and diagnose the lung nodules there are various imaging modalities available. Among them CT is proved to be the most reliable modality as it is a painless and non-evasive procedure to create precise slices of the lung organ. The imaging in CT scans is based on series of X-ray images generated from different orientations leading to the construction of 3D view of the lung. The 3D view has the advantage of revealing the shape, size, location, density of SPN and the details of other lung artifacts with greater degree of precision. A sample of CT scan with a nodule is shown in the Fig. 1. The biggest challenge faced by the radiologists while interpreting the CT scans is the effort that it demands to manage and analyze the large volume of data in the form of slices. To assist the radiologists, there are computational systems such as CAD system and Computer Aided Diagnosis (CADx) systems, which help in detecting and characterizing the lung lesions.
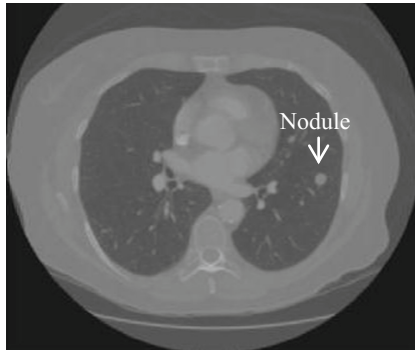
Fig. 1. Sample CT image with a nodule

Given a CT scan, a usual CAD system for lung cancer will have the following processing phases.

(i) Lung Parenchyma Segmentation
(ii) Candidate Nodule Detection
(iii) False Positive Reduction
(iv) Lung Nodule Classification (benign or malignant) leading to a decision making by the domain experts.

This paper presents a systematic procedure to detect the candidate nodules from a given segmented lung parenchyma in a 3D CT image. The segmented lung image is taken as the input and the minimum area filter is applied to all the slices of the image to eliminate small and irregular artifacts. Layer based filtering technique is used to eliminate blood vessels which appear as cylindrical structures in 3D scan. Finally a novel algorithm is developed to retain the spherical and speculated structures (candidate nodules) which define the aim of the presented work.

1.1   Literature Review

In the past few decades CAD systems for cancer in general and lung cancer in particular has attracted the attention of a galaxy of researchers. In [1–3] the extensive advancements in the field of CAD systems for lung cancer have been elaborated. The [4] proposed a method to segment the lung region and detect the lung nodules which was based on multiple region growing and thresholding, based on histogram analysis of an image. In [5] a completely automated methodology to segment the lung parenchyma based on active shape model was proposed. Two stage- CAD system proposed by [6] could detect and diagnose CT images of the lung with benign and malignant nodules. The sensitivity obtained with that method was 90%. Another approach by [7] which extracted textual features using GLCM from volumetric CT scan of the lung followed by classification using these 3D texture features yielded 89.1% sensitivity. A method was proposed to detect the candidate nodules based on active contour model and adaptive thresholding and applied them on a binary image [8].

2   Methodology

The work flow model of the candidate nodule detection scheme is shown in Fig. 2. The scheme has the following processing phases: a minimum area filtering to eliminate the irregularities in the segmented lung image, layer based filtering to eliminate the blood vessels which are cylindrical in shape in 3D view and finally a novel technique to eliminate the non-spherical structures to retain the candidate nodules which are either spherical or speculated in shape. Each of these phases progressively contributes in eliminating the non-nodule artifacts.
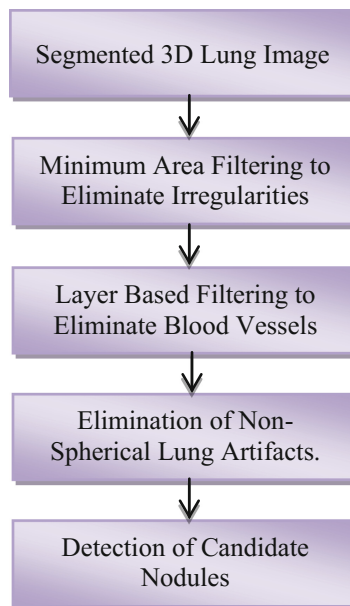
Fig. 2. Work flow diagram for candidate nodule detection system

2.1   Minimum Area Filtering to Eliminate Irregularities

Segmented lung image from the previous work [9] is considered as the input to the candidate nodule detection system. The segmented lung is a binary image which would contain blood vessels, irregularities in scanning, remaining of bronchus, i.e. after the trachea divides in to two branches and nodules.

The irregularities mentioned may appear due to

- Lack of clarity in scanning
- Possibility of slight error in choosing threshold value during the process of segmentation.

The irregularities consist of parts such as blood vessels and bronchi which cannot be processed as single connected components. They can also include noise which gets induced during scanning, which would have been retained even after the process of segmentation. As a first step in eliminating nodule like artifacts, these irregularities need to be eliminated. Minimum area filter is designed to accomplish this task.

The basic idea here is to obtain the area of each component in the segmented lung image and compare it against a threshold value. This threshold value is decided based on a series of experiments conducted; where in the area of irregularities were found to be lesser than the area of nodules. The components that have the area lesser than the computed threshold value are eliminated, leading to elimination of irregularities. It is to be noted that no nodules get eliminated as even the smallest nodule has the area bigger than the threshold value.

The minimum area filter is given in Eq. (1).

$$area(x) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{L} f(i,j,k) \tag{1}$$

where f is a binary image with the component x,
then $\forall x \in \{1, nC\}$, $if area(x) \leq T$, then eliminate the component x,
where nC = number of components,
T = threshold value computed.

This process eliminates only irregularities but not the blood vessels.

2.2   Layer Based Filtering to Eliminate of Blood Vessels

Blood vessels appear like cylindrical structures which run throughout the length of lungs. They may look like disconnected due to some irregularities. Nonetheless a single component of blood vessel will span across many slices of CT scan. These blood vessels can be eliminated by placing a maximum threshold value for the number of slices a component can span across.

The criterion to eliminate these blood vessels is "number of slices the blood vessel spans across". By taking the number of slices as the threshold, it is possible to eliminate the blood vessels. The process of elimination is given by the Eqs. (2) and (3).

$$\sum\nolimits_{z=1}^{L} f(x, z) \tag{2}$$

Where z indicates current slice, L signifies total number of slices and x denotes component number. The component x is eliminated if

$$\sum\nolimits_{z=1}^{L} f(x, z) > T, \ 1 \leq x \leq nC \tag{3}$$

The sample result obtained after applying the above filter is shown in Fig. 3.

| DICOM Image | Segmented Lung | Blood Vessel | Removed blood vessel |
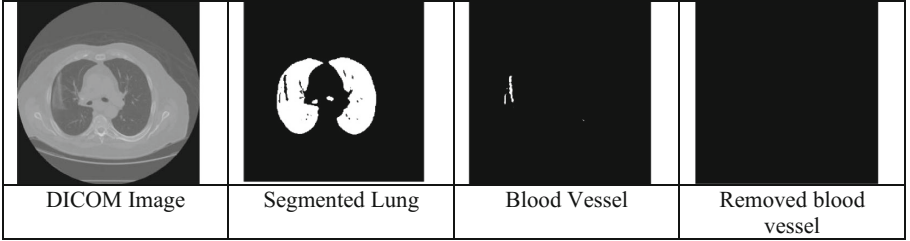
Fig. 3. The result of layer based filtering to eliminate blood vessels

After the above filter operation, the components present in the images are either the nodules or other disconnected blood vessel components. Nodules are either completely spherical or spiculated in shape, depending on the gravity of the cancer i.e. benign or malignant.

2.3 Elimination of Non-spherical Lung Artifacts

If a component is spherical in nature then its diameter across its volume will be almost equal (i.e. diameter across x and y axis). If the diameter of a component across x-axis/y-axis dominates over the diameter across the other dimension, then the component is not spherical and hence not a nodule. Even if the nodule is spiculated the difference in diameter doesn't exceed a certain threshold.

The following algorithm illustrates the computation of the diameter across X axis.

```
Algorithm Compute_Dia_X

    Step 1: Initialize minimum and maximum value of X coordinate to infinity and
    zero respectively.

    Step 2: Find the minimum and maximum value of X coordinate in the first layer
    and store it.

    Step 3: Repeat Step 2 for each layer and compare the obtained values with the
    existing values. If the newer value is a maximum point, store its X coordinate.

    Step 4: Compute X_Dia by taking the difference between maximum and minimum X
    coordinate of the nodule.
```

Similarly the diameter across Y axis, Y_Dia should be computed.

Now, for any component, say q, if *abs (X_Dia-Y_Dia)* is greater than the threshold value, then the particular q is eliminated. i.e. if the computed difference of diameters across X and Y axis exceeds the threshold then the component is not spherical and hence can be eliminated. Using this technique the nodule like structures are retained including the spiculated ones.

3 Experimental Verification

3.1 Results

The algorithms proposed in Sect. 2 (A, B and C) are implemented and are tested on the LIDC data sets whose specification is given in Table 1. The Fig. 4 shows the results obtained for 5 cases with different slices. Each column defines the results obtained for an input DICOM image after applying the successive processing steps as given in Fig. 2. It could be spotted in the last row of Fig. 4, that, among hundreds of components only a very few nodules candidates are retained.
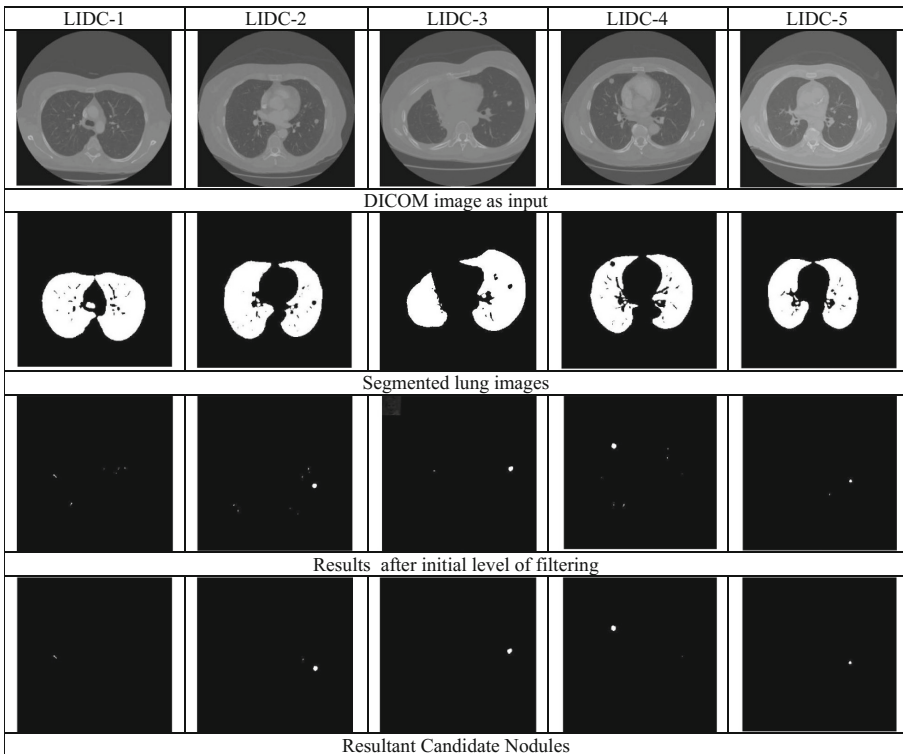
Fig. 4. Results of candidate nodule detection system

3.2   Data Set

The data sets, i.e. CT scans, for the experiments are selected from Lung Image Database Consortium (LIDC) which is available in public domain. Among high resolution CT scans available, 15 cases each having an average of 250 slices are considered for the experiments. The particulars of the data set used are given in the Table 1.

Table 1. Particulars of the CT scans used

Avg. no. of slices	Size	Color type	Bit depth	Scan option	kVp peak kilo voltage
15 * 250	512 * 512	Gray scale	16	Helical	120

3.3   Analysis

Table 2 shows the analysis of the results obtained that are given in Fig. 4. The efficiency of the algorithms was assessed by applying the simple mathematical formula, $\frac{NcB - NcA}{NcA} * 100$, where NcB = Number of components Before and NcA = Number of Components After, the application of algorithms.

Table 2. Analysis of the candidate nodule detection system

LIDC case	# of slices	# components before	# components after	Efficiency w.r.t. reduction of components
LIDC_1	155	249	16	93.5%
LIDC_2	166	299	27	90.9%
LIDC_3	145	510	29	94.3%
LIDC_4	160	394	21	94.6%
LIDC_5	164	355	24	93.2%

4   Conclusion and Discussion

This paper presents novel techniques to filter out the irregularities, blood vessels and other non-spherical lung artifacts. The work was based on analysis of some of the properties of nodules and nodule like structures in the 3D CT scans. When compared to the results in [8, 9], the results obtained in the proposed approach are highly encouraging to move forward in the direction. Though the result reveals the candidate nodules, there are still false positives and eliminating them would be the challenge further. The future work demands to extract the features, namely shape features, intensity features and texture features for the candidate nodules obtained and reduce the false positives to the maximum. An appropriate classifier needs to be designed to ultimately classify the nodule to be either benign or malignant.

References

1. Li, Q.: Recent progress in computer-aided diagnosis of lung nodules on thin-section CT. Comput. Med. Imaging Graph. **31**(4), 248–257 (2007)
2. Ambrosini, V., Nicolini, S., Caroli, P., Nanni, C., Massaro, A., Marzola, M.C., Rubello, D., Fanti, S.: PET/CT imaging in different types of lung cancer: an overview. Eur. J. Radiol. **81** (5), 988–1001 (2012)
3. Bhavanishankar, K., Sudhamani, M.V.: Techniques for detection of solitary pulmonary nodules in human lung and their classifications -a survey. Int. J. Cybern. Inf. (IJCI) **4**(1), 27–40 (2015)
4. Sousa, D.S.: J.R.F., Silva, A.C., de Paiva, A.C., Nunes, R.A.: Methodology for automatic detection of lung nodules in computerized tomography images. Comput. Methods Programs Biomed. **98**(1), 1–14 (2010)
5. Sun, S., Bauer, C., Beichel, R.: Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach. IEEE Trans. Med. Imaging **31**(2), 449–460 (2012)
6. Kumar, S.A., Ramesh, J., Vanathi, P.T., Gunavathi, K.: Robust and automated lung nodule diagnosis from CT images based on fuzzy systems. In: International Conference on Process Automation, Control and Computing (PACC), pp. 1–6. IEEE (2011)
7. Dhara, A.K., Mukhopadhyay, S., Khandelwal, N.: 3D texture analysis of solitary pulmonary nodules using co-occurrence matrix from volumetric lung CT images. In: International Society for Optics and Photonics SPIE Medical Imaging, p. 867039 (2013)
8. Keshani, M., Azimifar, Z., Tajeripour, F., Boostani, R.: Lung nodule segmentation and recognition using SVM classifier and active contour modeling: a complete intelligent system. Comput. Biol. Med. **43**(4), 287–300 (2013)
9. Bhavanishankar, K., Sudhamani, M.V.: 3-D segmentation of lung parenchyma in computed tomography scans. Int. J. Appl. Eng. Res. (IJAER) **86**, 477–481 (2015)

Liveness Detection Based on Eye Flicker

Rekha A. Shidnekoppa[1]([✉]) [iD], Manjunath Kammar[1] [iD],
and K. S. Shreedhar[2] [iD]

[1] Tontadarya College of Engineering, Gadag 582101, Karnataka, India
rshidnekoppa@gmail.com, kammar.manjunath@gmail.com
[2] University BDT College of Engineering, Davanagere 577004, Karnataka, India
ks_shreedhara@yahoo.com

Abstract. Many of the methods has been reported for detection of liveness based on eye flicker are intrusive, in our method detecting liveness using non intrusive. The proposed methodology uses eye flicker activity in an image sequences. The present generic camera will capture the face video with twenty five FPS, frame intervals difference is less than seventy milliseconds. Sequences of images are used in AdaBoost method and Viola Jones technique for the face identification. The Harr like components are figured by convolving the picture with layouts of various size and introduction. Next step is to detect eye region, the position of eye in a face is found on certain geometric dependencies known for human face. If iris is detected in eye region then eye is open. If not, it is closed. Here in this paper we are thinking about eye flicker and iris varieties as a confirmation of liveness to reject utilizing few fake. We tested around more than 20 natural videos to detect the liveness based on eye flicker and achieved 92.50% accuracy.

Keywords: Face detection · Iris recognition · Eye flicker · Haar features
Viola Jones · Liveness detection

1 Introduction

Face acknowledgment framework is an application for distinguishing somebody from picture or recordings. Face acknowledgment is characterized into three phases ie. Face identification, Feature Extraction, Face Recognition. Face identification technique is a troublesome assignment in picture investigation. Face discovery is an application for identifying object, breaking down the face, understanding the confinement of the face and face acknowledgment. It is utilized as a part of numerous applications for new correspondence interface, security and so on. Face Detection is technique for identifying faces from picture or from recordings. The face recognition calculation changes over the information pictures from a camera to parallel example. Figure 1 shows the various stages of face recognition system. Face Detection is in vital progress in the real world applications. Face Detection Technology is terribly vital in many fields like security services [1, 2]. Several different sorts of techniques are there, among these for training Adaboost algorithm is employed by Viola and Jones [3, 4].

Fig. 1. Example for face detection

In these qualities, iris acknowledgment is increasing more consideration in light of the fact that the iris of each individual is remarkable and it never shows signs of change amid a man's life time [5]. Summed up iris acknowledgment comprises of picture obtaining, iris division and limitation (pre-processing). Biometric based individual distinguishing proof requires to exact iris from an eye picture [6] Fig. 2 show example for iris detection.

Fig. 2. Examples for iris detection

Eye flickers are considered has a proof of liveness to reject utilization of fake images, for the most part 2D photos, to parody confront acknowledgment frameworks. Eye flicker is an action of fast shutting and opening of an eyelid. The rate of eye flicker varies based on stress, behaviour category, amount of sleep etc. [7, 8].

The greater part of the present face acknowledgment frameworks depend on pictures taken with a nonexclusive camera, since it could be effortlessly incorporated into the current face acknowledgment approach. Liveness detection method is based on eye flicker is presented for avoidance of fake photo. Two databases are assembled: photograph faker and live face video database. The most part for assessment of ability of against satirizing of photos; the second one is utilized to recreate video frauds of information. In Fig. 3 shows the real face videos and fake videos.
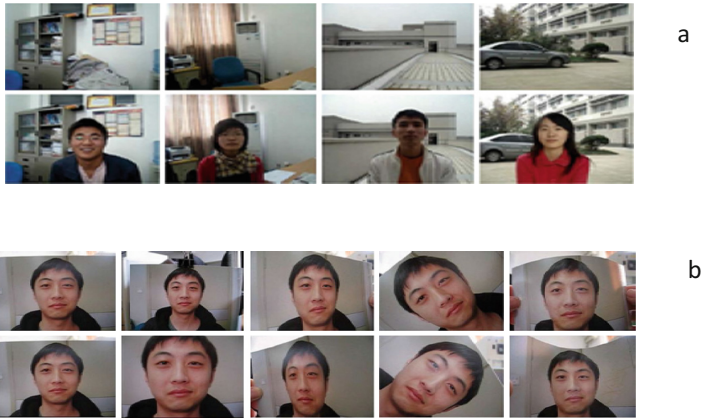
Fig. 3. Samples from (a) real face video database (b) fake video database

2 Related Work

[1] This paper proposes a facial recognition system for automatically identifying or verifying a person from a digital image or a video frame from a video source. Euclidean distance test is used for checking a person's aliveness which ensures the detection of fake/dummy images. [2] This paper proposes fake face detection based on skin elasticity. Biometric system provides a way of automatic verification or identification of a person. But nowadays due to lack of secrecy, there is lot of security threat due to spoofing. [3] This paper proposes a database of iris printouts and its application: development of Liveness detection method for iris recognition. [4] This paper proposes a single image-based face Liveness detection method for discriminating 2-D masks from the live faces. Still images taken from live faces and 2-D masks were found to bear the differences in terms of shape and detailedness.

3 Limitations of the Existing Systems

The increase of attack using fake biometric reduces reliability and security of biometric system, since many of these systems use biometric algorithms which are inefficient to differentiate live biometric characteristics from non-live biometric. In the face recognition system, numerous recognition approaches have been presented, but the efforts on anti spoofing are still very limited. The most common faking way is to use a facial photo of the authenticated user to spoof the face recognition system. It is a challenging task to detect whether the input face image is from a live person or from a photograph, since we need only a generic webcam which could be easily integrated into an existing face recognition system to perform liveness detection with greater performance, efficiency and reliability.

4 Proposed Methodology

In our work a videos are taken from Sony HD 5.3 uber pixel camera, these videos are gathered in various places to make the database and select videos, produce number of frames in a video which is at the very least 25 fps. In pre-handling input shaded picture changes over into gray scale picture. Then resize the image into standard size 320 × 240. The following Fig. 4 shows architecture of proposed methodology.
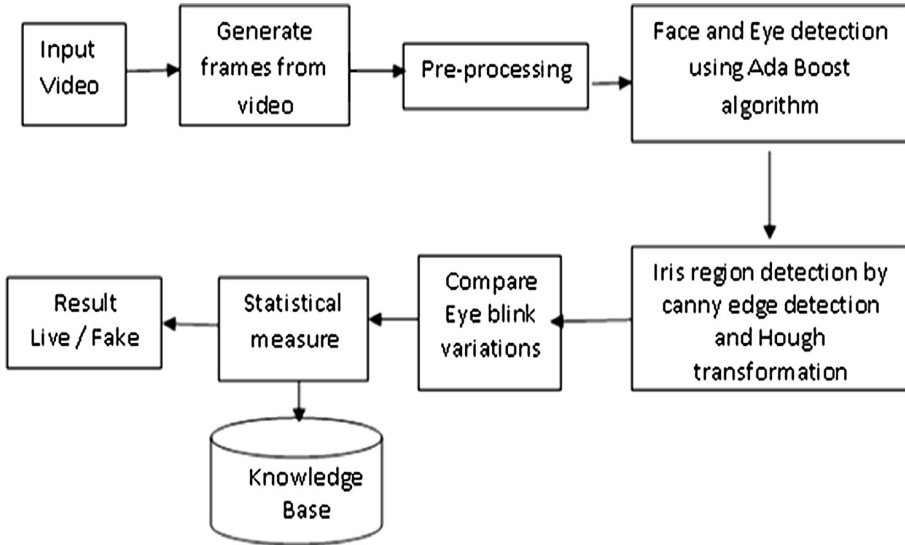
Fig. 4. Architecture of liveness detection

4.1 Face and Eye Detection

The Viola-Jones confront finder contains three fundamental thoughts that make it conceivable to fabricate an effective face locator that can keep running progressively: the Haar-like components, classifier learning with AdaBoost, and the attention course structure. Viola-Jones method depends on investigating the image by methods for sub window equipped for distinguishing highlights. This window is scaled and recognizes face in the frame. Viola Jones built up a image each time with different size. Being scale invariant, the locator need same number of computations paying little heed to the extent of the image. In the main stage comprise of basic identifier which takes out just those windows which don't contain faces. Figure 5 shows integral image.
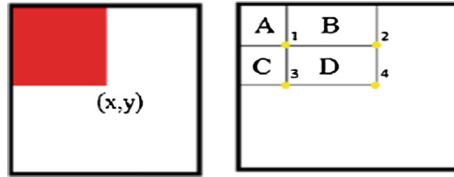
Fig. 5. Integral image construction in Viola Jones

Initial stage Viola Jones calculates the change of the information in a basic image. This is finished by making every pixel equivalent to the whole aggregate of all pixels above and to the concerned pixel. Thusly, whole of all pixels inside any given rectangle can be ascertained utilizing just four esteems. These elements comprise of at least two rectangles. Each component gives a solitary resultant esteem which is computed by subtracting the aggregate of white rectangle from total dark rectangle. Some features are shown in Fig. 6.
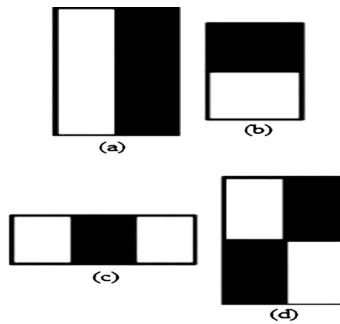
Fig. 6. Haar like features

4.2   Iris Recognition

Once the pre-processing step is achieved, after detecting the eye images we can extract iris structure. In *localization,* the iris inner and outer boundaries are located. *Segmentation* process, partitioning a digital image into multiple regions. *Canny Edge Detection* is Method for noisy images as the bridge gap between strong and weak edges of the image. It connects the weak edges in the output if they are connected to strong edges. The canter of the pupil and its radius can be finding using *Hough transform.*

4.3   Compare Eye Flicker by Iris Code Matching

The two iris code layouts are processing by hamming separation between them utilizing condition.

$$HD = \frac{1}{N} \sum_{j=1}^{N} X_j (XOR) Y_j \tag{1}$$

Xj and Yj are the two iris codes, and N is the number of bits in each template in above Eq. (1).

4.4   Eye Flicker for Liveness Detection

The proposed philosophy is one which utilizes eye flicker action are spoken to by a picture successions, the regular eye states are opening and shutting. The unconstrained resting flicker rate of some person's being is kind of from fifteen to thirty eye squints for each moment. That is, an individual flickers around once each a couple of to four seconds, and furthermore the normal flicker endures in regards to 250 ms. The present non specific camera will basically catch the face video with twenty five FPS.

5   Results and Experiments

In this paper a UI is intended to furnish the menu with the choice of enrolment like begin camera to catch the ongoing video, peruse database for choosing standard face video. Prepare for sparing the face video, begin and confirm Liveness discovery is whether the individual is a live or fake. As shown in Fig. 7.

5.1   Statistical Measurements

Figure 8 Shows ROC graphs are a useful technique for organizing classier and visualizing their performance.

The proposed approach applies two execution measurements, the review rate, and the accuracy rate, for the assessment of live and fake recognition in video. The Table 1 demonstrates the analyses of test outcomes.

Following Fig. 9 shows percentage of live and fake detection of proposed method. Here we tried around 20 pictures, got great genuine positive, negative esteems and furthermore have watched that less false positive and negative esteems. This gives great accuracy rate and review rate in discovery of live and fake.
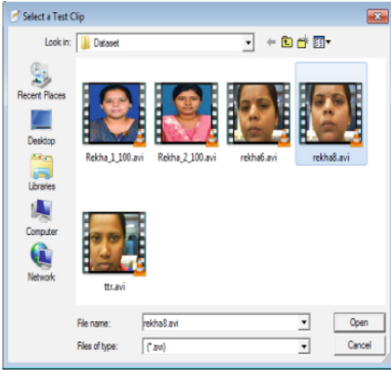
(a). Selection of Live video stored in dataset
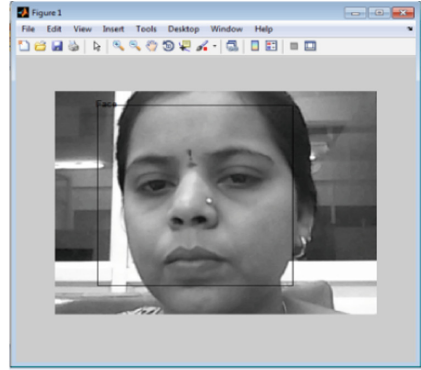
(b). Live face in a video
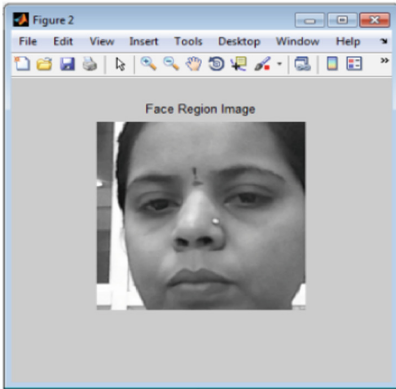
(c). Face region resize standard formats
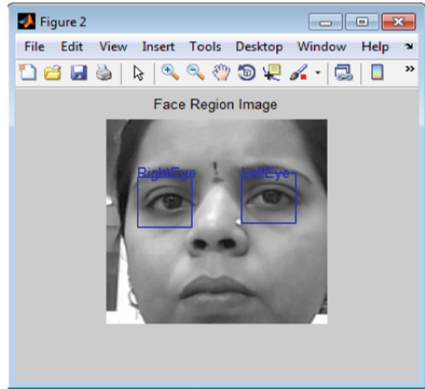
(d). Marking left and right eyes
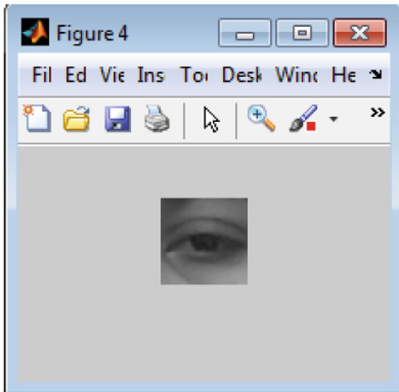
(e). Right eye
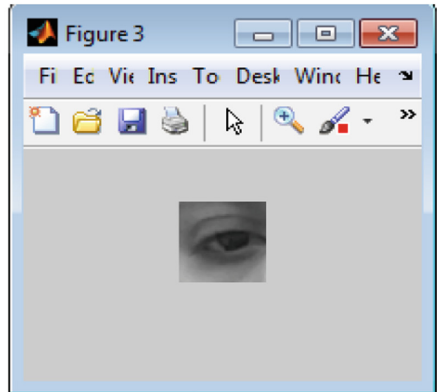
(f). Left eye

Fig. 7. Steps involved in proposed methodology
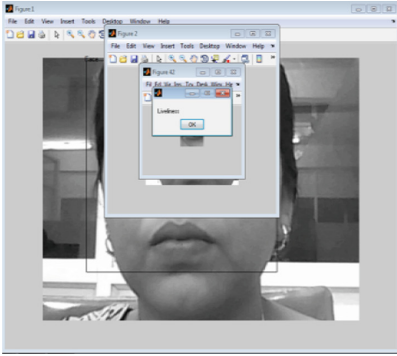
(g). Person in a video is live   (h). Person in a video is fake
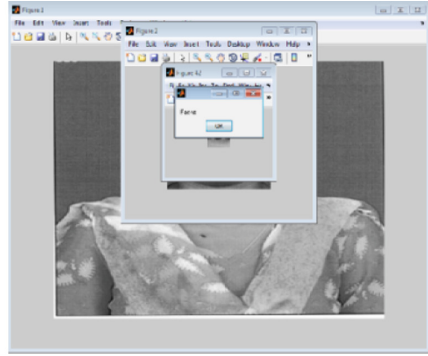
Fig. 7. (*continued*)
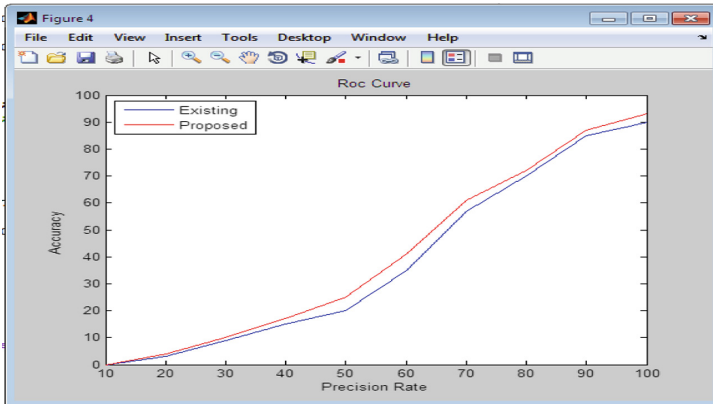
Fig. 8. Receiver operating characteristics (ROC) curve

Table 1. Experimental test results

Number of images/frames studied	20
True positive	18
False positive	2
False negative	1
True negative	19

Accuracy of test result is 92.50%.
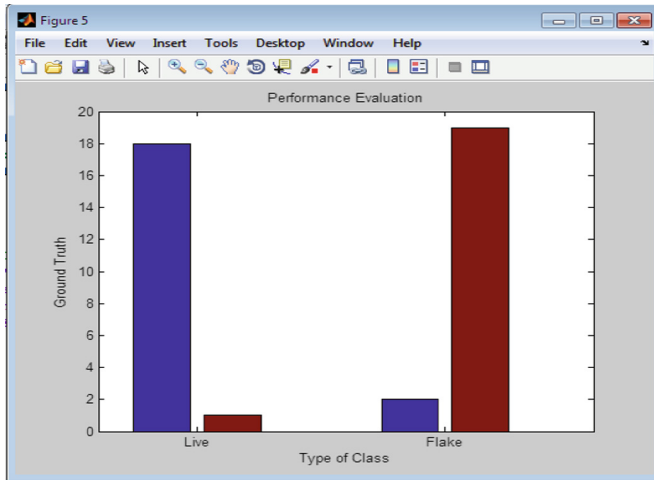
Fig. 9. Performance evaluation

6   Conclusion

In proposed method eye-flicker is used for liveness detection. Haar like features are used to detect the face. Once face is detected successfully, then eye region is extracted from face. Iris is detected from the given eye region. The overall work will be based on Comparing eye flicker variation condition which is useful to find the person liveness. Finally live or fake result is displayed. The overall accuracy obtained using proposed method is 92.50%. This work discuss only about eye flicker to recognize liveness. In future there is need to increase number of frames in video and also attempt has made to implement the work on facial muscle and other face considerations is used to find the person liveness.

References

1. Annu, Kant, C.: Liveness detection in face recognition using euclidean distances. Int. J. Adv. Res. Eng. Technol. **1**, 1–5 (2013)
2. Kant, C., Sharma, N.: Fake face detection based on skin elasticity. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **3**, 1048–1051 (2013)
3. Czajka, A.: Database of iris printouts and its application: development of liveness detection method for iris recognition. Institute of Control and Computation Engineering Warsaw University of Technology, ul. NowowiejskaWarsaw, Poland, pp. 26–29 (2013)
4. Kim, G., Eum, S., Suhr, J.K., Kim, D.I., Park, K.R., Kim, J.: Face liveness detection based on texture and frequency analyses. Research Institute of Automotive Electronics and Control, Hanyang University, Republic of Korea, pp. 1–5 (2012)
5. Bhatt, B.G., Shah, Z.H.: Face feature extraction techniques: a survey. In: National Conference on Recent Trends in Engineering & Technology, pp. 13–14 (2011)

6. Zhang, H., Sun, Z., Tan, T., Wang, J.: Learning hierarchical visual codebook for iris liveness detection. Shanghai Institute of Technical Physics, Chinese Academy of Sciences, pp. 1–5 (2011)
7. Nadimi, S., Bhanu, B.: Physical models for moving shadow and object detection in video. IEEE Trans. Pattern Anal. Mach. Intell. **26**, 1079–1087 (2004)
8. Ni, B., Kassim, A.A., Winkler, S.: A hybrid framework for 3D human motion tracking. IEEE Trans. Circ. Syst. Video Technol. **18**, 1075–1084 (2008)

Influence of Health Service Infrastructure on the Infant Mortality Rate: An Econometric Analysis of Indian States

C. Arun$^{(\boxtimes)}$, Sangita Khare , Deepa Gupta ,
and Amalendu Jyotishi

Amrita Vishwa Vidyapeetham, Bengaluru, India
arun.aradhya2005@gmail.com,
amalendu.jyotishi@gmail.com,
{k_sangita,g_deepa}@blr.amrita.edu

Abstract. This study is intended to identify the influence of health service infrastructure, including drinking water and sanitation provision on the infant mortality rate, thus facilitate the informed spending decision by the government. The study also provides an insight into how the states fares in India, cluster on these parameters with respect to the Infant Mortality Rate. The study involves analysis of the pooled data of each state from the census surveys of 2001, 2011 and the latest NFHS-4 data pertaining to the year 2015–16. In this paper OLS regression and K-Means clustering methods are used to arrive at results. The results indicate that provision of Community Health Centers (CHC) and Sanitation facilities among other variables play a crucial part in reducing the infant mortality rate. The clustering of the variables at three different points of time suggest very few states move out of one cluster to another, indicating very few states move out of their reference status in terms of improvement or deterioration.

Keywords: Infant Mortality Rate (IMR) · NFHS
Econometric analysis · India

1 Introduction

Infant mortality is the term assigned to the death of children, whose age are less than one year, and it is measured using the metric Infant mortality rate (IMR). IMR is defined as the count of deaths of children under the age of one year for every 1000 live births [1]. Infant mortality rate varies across nations based on several factors prevalent in that country, and this can vary based on medical infrastructure, maternal education, prevalent political scenario, environmental and many more conditions. It is also imperative to understand that the infant mortality rate is a direct reflection of the quality of life in these nations. In other words, the quality of life directly impacts the Infant mortality rate. Developed and developing countries show differential IMR that can be attributed to health service infrastructure among other important variables. Thus, it is useful to understand the role of health service infrastructure on IMR.

Infant mortality can be caused due to several factors as mentioned above. Factors affecting IMR can be broadly classified into following groups including medicine and biology, environmental, socio-economic factors, war, birth spacing and education. To make it even more precise the WHO has provided a framework which provides the explanatory variables necessary for a health system as shown in Fig. 1. The framework provides five major categories that include Human health resources, Health service coverage, Medical products vaccines and technology, Health financing, and the Leadership and the governance in that region [2].
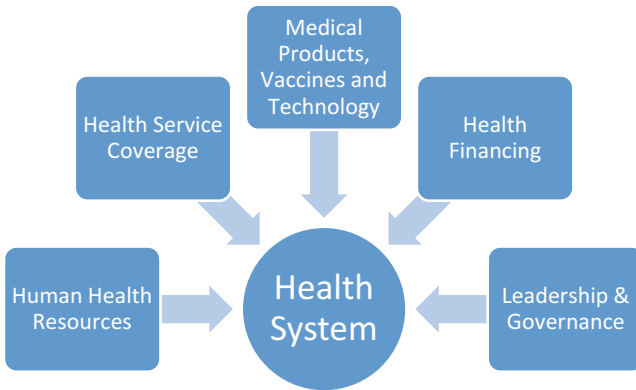
Fig. 1. Factors which influence the health system based on work by Muldoon et al. [2].

The focus of this study is limited to the Health service coverage and its influence on IMR across states at different points of time. In our study we term these characteristics as health infrastructure which also include the basic facilities such as sanitation and drinking water.

2   Related Literature

Raj and Boehmer [3] in their study found that the countries with greater rate of child marriage amongst the girls, resulted in a greater rate of infant and maternal mortality rates. The study was conducted considering the data of 97 countries and using linear and logistics regression models for the study. Saikia et al. [4] in their study have found that gap between the rural region and the urban region has gone up over the time. The main issue here was the disadvantage in the household wealth, maternal education. These two factors have contributed a lot to the infant mortality rate gap, whereas breastfeeding and ORS knowledge has reduced the IMR. Zakir and Wunnava [5] in their study used a cross sectional data set of the year 1993, where the observations are from 117 countries from 6 continents from all low, middle, and high income strata. A regression model is used with the dependent variable IMR and several independent variables as the fertility rates, female literacy rates, government expenditure on health, etc. On tuning the model for heteroscedasticity, it turned out that all factors, other than

expenditure on health programmes, significantly affected infant mortality rates. Muldoon et al. [2] have found that the higher physician density, higher sustainable access to water and sanitation and having a less corrupt government, all leads to significant variation of the IMR.

Swaminathan and Thomas [6] have studied the impact of the political capacity on IMR rates. The political capacity is defined as the taxable capacity of a state. The authors have found that the political capacity of low and middle income states reduces IMR, but the effect weakens as income of the state increases. Most important finding is that the female literacy has a negative effect on IMR in urban areas, however the effect reduces in rural areas. Sahu et al. [7] in their study have found improvements in the reduction of IMR over the period 1992–2006. The factors which were found to be associated with IMR are birth interval, household wealth and the region. The IMR hazard was highest in births from mothers over age of 30 years. Male children were at a risk of higher IMR and the literate women had a lesser risk. The results suggest that the mother's age at first birth, and spacing between births are equally important. Gupta et al. [8] used a Cox proportional hazard model to investigate into the factors leading to infant mortality. They also use a Kaplan-Meir estimation to figure out the survival pattern in the infants. The paper has compared the differences in factors causing IMR between households with a male and a female head and they have found that the size of the household and the count of children below five years of age are the factors for a male headed household and the previous birth interval is found to be the major risk factor of an infant death in a female headed household. The conclusion is that the female headed household have a higher survival probability than a male headed household.

Jain [9] used an analytical framework which is distinguished between levels: village, household and individual. Infant mortality is taken to be the dependent variable and the infant care which is at an individual level and the household and village level factors are taken to be the independent variables. A regression analysis is performed in order to ascertain the extent of the effects of each of the independent variables. The paper concludes simple medical practices such as pregnancy period vaccinations, post-birth vaccinations and proper birth practices can reduce IMR by a great extent. Along with it, it is essential to improve female education and the economic situation prevalent in the society and at each household level, which would in turn increase the effect of preventive medical interventions. Krishna et al. [10] in their study find that there are large inequities in under 5 mortalities across states and also between the economic and the social groups, however there has been some reduction in the socio-economic inequalities over the last 30 years. The factors which are related to the infant and under 5 mortalities are low levels of maternal education, early childbearing, and inadequate birth spacing. The study also throws light on some disturbing facts such as the slower decline of under-five mortality in urban areas, the linkage between the neonatal mortality with maternal obesity. This when coupled with the rapid increase in urbanisation is sure to become very important in evaluating the IMR in the future.

Koenig et al. [11] tried to decipher the effects of domestic violence on the perinatal, neonatal and infant mortality in the rural parts India. They used the NFHS-2 dataset with the birth data over a period of 4 years span for the study. The methodology used was bivariate analysis and hazards regression analysis. The study suggests that the women who had a prior experience of violence had greater risk of infant mortality. Women who had experienced domestic violence just once were not at higher risk of mortality. Chandhiok et al. [12] in their study check whether below par childhood feeding contributes to the child malnutrition and infant mortality. As a part of their study they observe the effect of breastfeeding on the IMR along with other related factors. They use the data from NFHS-1 which is for the period 1992–1993, and NFHS-3 which is for 2005–2006. The methodology used in the study is a multivariate extended cox regression model. The results indicated that the children who were breastfed were at a risk level which was 97% less as compared to those who were not breastfed and thus vouch for the promotion of proper feeding practices to reduce the cases of child under-nutrition and IMR in India.

3   Objectives

The study objective is to assess the impact and effects of health service infrastructure including healthcare facilities, potable water, and sanitation on the infant mortality rate. These independent variables were chosen primarily because it covers most of the major causes of the IMR as prescribed by the WHO framework of explanatory variables for health system. The second part of the study is to cluster the states based on these identified variables at different point in time. It essentially involves clustering the states based on the IMR rates to see how many clusters of states are present and to analyse the movement of states across the clusters over time. This would be useful for a targeted approach at policy and implementation level to address the Infant Mortality Rate [13, 14].

4   Methodology

The first step involved finding the required relevant data across the various timelines. It is indeed very difficult to find the continuity in the variables especially relating to healthcare infrastructure, due to the various agencies involved to carry out the surveys or collect data at varying time points. Our choice of variables hence, was restricted by their availability at different times. The other problem faced was that the surveys conducted by any agency which has the values of IMR had a gap of 10 years. For instance, the Census data contained the IMR values for the years 2001 and 2011. The census data is collected from the entire population thus would represent more accurate value. However, the need for a more recent data led us to the NFHS-4 data set that contained the latest IMR figures of 2015–16 and also contained the other predictor variables which were chosen for the other two time-lines. Thus, the analysis is done using the data sets which comprise of the Census data from 2001 and 2011, and the

NFHS-4 data set which consists of the 2015–16 data. Since the analysis uses the data sets which are sourced from two different types of data collection methods, an assumption that the NFHS-4 data which is a sampled data, truly represents the characteristics of the Census data.

4.1   Independent Variables Selection

Before continuing with the variable selection it is essential to know about the dynamics of the health care facilities in India. The health care system in India has a three layered system consisting of the primary, secondary and tertiary health care facilities. The primary tier is the one which has the three health care institutions and also the one which is considered in our study. The three institutions in the primary care are Sub-Centre (SC) which caters to about 3000–5000 people, a Primary Health Centre (PHC) which caters to 20000 to 30000 people and a Community Health Centre (CHC), one for every four PHCs catering to around 1lakh people. The district hospitals form the secondary layer in the rural region and is the first layer in the urban region and finally the tertiary health care is the health care facilities in urban areas which are modern.

The availability of variables is important criteria for the years taken into consideration. Thus, the common variables which were available for all states, for state-wise analysis relating to healthcare infrastructure are number of community health centres and sub centres, the percentage of households with facilities of potable water and the percentage of household with proper sanitation [15]. The details of the variables considered are shown in Table 1. There were a few more variables including availability of primary health centre for these points of time across states. However, in the final model they did not appear as those did not get through the robustness of the model.

The number of community health centres and the sub centres are taken as the number of centres available for every 100,000 people, this ensures that the size of the states do not provide any bias in the output. The target variables or the variable of interest is IMR, whose mean and standard deviation is also shown in Table 1. Similarly, drinking water and sanitation facilities were factored by respective population of the State or Union Territory. For 2016, the project population for the states were used as provided by the Health Management Information System website and the projections are made based on the component method which is a universally accepted method of making projections and the projections are based on the 2001 census and projections were made in 2006.

4.2   Linear Regression Model

The methodology which has been followed to analyse the data is a regression model which could be either a linear regression model or a nonlinear regression model, however using the MacKinnon, White, Davidson test (MWD test) the model to be used

Table 1. The mean and standard deviation of the variables across all timelines

Variable	Year	Mean	StDev
CHC	2001	0.441	0.271
	2011	0.5318	0.2863
	2016	0.5075	0.261
	overall	0.4947	0.2725
SC	2001	16.19	8.1
	2011	15.2	7.3
	2016	14.75	7.06
	overall	15.356	7.421
Total Water	2001	73.99	20.4
	2011	79.99	17.56
	2016	88.72	7.65
	overall	81.14	16.93
Total Sanitation	2001	49.87	20.78
	2011	60.32	22.49
	2016	57.66	19.52
	overall	56.1	21.18
IMR	2001	48.66	22.4
	2011	35	14.06
	2016	33.03	12.54
	overall	38.62	17.91

was identified to be a linear regression model. The final model was settled with after trial of several iterations and combinations. A linear regression model needs to confirm to some assumptions. The assumptions which are made as per the process are explained here.

Assumption 1: Linear Regression model-the model is linear in parameter, is an assumption which is made and this is validated with the r-square value which indicates the goodness of fit of the model. All the models which has been created in this study have an r-square value of over 0.5, thus indicating a good fit.

Assumption 2: The X values or the independent variables are independent of the error term. This can be verified by finding out the Covariance of independent variables and the residuals. $Cov(X, U) = 0$. This is verified by finding the covariance of the

independent variables along with the residual value. The tests performed indicate that the assumption is satisfied.

Assumption 3: The mean value of the residuals is equal to zero. The study also shows that this condition is satisfied.

Assumption 4: Variance of the error term is equally distributed (Homoscedastic in nature). To verify this assumption, the Parks test is used. The parks test runs a regression which is as shown below.

$$ln\left(u^2\right) = ln\left(\sigma^2\right) + \beta \, ln \, Xi + vi = \alpha + \beta \, ln \, Xi + vi \tag{1}$$

'u' or the residual obtained from the main regression is squared and used in the above equation to regress against the log of independent variables. If the Beta values turn out to be not-significant, then the assumption of homoscedasticity is not rejected. On running the required regression, the Beta values turn out not to be significant thus indicating the homoscedastic nature.

Assumption 5: No Autocorrelation between the disturbances. For this the Durbin Watson Statistic is used. The DW Statistic for the regression model is obtained as 1.94. Thus, it can be assumed that there is no Auto Correlation.

Assumption 6: The nature of X variables (multi-collinearity) which is tested for Variance Inflation Factor (VIF) to arrive at the final model.

Assumption 7: No of observations is greater than the number of parameters. This is true in this case as there are 5 parameters and 92 observations in the data.

5 Results

The model which has been fitted consists of the number of community health centres (CHC), sub centres and the availability of drinking water and also the sanitation. The values in Table 2 indicate the model summary and also the coefficient of the variables and their significance levels. It can be seen that the R square values which is obtained for this model is 51.17%, this indicates that 51.17% of the variations in the IMR can be explained by using the above mentioned variables alone. Also from the p-values of the coefficients it is apparent that at significance levels of 5%, the variables CHC, sub centres and sanitation facilities are significant. Also it is to be noted that the signs of the coefficients of the community health centres, water facility and sanitation facilities are negative whereas the sub centres have a positive coefficient. The complete regression equation for the models is as shown in the equation in the Table 2.

Table 2. The regression model

```
Regression Output:

Model Summary

     S    R-sq  R-sq(adj)  R-sq(pred)
12.7959  51.17%    48.93%      45.28%

Coefficients

Term                 Coef  SE Coef  T-Value  P-Value   VIF
Constant            81.17     9.32     8.71    0.000
CHC                -14.86     6.71    -2.21    0.029  1.86
SC                  0.515    0.248     2.08    0.041  1.88
Total water       -0.1256   0.0839    -1.50    0.138  1.12
Total Sanitation  -0.5868   0.0643    -9.13    0.000  1.03


Regression Equation

IMR = 81.17 - 14.86 CHC + 0.515 SC - 0.1256 Total water -
 0.5868 Total Sanitation


Durbin-Watson Statistic

Durbin-Watson Statistic =  1.93794


Fstat:
Analysis of Variance

Source              DF   Adj SS   Adj MS  F-Value  P-Value
Regression           4  14928.7   3732.2    22.79    0.000
  CHC                1    802.2    802.2     4.90    0.029
  SC                 1    706.2    706.2     4.31    0.041
  Total water        1    366.9    366.9     2.24    0.138
  Total Sanitation   1  13639.2  13639.2    83.30    0.000
Error               87  14245.0    163.7
Total               91  29173.7
```

To find out if there is probability of clusters existing between states based on the similarity of IMR rates and the variables, a k-means cluster analysis was conducted separately for the three points of time viz., 2001, 2011 and 2016. The final number of clusters were chosen based on 'elbow method'. All the three time points provided three clusters with the mean values of the IMR falling in high, medium and low areas.

Table 3. The clusters across which the States are spread in 2001, 2011 and 2016

Cluster 1 : Low IMR	2001	Kerala, Manipur, Meghalaya, Mizoram, Tripura
	2011	Kerala, Manipur, Meghalaya, Mizoram, **Nagaland**, Tripura
	2016	**Andaman & Nicobar Islands, Goa, Haryana, Himachal Pradesh,** Kerala, Mizoram, Nagaland, **Punjab, Sikkim**
Cluster 2: Medium IMR	2001	Andaman & Nicobar Islands, Chandigarh, Daman and Diu, Delhi, Goa, Himachal Pradesh, Jammu & Kashmir, Maharashtra, Puducherry, Punjab, Sikkim, Tamil Nadu, West Bengal
	2011	Andaman & Nicobar Islands, **Andhra Pradesh, Assam**, **Dadra and Nagar Haveli**, Daman and Diu, Delhi, Goa, **Gujarat**, **Haryana**, Himachal Pradesh, Jammu & Kashmir, **Karnataka**, Maharashtra, Puducherry, Punjab, Sikkim, Tamil Nadu, West Bengal
	2016	Andhra Pradesh, Daman and Diu, Delhi, Gujarat, Jammu & Kashmir, Karnataka, Maharashtra, **Meghalaya**, Puducherry, Tamil Nadu, **Tripura**, **Uttarakhand**, West Bengal
Cluster 3: High IMR	2001	Andhra Pradesh, Assam, Bihar, Dadra and Nagar Haveli, Gujarat, Haryana, Karnataka, Madhya Pradesh, Odisha, Rajasthan, Uttar Pradesh
	2011	Bihar, **Chhattisgarh**, **Jharkhand**, Madhya Pradesh, Odisha, Rajasthan, Uttar Pradesh
	2016	**Assam**, Bihar, Chhattisgarh, **Dadra and Nagar Haveli**, Jharkhand, Madhya Pradesh, Odisha, Rajasthan, **Telangana**, Uttar Pradesh

Cluster 1 forms the regions with Low IMR values, and it can also be seen that Kerala has been one such state which has maintained a low IMR throughout. Cluster 2 has a medium range of IMR whereas cluster 3 has a higher mean value of the IMR. Also to be noted is that the households with drinking water and sanitation facilities is the highest in cluster 1 and the lowest in the cluster 3. The states which have been highlighted have shifted the clusters when compared to the base year of 2001. The Table 3 shows the clusters at three points of time. The cluster plots which were obtained during clustering is as shown in Fig. 2.
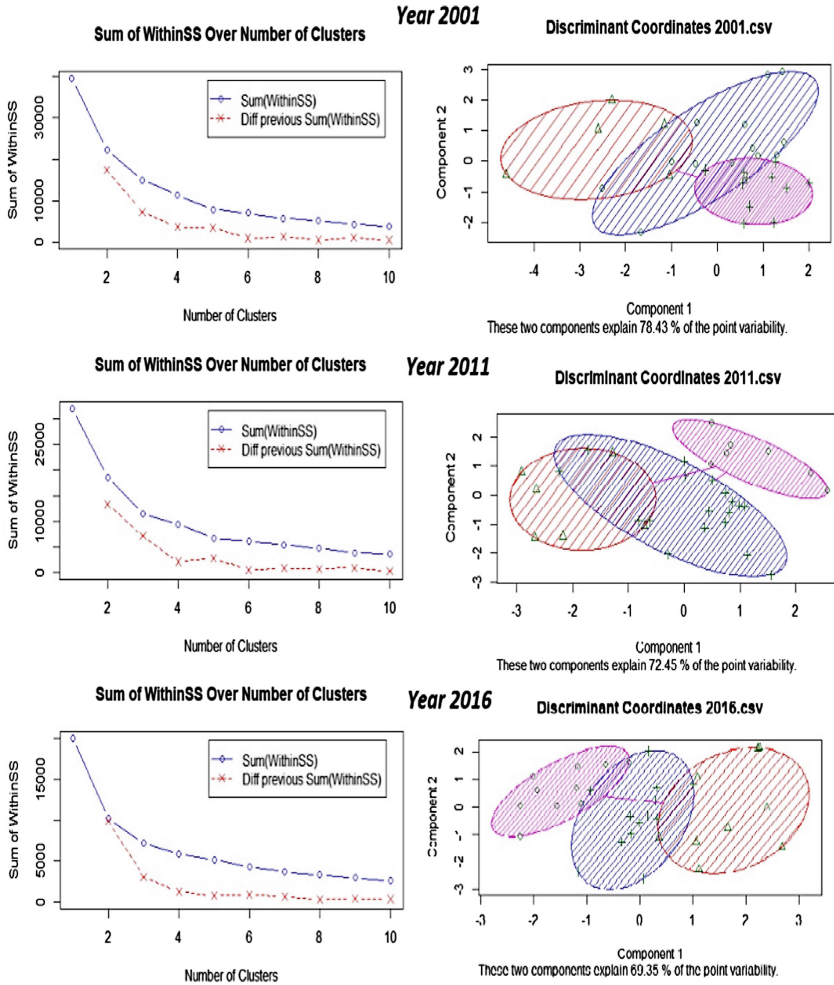
Fig. 2. The elbow plot of 10 cluster iteration and the Discriminant coordinates of the 3 cluster model

6   Discussion and Conclusion

It is evident from the above model that the basic amenities and the socio economic characteristics such as proper healthcare and sanitation facilities do play an important role in the variations of the Infant Mortality rate. When we analyse the results obtained in the model, it is apparent that the number of sub centres has a counter intuitive effect on the infant mortality rate and the IMR increases. This can lead us to make an inference that the quality of the healthcare provided at these facilities is of low quality or perhaps the medical expertise and facilities which are present at these sub-centres are inadequate, although the community health centre plays an important role in reducing

the IMR to a great extent. As discussed earlier the CHC is present for every four PHC and caters to about 1 lakh people, and thus it has better infrastructure and probably better health care personnel thus resulting in better care of the mother and the child during the post pregnancy period, thus making it one of the important factors in the reduction of the infant mortality rate. The other thing which stands important in the model is the percentage of households with sanitation, and it causes a considerable amount of reduction in IMR with even slight percentage increase. Thus it is essential for the government to ensure that proper sanitation facilities are present across the country.

When we look at the clusters of the states based on the IMR and all the other factors considered here, Punjab and other small states have shown decline in IMR whereas Assam and West Bengal have shown increased IMR. Most of the states remained in their respective positions on IMR, though there was overall decline in IMR. Therefore, the cluster 2 and 3 states need more attention especially on these parameters of healthcare infrastructure including drinking water and sanitation facilities to reduce IMR to MDG standard.

References

1. UNICEF Basic Indicators. https://www.unicef.org/infobycountry/stats_popup1.html
2. Muldoon, K.A., Galway, L.P., Nakajima, M., Kanters, S., Hogg, R.S., Bendavid, E., Mills, E.J.: Health system determinants of infant, child and maternal mortality: a cross-sectional study of UN member countries. Glob. Health **7**(1), 42 (2011). https://doi.org/10.1186/1744-8603-7-42
3. Raj, A., Boehmer, U.: Girl child marriage and its association with national rates of HIV, maternal health, and infant mortality across 97 countries. Violence Against Women **19**(4), 536–551 (2013). https://doi.org/10.1177/1077801213487747
4. Saikia, N., Singh, A., Jasilionis, D., Ram, F.: Explaining the rural-urban gap in Infant mortality in India. Demogr. Res. **29**, 473–506 (2013). https://doi.org/10.4054/demres.2013.29.18
5. Zakir, M., Wunnava, P.V.: Factors affecting infant mortality rates: evidence from cross–sectional data. Appl. Econ. Lett. **6**(5), 271–273 (1999). https://doi.org/10.1080/135048599353203
6. Swaminathan, S., Thomas, J.: Saving the next generation: political Capacity and Infant Mortality Decline in India's states. Int. Interact. **33**(3), 217–242 (2007). https://doi.org/10.1080/03050620701449017
7. Sahu, D., Nair, S., Singh, L., Gulati, B.K., Pandey, A.: Levels, trends & predictors of infant & child mortality among Scheduled Tribes in rural India. Indian J. Med. Res. **141**(5), 709 (2015)
8. Gupta, A.K., Borkotoky, M.K., Kumar, A.: Household headship and infant mortality in India: evaluating the determinants and differentials. Int. J. MCH AIDS **3**(1), 44–52 (2015). https://doi.org/10.21106/ijma.37
9. Jain, A.K.: Determinants of regional variations in infant mortality in rural India. Popul. Stud. **39**(3), 407–424 (1985). https://doi.org/10.1080/0032472031000141596
10. Krishna, V.R., Shashidhar, R., Smitha, D.: Infant mortality trends in India: a review of health system. OIDA Int. J. Sustain. Dev. **9**(7), 15–22 (2016)

11. Koenig, M.A., Stephenson, R., Acharya, R., Barrick, L., Ahmed, S., Hindin, M.: Domestic violence and early childhood mortality in rural India: evidence from prospective data. Int. J. Epidemiol. **39**(3), 825–833 (2010). https://doi.org/10.1093/ije/dyq066
12. Chandhiok, N., Singh, L., Singh, K.J., Sahu, D., Pandey, A.: Does breastfeeding have an effect on infant mortality in India? An analysis of national family health survey data. Open J. Prev. Med. **05**(09), 359–369 (2015). https://doi.org/10.4236/ojpm.2015.59040
13. Suriyakala, V., Deepika, M.G., Amalendu, J., Deepa, G.: Factors affecting infant mortality rate in India: an analysis of Indian states. In: Corchado Rodriguez, J., Mitra, S., Thampi, S., El-Alfy, E.S. (eds.) Intelligent Systems Technologies and Applications 2016. AISC, vol. 530, pp. 707–719. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47952-1_57
14. Anilkumar, N.A., Gupta, D., Khare, S., Gopalkrishna, D.M., Jyotishi, A.: Characteristics and causes of malnutrition across Indian states: a cluster analysis based on Indian demographic and health survey data. Paper Presented at the 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Manipal University, Karnataka, India, 13–16 September 2017. https://doi.org/10.1109/icacci.2017.8126158
15. Functioning of Community Health Centres. http://planningcommission.nic.in/reports/peoreport/peo/peo_chc.pdf

Health and Nutritional Status of Children: Survey, Challenges and Directions

Sangita Khare[(✉)] , Deepa Gupta , K. Prabhavathi ,
M. G. Deepika , and Amalendu Jyotishi

Amrita Vishwa Vidyapeetham, Bengaluru, India
{k_sangita, g_deepa}@blr.amrita.edu,
asbblrpk@gmail.com, mgdeepika@gmail.com,
amalendu.jyotishi@gmail.com

Abstract. Globally, there is deep concern about nutrition and health situation of children in the age group under five year. Malnutrition is found to be one of the leading problems in developing nations and is one of the prominent cause of infant mortality and morbidity. It also reflects the socioeconomic well-being of the country. However, literature is not unanimous in terms of definition, causes and methodology used in understanding malnutrition. Through an extant review of literature, this study aims to do a systematic inquiry on characteristics, causes and methodologies to deal with malnutrition in children under age of five. This paper is an attempt to identify these features and techniques in explaining malnutrition and is making a case for usage of artificial intelligence in identifying the causes of malnutrition.

Keywords: Malnutrition · Nutritional status · Datamining · Meta-analysis
IDHS data set

1 Introduction

Nutrition refers to the process of providing the food necessary for health and growth. For a child right balance of carbohydrates, proteins, fats, vitamins and minerals are important to stay healthy. Malnutrition is one of the worldwide health problem specifically in the area of child survival. Malnutrition refers to both undernutrition and overnutrition. Globally, majority of all deaths among children under five are mainly because of undernutrition. Undernutrition enhances the health risks among the children under five and such children are prone to frequent and severe infections and delayed recovery [1]. Malnourished children are more prone to frequent illness like reduced mobility, increased risk of fractures, infections, muscle wasting, weight loss, low energy, more length of stay in hospital, reduced wound healing etc. These are factors which directly affect the social and economic well-being of the country and have the major effect in the developing countries. Therefore, there is an urgency to deal with malnutrition situation. In this context, the Millennium Development Goal (MDG) by WHO has appropriately identified the importance of reducing mortality in children by two third in duration from 1990 to 2015 (MDG 4). Child intelligence level and physical activity depends on nutritional status of the child. The intelligence level of child under

the age of five is very high and it directly affects the child's future brain development. Lack in in-take of food and micronutrients in the early stages of life hampers the growth of the child, affects the cognitive development as an adolescent and makes the person less productive as an adult. Therefore, maternal and child health need early attention. This problem is especially critical in the low and middle income countries [2, 3]. Malnutrition among children under age of five is a serious problem which need to be addressed at root level. There are number of studies on malnutrition in various parts of the world. These studies identify causes of malnutrition using DHS - demographic and health survey data of the respective country. There are different methods used to identify the causes of malnutrition. Besides, different indicators are used to define malnutrition.

This paper, through an extant review of literature, attempts to systematically identify approaches to malnutrition in different studies, factors causing malnutrition, and methods used for their approach. Our attempt would be to identify the prominent factors among the defining factors of malnutrition. We also intend to systematically characterize the causes of malnutrition as identified in the literature. This paper is organized in the following way. Next section identifies the defining factors describing the nutritional status of a child. In subsequent sections factors explaining malnutrition and methodologies explaining malnutrition are identified, followed by discussion and conclusion.

2   Defining Malnutrition

Malnutrition refers to the situation where nutrients are not balanced in diet which means that there are some nutrients which are in excess or some nutrients are lacking or may be in wrong proportion. According to WHO (World Health Organization) deficiency of nutrition can be undernutrition (malnutrition) and overnutrition (mall-nutrition) [4]. This review suggests, there is no single indicator to define malnutrition. Though BMI and stunting (HAZ) are the most often used indicators of malnutrition, there are at least four different indicators which are used in defining malnutrition. However, at the core of these definitions are three important parameters viz., height, weight and age of the person. These definitions are explained below.

2.1   Body Mass Index (BMI)

BMI which quantifies the fat in the body based on two factors weight and height, and is calculated by dividing the weight of body in kilograms by square of the height in unit meter square. BMI can be scaled as $18.5 \leq BMI \leq 24.9$, were the BMI count below 18.5 is considered as underweight [5], in between 18.5 to 24.9 is considered as normal condition, and above 24.9 is overweight [6]. It can also be represented in the form of standard deviation or Z-score of BMI in the ranging of ± 2, which represents the gap between the raw score and the population mean in terms of the standard deviation [7].

2.2 Height for Age Z Score (HAZ)

Low height for a given age is also known as stunting. HAZ is also represented in the form of standard deviation of range ± 2, where if the value is less than −2 then it's a case of stunted, in between −2 to +2 it's a normal state and above +2 is considered as overnutrition.

2.3 Weight for Age Z Score (WAZ)

Low weight for a given age is also known as underweight. WAZ is also represented in the form of standard deviation of range ± 2 and the interpretation is same as HAZ.

2.4 Weight for Height Z Score (WHZ)

Low weight for a given height is also known as wasted. Similar to HAZ and WAZ, WHZ is also represented in the form of standard deviation of range ± 2 and is interpreted in the same way.

3 Factors Explaining Malnutrition: A Review

The effects of malnutrition on human health of all age groups has been the subject of extensive research for several decades. There are various types of study conducted which are focused on malnutrition and child mortality. Health Surveys are done in various parts of the world [8], collecting data from families on various aspects of health, socioeconomic situations, accessibility to basic needs and availability of medical and para-medical services. DHS are series of standardized, nationally representative surveys that have been put in effect in around 70 countries since 1984. These periodic survey provide data relating to malnutrition, apart from various other health indicators. From these data set, one can also draw causal relationship of malnutrition with various other indicators. These indicators can be broadly categorized as demographic, access to basic needs, health indicators including health services indicators, household characteristics, and socioeconomic factors. From the extant review of recent literature using DHS data across countries, possible causal indicators of malnutrition were identified. It is also important to remember that four different definitions are used intermittently to define malnutrition. The literature survey and meta-analysis suggest that 164 explanatory variables were used globally to explain malnutrition. These include mother's education, wealth index, adequate sanitation, sex of the head of the household, does the household has the refrigerator, mother's weight, Antenatal care, Iron supplementation etc. Several approaches have been explored based on statistical analysis. Several biological factors like child's age, mother's height as well as social economic factors such as mother's education and household wealth are identified as important causes of a child's nutritional status [9, 10]. Access to basic services such as water and sanitation also influence the probability of nutritional indicators of children viz., stunted and underweight [11]. Absence of toilet and sanitation facility, garbage disposal and prevalence of major disease in children contribute to malnutrition.

Globally, neonatal mortality (NNM) accounts for more than thirty percent of all the deaths in children under age of five. In absolute terms, this is a staggering number of nearly 11 million deaths. Some studies find that there is association of NNM and maternal background, antenatal care, post-natal care and child birth related characteristics [12]. Mothers' perception on the size of the child is also an important indicator. Among all the deaths under the age of five, around 41 percent accounts for NNM [13]. A study on malnutrition in Kenya among 0–59 month old children, finds its association with selected households, maternal and child characteristics. The main determinants for undernutrition in Kenya include household wealth index, location, maternal education, BMI, sex, age and size at birth of the child [14]. In a study from Senegal for children up to age of 35 months shows the impact of NGO services and public clinics on nutritional status of mother and child [15]. In another study, though marked differences were found in the levels of women's education, availability of water and sanitary facilities, and women's relative decision making power between urban and rural areas, but this was not reflected on nutritional status of the children [16]. According to one study in India, it is found that prolonged breastfeeding and undernutrition are negatively associated with each other [17].

There are studies done for predicting undernutrition among the children under the age of five. Some of the key explanatory factors of undernutrition are identified irrespective of temporal or spatial dimension of the study. According to one study residence, region, mother's education, wealth index of household and child age are major contributors to malnutrition [18]. This model predicts the nutritional status of under five children by analyzing the relationship between two measures of child health- HAZ and WAZ. Physical and cultural determinants of child nutrition including mother's characteristics, household assets and access to public services are identified as important causes of child health [19]. It is also found that malnutrition is directly linked to low birth weight which, in some way is related to mother's nutritional status during pregnancy/preconception period, infection and antenatal care [20]. The main contributing factor to malnutrition according to another study are previous birth interval, size at birth, mother's BMI and parent's education [21]. According to a study on DHS data from 11 countries that examines the association between dietary diversity and HAZ score for children upto two year old, improving diet quality is a critical factor to reduce the incidents of malnutrition [22]. Dietary diversity has more importance for non-breast fed children as they rely on complementary food for all their nutritional requirements. Another interesting study examines the risk factors of malnutrition among children of HIV infected mothers in sub-Saharan Africa, which shows that such children are at lower risk of malnutrition. It is because of more awareness among the people of countries with high HIV prevalence compared to others. It is possibly because the countries in the advanced stages of the HIV epidemic have overcome the expansionary stage and now have proper measures to mitigate the negative impact of HIV/AIDS [23].

In Indian context, very few studies are available to find risk factors of malnutrition and diet planning [24, 25]. There is an emerging nutrition transition for women living in rural and urban communities which affects the malnutrition condition in children. A study in Andhra Pradesh, compares overweight and obese women with those of normal weight and find factors for obesity which includes age, religion and other

socioeconomic and environmental conditions [26]. Another study finds that malnutrition in 6 to 12 month old infants was associated with mothers with history of depression, low intelligence, and low birth weight [27]. The WHO estimates that India is one of the top-ranking countries in case of count of malnourished children, almost double that of Sub Saharan Africa, with critical consequences on mobility, mortality, productivity and economic growth [28]. It also states that situation in India is even worse than Bangladesh, Haiti, Burkina Faso or North Korea. There are 48% of children who are stunted, 20% are wasted and 43% are underweight in India and 7 out of every 10 children under five are undernourished [29]. The 2015 Global Hunger Index (GHI) Report placed India 20th amongst leading countries having serious hunger situation. In South Asian nations, it ranks third after Afghanistan and Pakistan having a GHI score of 29.0 ("serious situation") [30]. As per report published by Child in Need Institute (CINI) around 40% of world's severely malnourished children under five live in India [31]. Close to 50% of all childhood deaths are because of malnutrition, as per UNICEF. In India 48% of children under the age of five are stunted, which is sign of chronic malnutrition. Girl children are more vulnerable than boys to malnutrition and mortality [32].

The incidents of malnutrition is prominent in the states of Uttar Pradesh, Tamil Nadu which reveals that 23% of children are underweight while 25% of children are stunted, in Madhya Pradesh 60% of children are dealing with malnutrition- India's highest number of malnourished children, Jharkhand at 56.5% of malnourished children- India's second highest number of malnourished children followed by Bihar at 55.9% of malnourished children [33]. One of the major causes for malnutrition in India is economic inequality. There is huge imbalance in economic status of various groups of people. The one having low social status does not get diet with proper nutrients, especially women does not get proper quantity of food as well, so new born infants of such women are generally malnourished. Also mothers do not have proper knowledge about feeding children.

According to recent NFHS 4 (2015–16) report the infant mortality rate in India has reduced from 57% to 41% and under five mortality rate has reduced from 74% to 50% from last NFHS 3 (2005–06) survey [34].

From the review of extant literature on malnutrition using Demographic Health Survey datasets, we could identify the important features defining and influencing malnutrition. Figure 1 lists the 164 features identified from literature and the frequency of the feature appearing in number of papers among our shortlisted research papers. In these listed features, some features are considered as defining or dependent features viz., BMI, HAZ, WAZ and WHZ. Other features are independent or explaining factors. Among the explaining features, mother's education is identified as the most important influencing factor explaining malnutrition confirming to this study [35]. Children born to educated women suffer less from malnutrition comparatively. Apart from maternal education few other features like, sex of the child and mother's weight are also having considerably higher frequency when compared to other features. Similarly among the defining factors WHZ is the most used indicator of malnutrition followed by BMI and HAZ.
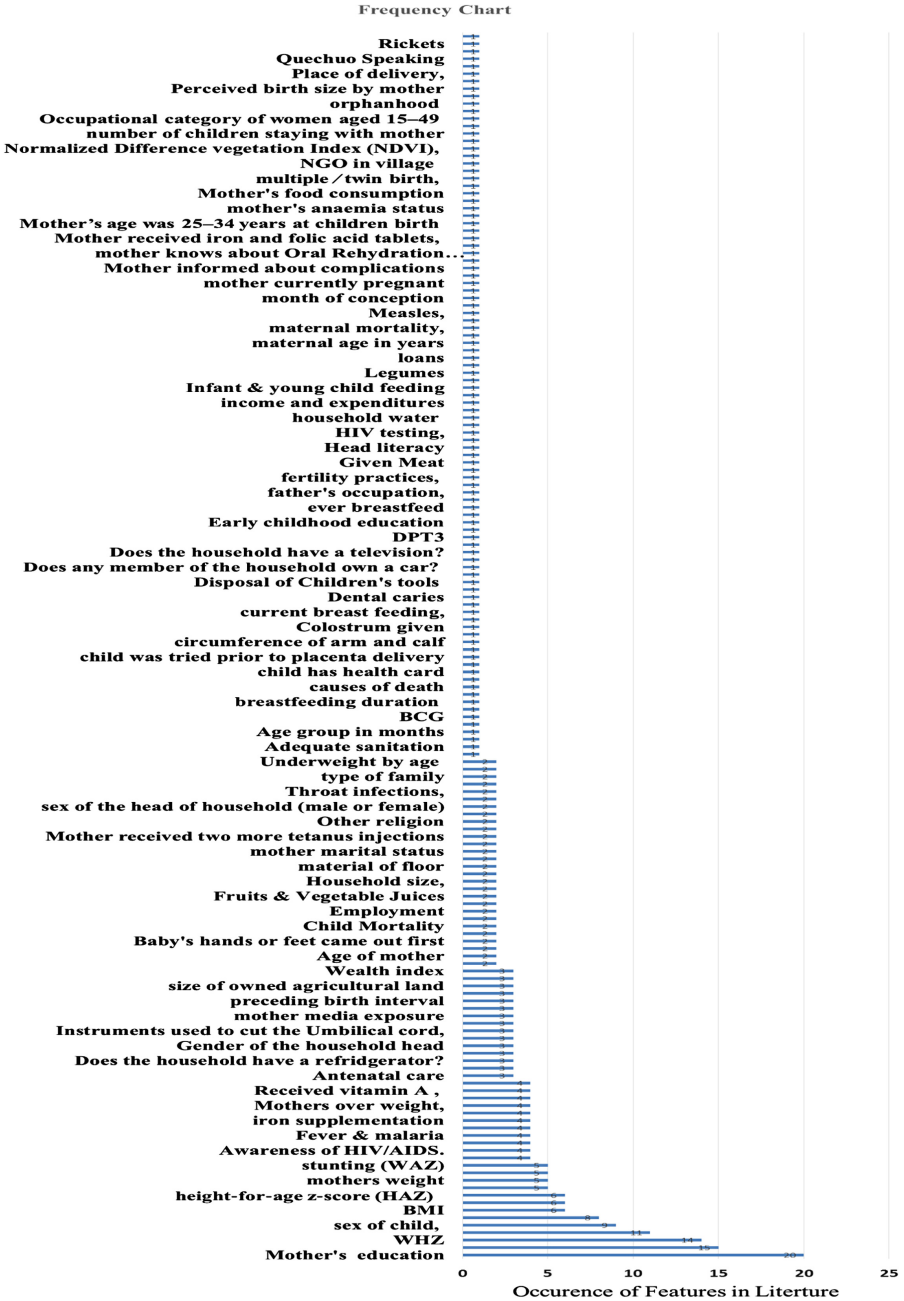
Fig. 1. Frequency chart of list of features identified in the literature

4 Methodology Explaining Malnutrition: A Review

In terms of methodologies that have been used for study of malnutrition, most of the techniques used in past are based on statistical analysis and techniques. Logistic regression and linear regression methods are often used, to analyze the determinants of child nutrition and to find the probability of under-nutrition among children under the age of five [10, 11]. Multivariate regression analysis is used to find determinants of malnutrition, which is based on DHS data from 36 developing countries [16]. Multiple regression technique is used to find association between malnutrition and prolonged breastfeeding in one paper [17]. The report on 0–59 month's old children in Kenya, analyses the determinants of child undernutrition & overnutrition at the descriptive, univariate, bivariate and multivariate levels. The mean z-scores and standard deviations were calculated for the HAZ, WAZ, WHZ, and BMI. Bivariate analysis was done and Pearson's chi square test is used to check for links between malnutrition and household, maternal and child characteristics. Multivariate analysis and logistic regression is used to learn the determinants of malnutrition [14]. In Senegal for children up to age of 35 months, investigation is done to find the impact of NGO services and public clinics for outcomes in terms of nutritional status. It uses ordinary least square method, R square and F-test method for studying nutritional status [15].

OLS (ordinary least square) models were estimated for HAZ and WAZ in a set of six regression models [19]. Regression model is used for study of infant mortality rate [36, 37]. In one of the study conducted on children under age of five in Bangladesh bivariate analysis and multivariate analysis-cox's linear logistic regression model were used to find the detriments of malnutrition. The main contributing factor according to this study are previous birth interval, size at birth, mother's BMI and parent's education [21]. Factor analysis and least square method are used to find the association between the selected features or attributes with malnutrition. This paper takes in account DHS data from 11 countries to examine the association between dietary diversity and HAZ score for children upto two year old [22]. This paper studies association of feeding practices and HAZ using multiple regression analysis and suggests that improved child feeding practice, nutrition education and behaviour change interventions can be useful globally [38]. Multilevel logistic regression is used to examine the risk factors of malnutrition among children whose mothers are infected with HIV in sub-Saharan Africa (SSA) [23]. There are very few attempts made using Machine Learning techniques like Association Rule Mining to find causes of malnutrition [9]. In an study done on under age of five Ethiopian children, data mining approach was used for predictive model building using PART pruned rule induction classifier, Naive Bayes and J48 Decision tree to predict the nutritional status of children [18].

One of the study in Indian context uses additive quintile regression for finding risk factor for childhood malnutrition [24]. Another study uses decision tree ID3 algorithm [25], to understand diet planning for children to deal with malnutrition. Logistic regression technique was used in a few study [26, 27]. Nutritional outcome of children is generally measured in terms of HAZ, WAZ and WHZ. These are assessed using Seemingly Unrelated Regression Model (SUR model) for children under age five in one of the studies in Indian context [39].

Selection of features in earlier literature is largely based on a priori or logical deduction. The models frequently used in the literature are based on statistical analysis method, probability strategy, nearest neighborhood technique, logical analysis, comparison technique, logistic regression, multilevel logistic regression, descriptive analysis, association rules and decision tree. Very few Machine Learning techniques are explored. Table 1. lists the techniques adopted by researchers on different datasets for various aspects of malnutrition using some statistical or machine learning techniques.

Table 1. Methodologies used in different countries for Malnutrition study

	Application/Task	Country (year)	References
1	Method of maximum likelihood, PART pruned rule induction classifier, Naive Bayes and J48 Decision tree	Ethiopia (2000) (2011)	11, 18
2	Association rule mining	Bangladesh, India, Indonesia, Nepal & Philippines (2003-06)	12
3	Analysis at the descriptive, univariate, bivariate and multivariate levels	Kenya (1993–2009)	14
4	Ordinary least square method, R square and F-test method.	Senegal (2004)	15
5	Multiple logistic regression, linear regression	Vietnam (2004)	10
6	Multivariate regression analysis	36 developing countries	16
7	Regression models	Bolivia (1998)	19
8	Multiple regression technique	19 DHS Surveys(1987)	17
9	Regression model	Brazil (1996–97)	36
10	Association rule mining (Apriori Algorithm), Decision tree analysis using J48	Srilanka (2006–07)	9
11	Bivariate analysis and multivariate analysis- cox's linear logistic regression model	Bangladesh (1999–2000)	21
12	Factor analysis and least square method are used to find the association between the selected features or attributes with malnutrition	DHS data from 11 countries	22
13	Multilevel logistic regression	Sub-Saharan Africa	23
14	Regression model	India (2001–11)	37
15	Boosting algorithm for estimation in additive quantile regression model	India (2005-06)	24
16	Decision tree learning algorithm ID3 algorithm	Indian food database	25
17	Logistic regression technique	South India	27
18	Seemingly unrelated regression model	India (NFHS data of 2005–06)	39

5   Discussion

From the extant review of literature, it is found that there are few Machine Learning techniques used for malnutrition study of children under the age of five. Machine Learning techniques are being used in many applications from clinical data analysis to Robotics [40–43]. Therefore, it would be interesting to understand if the artificial intelligence (or Machine Learning techniques) can imitate the a priori features and/or throw up new features that may need attention of the researchers and policy makers. The literature is scanty where the data mining or machine learning techniques have been implemented. The reasoning for choosing explanatory variables are weak and inductive in most of the existing literature. There is limited literature focusing on India especially on nutrition research [44] and hence there is possibility of disconnect between the research findings and chosen policy paths.

It is of critical importance to address the malnutrition situation in Indian context. India being a large country having diversity in physical features and cultural patterns there are variations in nutritional status of children. Different conditions may be responsible for such variations. It is not possible to apply same policy to all the regions in India. There is need to study region-wise nutritional status of children to gain in-depth knowledge of causes [45]. Accordingly, the policies need to be customized for the regions to effectively address malnutrition issues.

There is always ambiguity in identifying the factors that directly or indirectly cause malnutrition. Normally, the organizations collect the huge amount of relevant data about children to identify various nutritional parameters and report national nutritional standing of children, using certain statistical analytical techniques and restricted tools. These limited parameters or attributes which are identified can be improved using technology in a better way, as the technology has the capacity to extract hidden valuable information. Machine Learning is an important tool which can be explored and used for precise specifications or features of malnutrition [46]. Also it is equally important to have knowledge of precise features and their influence on malnutrition to make the policies more effective.

6   Conclusion

As per recently released National Family Health Survey data (NFHS 4) 2015-16 results there are more than 58% of children below age of five that are anemic [47]. This definitely calls for serious steps that need to be taken to address the poor health condition of children in India. There are some studies that do show that infant and child mortality have declined over past few years in India [48], but overall population has also increased. According to this survey, percentage of stunting is reduced by approximately 9.6% as compared to 2005–06 survey results, underweight has reduced by 6.8% but wasted category has shown increase by 1.2%. By all means, it is still not very good situation with current figure for stunting at 38.4% and underweight at 35.7%. If we combine wasted and severely wasted kids count then it comes out to 28.5%. These proportions are reasonably high for any fast growing economy. Malnutrition kills more Indians than any specific disease, yet governments have failed to effectively implement

the two largest programs for addressing malnutrition, the Integrated Child Development Scheme (ICDS) to provide nutritional diets to infants, pregnant and nursing mothers, and the Midday meal scheme to provide 13 crore school children hot cooked food [49]. Governments invest vast sums on treating NCD's (noncommunicable diseases) but overlook the link between them and undernutrition. As this new dataset is released recently it can be analyzed using Machine Learning for current state of the art situation, for effective solution to malnutrition.

In this study various factors that are responsible for malnutrition condition across various countries have been analyzed. The main goal was to find appropriate causes of malnutrition. Broadly 164 features for malnutrition were found, but not all were explored. Most of the studies made use of limited number of features and there is ambiguity in selection of the features. Very few automated Machine Learning techniques were explored. Therefore, this survey based study makes a compelling case for considering Machine Learning techniques to explore various parameters and take appropriate steps to address malnutrition.

References

1. Undernutrition contributes to nearly half of all deaths in children under 5 and is widespread in Asia and Africa. https://data.unicef.org/topic/nutrition/malnutrition/
2. Sinha, K.: First 1,000 days shape health for life. http://timesofindia.indiatimes.com/india/First-1000-days-shape-health-for-life/articleshow/12349751.cms?referral=PM
3. Houweling, T.A.J., Kunst, A.E., Huisman, M., Mackenbac, J.P.: Using relative and absolute measures for monitoring health inequalities: experiences from cross-national analyses on maternal and child health. Int. J. Equity. Health. **6**, 15 (2007). https://doi.org/10.1186/1475-9276-6-15
4. Yach, D., Hawkes, C., Gould, C.L., Hofman, K.J.: The global burden of chronic diseases: overcoming impediments to prevention and control. JAMA **291**(21), 2616–2622 (2004)
5. WHO 1995 Report, http://www.who.int/whr/1995/en/whr95_en.pdf?=1
6. Cole, T.J., Flegal, K.M., Nicholls, D., Jackson, A.A.: Body mass index cut offs to define thinness in children and adolescents. BMJ **335**(7612), 166–167 (2007). https://doi.org/10.1136/bmj.39281.439178.80
7. Statistics Guide. https://statistics.laerd.com/statistical-guides/standard-score.php
8. The DHS Program: Demographic and Health Surveys. http://www.dhsprogram.com/
9. Ariyadasa, S.N., Munasinghe, L.K., Senanayake, S.H.D., Fernando, N.A.S.: Data mining approach to minimize child malnutrition in developing countries. In: The International Conference on Advances in ICT for Emerging Regions (2016)
10. Hien, N.N., Kam, S.: Nutritional status and the characteristics related to malnutrition in children under five years of age in Nghean, Vietnam. J. Prev. Med. Public Health **41**(4), 232–240 (2008). https://doi.org/10.3961/jpmph.2008.41.4.232
11. Environmental factors and children's Malnutrition in Ethiopia, World Bank Policy Research Working Paper 3489 (2005)
12. Fort, A.L., Kothari, M.T., Abderrehim, N.: Association between maternal, birth and newborn characteristics & neonatal mortality in five Asian countries, DHS Working Paper (2008)

13. Oestergaard, M.Z., Inoue, M., Yoshida, S., Mahanani, W.R., Gore, F.M., Cousens, S., et al.: Neonatal mortality levels for 193 Countries in with Trends since 1990: A Systematic Analysis of Progress, Projections, and Priorities. PLOS Med. J. (2011). https://doi.org/10.1371/journal.pmed.1001080

14. Masibo, P.K.: Trends and determinants of malnutrition among children age 0–59 months in Kenya. DHS Working Paper (2013)

15. Linnemayr, S., Anderman, H., Ka, A.: Determinants of malnutrition in senegal: individual, household, community variables, and their interaction. Econ. Hum. Biol. **6**(2), 252–263 (2008)

16. Smith, L.C., Marie, T.R., Ndiaye, A.: Why is child malnutrition lower in Urban than in Rural areas? Evidence from 36 developing countries. World Dev. **33**(8), 1285–1305 (2005)

17. Laura, E.C., Margaret, E.B., Saifuddin, A.: Is prolonged breastfeeding associated with malnutrition? Evidence from nineteen demographic and health surveys. Int. J. Epidemiol. **25**, 693–703 (1996)

18. Markos, Z., Doyore, F., Yifiru, M., Haidar, J.: Predicting under nutrition status of under-five children using data mining techniques: the case of, ethiopian demographic and health survey. J. Health Med. Inf. **5**, 152 (2014). https://doi.org/10.4172/2157-7420.1000152

19. Morales, R., Aguilar, A.M., Calzadilla, A.: Geography and culture matter for malnutrition in Bolivia. Elsevier Journal. Economics and Human Biology **2**(3), 373–389 (2004)

20. Dharmalingam, A., Navaneetham, K., Krishnakumar, C.S.: Nutritional status of mothers and low birth weight in India. Matern. Child Health J. **14**(2), 290–298 (2010). https://doi.org/10.1007/s10995-009-0451-8

21. Rayhan, M.I., Khan, M.S.H.: Factors causing malnutrition among under five children in Bangladesh. Pak. J. Nutr. **5** (2006) https://doi.org/10.3923/pjn.2006.558.562

22. Arimond, M., Ruel, M.T.: Dietary diversity is associated with child nutritional status: evidence from 11 demographic and health survey. J. Nutr. **134**(10), 2579–2585 (2004)

23. Magadi, M.A.: Cross-national analysis of the risk factors of child malnutrition among children made vulnerable by HIV/AIDS in sub-Saharan Africa: evidence from the DHS. Trop. Med. Int. Health **16**(5), 570–578 (2011)

24. Fenske, N., Kneib, T., Hothorn, T.: Identifying risk factors for severe childhood malnutrition by boosting additive quantile regression, Technical report No. 052 (2009)

25. Kalea, A., Autib, N.: Automated menu planning algorithm for children: food recommendation by dietary management system using ID3 for Indian food database. Proc. Comput. Sci. **50**, 197–202 (2015). https://doi.org/10.1016/j.procs.2015.04.070

26. Griffiths, P.L., Bentley, M.E.: The nutrition transition is underway in India. J. Nutr. **131**(10), 2692–2700 (2001). The American Society of Nutritional Sciences

27. Anoop, S., Saravanan, B., Joseph, A., Cherian, A., Jacob, K.: Maternal depression and low maternal intelligence as risk factors for malnutrition in children: a community based case-control study from South India. Arch. Dis. Child. **89**(4), 325–329 (2004). https://doi.org/10.1136/adc.2002.009738

28. World Bank Report on Malnutrition in India (2009). http://siteresources.worldbank.org/INTINDIA/Resources/ChildMalnutrition

29. Arnold, F., Parasuraman, S., Arokiasamy, P., Kothari, M.: Nutrition in India. National Family Health Survey (NFHS-3) India 2005–2006

30. Grebmer, K.V., Bernstein, J., Waal, A.D., Prasai, N., Yin, S., Yohannes, Y.: Global hunger index: armed conflict and the challenge of hunger. http://www.ifpri.org/publication/2015-global-hunger-index-armed-conflict-and-challenge-hunger (2015)

31. Nutrition. http://www.cini.org.uk/nutrition.html

32. NFHS 3 India Fact Sheet (2005–2006). http://rchiips.org/nfhs/nfhs3.shtml

33. Malnutrition in India Statistics State Wise. https://www.savethechildren.in/resource-centre/articles/malnutrition-in-india-statistics-state-wise?gclid=CjwKEAjwl4q-BRDtzJmSk-uGunkSJACmCOYgW5EwggIsCGkTXBqXl0TD1QEsFQx4oEHxt5VlVFNzhoCaiTwwcB

34. NFHS 4: India Fact Sheet (2015–2016). http://rchiips.org/NFHS/pdf/NFHS4/India.pdf

35. Kumchulesi, G.: Can a mother's education effect child's malnutrition? Policy Brief (2010)

36. Alves, D., Belluzzo, W.: Child health and infant mortality in Brazil. IDB Working paper No. 196. http://dx.doi.org/10.2139/ssrn.1814748 (2005)

37. Suriyakala, V., Deepika, M.G., Amalendu, J., Deepa, G.: Factors affecting infant mortality rate in india: an analysis of indian states. In: Corchado Rodriguez, J., Mitra, S., Thampi, S., El-Alfy, E.S. (eds.) Intelligent Systems Technologies and Applications 2016. AISC, vol. 530, pp. 707–719. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47952-1_57

38. Ruel, M.T., Menon, P.: Child feeding practices are associated with child nutritional status in Latin America: innovative uses of the demographic and health surveys. Am. Soc. Nutr. Sci. **132**, 1180–1187 (2002)

39. Naline, G., Viswanathan, B.: Revisiting the detriments of child anthropometric indicators in India using seemingly unrelated regression model. Working Paper (2016)

40. Vinitha, D., Gupta, D., Khare, S.: An effective performance analysis of machine learning techniques for cardiovascular disease. Appl. Med. Inf. **36**, 23–32 (2015)

41. Gupta, D., Khare, S., Aggarwal, A.: A method to predict diagnostic codes for chronic diseases using machine learning techniques. In: International Conference on Computing, Communication and Automation, pp. 281–287 (2016). https://doi.org/10.1109/ccaa.2016.7813730

42. Khare, S., Gupta, D.: Association rule analysis in cardiovascular disease. In: International Conference on Cognitive Computing and Information Processing (2016). https://doi.org/10.1109/ccip.2016.7802881

43. Vinitha, D., Gupta, D., Khare, S., Aggarwal, A.: 2nd Investigation of chronic disease correlation using data mining techniques. In: International Conference on Recent Advances in Engineering and Computational Sciences 7453329 (2015)

44. Khandelwal, S., Seigel, K., Narayan, K.M.: Nutrition research in india underweight, stunted, or wasted? Glob. Heart **8**(2), 131–137 (2013)

45. Akash, N., Gupta, D., Khare, S., Deepika, M.G., Jyotishi, A.: Characteristics and causes of malnutrition across Indian states: a cluster analysis based on Indian demographic and health survey data. In: International Conference on Advances in Computing, Communications and Informatics (2017). https://doi.org/10.1109/ICACCI.2017.8126158

46. Khare, S., Kavyashree, S., Gupta, D., Jyotishi, A.: Investigation of nutritional status of children based on machine learning techniques using Indian demographic and health survey data. In: 7th International Conference on Advances in Computing and Communications, vol. 115, pp. 338–349. Elsevier Procedia Computer Science (2017). https://doi.org/10.1016/j.procs.2017.09.087

47. Times of India Newspaper Article: More than half of under-5 kids in India are anaemic, 6 March 2017

48. Claeson, M., Bos, E.R., Mawji, T., Pathmanathan, I.: Reducing child mortality in India in the new millennium. Bull. World Health Organ. **78**(10), 1192–1199 (2000)

49. Times of India Newspaper Article: Malnutrition kills more Indians than any specific disease, yet successive governments pay scant attention, 24 October 2017

An Empirical Analysis of Machine Learning Classifiers for Clinical Decision Making in Asthma

M. R. Pooja[1(✉)] and M. P. Pushpalatha[2]

[1] Vidyavardhaka College of Engineering, Mysuru, Karnataka, India
poojamr.vvce@gmail.com
[2] Sri Jayachamarajendra College of Engineering, Mysuru, Karnataka, India
mppvin@gmail.com

Abstract. An empirical analysis of the various machine learning classifiers on Asthma data and their performance assessment are presented. As an integral part of preprocessing the data, feature reduction and outlier detection are carried out, following which classifiers are applied to the same to distinguish asthmatic subjects from those of non-asthmatic ones. In cases where the classifiers perform poorly, their performance is boosted via adaptive classifiers. Feature ranking is adopted to reduce the number of features that are considered for the final process of classification while support vector machine with radial basis function is deployed to eliminate outliers. The performance is evaluated via cross validation and random sampling with variable training sizes. The various performance metrics evaluated for the system including precision, recall, F1-measure and classification accuracy indicate good results. The study can be used as a guideline for the effective application of knowledge discovery techniques on clinical data in order to explore hidden knowledge representations.

Keywords: Sampling · Cross validation · Bootstrap · Outliers
Knowledge discovery

1 Introduction

Asthma is basically a chronic respiratory disease which is heterogeneous in nature and covers a number of distinct diseases that often exist as comorbidities along with it. By being heterogeneous, the asthma subjects do not respond in a common way for any treatment and do not show common symptoms which are typical of the disease. The disease if not treated in early stages can manifest itself as a chronic condition which can be detrimental to the lives of the patients. Asthma can easily persist throughout the childhood and even stretch itself into adult hood if not detected at earlier stages and treated. It is often considered as a health problem at the global level affecting individuals across all ages, ethnicities and countries. The persistence of the disease is even higher with increased risk factors, severity level and sensitization. Asthma seen in children often referred to as pediatric asthma is even more of concern as larger section of the society represents school going children who are especially associated with low health related quality of life. Thus there exists a strong need to investigate the factors

underlying the disease and predict the likely outcome of the disease which will further lead to better decisions regarding the prognosis of the disease.

Clinical data sets including those related to asthma and allergy have been to date more or less analyzed using statistical techniques which are traditional in nature. Sophisticated tools that can infer useful information, even when nonlinear relationships exist between the data are required in order to investigate diseases such as Asthma. Knowledge discovery techniques that encompass various learning methods can be deployed to deduce or mine hidden information from the clinical datasets that are often made available following surveys conducted across affected populations. In cases where a larger section of the population represents affected subjects, computerized applications that exist as decision support systems/intelligent systems or predictive models can be helpful as there could be scarcity of physicians at the right time and right place. However this requires the computerized systems to be very accurate as wrong decisions could cost the life of patients and hence the performance of such systems have to be assessed via a number of useful reliable metrics.

Our objective is to analyze the performance of the various classifiers following the pre-processing phase where we deploy suitable feature extraction techniques and also eliminate outliers that could possibly lead to misclassifications. The design choices underlying the different machine learning variants are suitably made so as to result in better classification performance. The performance is assessed using random sampling and cross validation and the results show optimal performance for almost all the classifiers employed in this paper.

The rest of this paper is organized as follows. In Sect. 2, we review the related works in the area of clinical decision making in asthma via machine learning methods. In Sect. 3, we discuss different learning methodologies that are deployed for the asthma dataset under study. In Sect. 4, we present the results obtained by applying the different classifiers on the preprocessed data under both the evaluation schemes including cross validation and random sampling and comment on the performance of the two. Section 5, provides concluding remarks of the study carried out.

2   Related Work

Asthma is a disease related to the lungs and often characterized as a respiratory disease which is chronic in nature described through features such as wheezing, tightness of the chest, coughing and shortness of breath. Asthma can arise at any age but, commonly affected age groups include children and young adults [25]. Around 5 to 10% of children with mild asthma go on to develop severe asthma later in life. Skin allergy may coexist with asthma.

Most studied clinical studies of Asthma have been related to age and atopy [5]. It is also one of the most common illness with an incidence of 10–15% in the pediatric population. Data mining involving predictive analytics is an intelligent means to extract hidden knowledge representations from raw data. Existing work shows that it is very difficult to name a single mining algorithm as the most suitable for the diagnosis and/or prognosis of heterogeneous diseases and as such it would be preferable to use a combination of the best properties of ensemble algorithms that would result in a more

effective outcome. Since the major approach to predictive modelling involves machine learning, the challenges facing the same must be addressed, the major challenges being: typically quantifying the hyper-parameters which characterize the model and transforming the original format in which the data is clinically stored into the format that is compatible for predictive modeling [22]. Visual analytics can be used to facilitate the process of analytical reasoning and easy interpretation of the results obtained by running the machine learning variants [23, 24].

Development of decision support sytems/predictive models in the area of healthcare systems can be an intricate task, followed by a danger of not being accepted by the consumers [7]. Thus, we adopt a strategy which aims to improvise the performance of the system by suitably tailoring the methods according to the requirements of the study. Here, an attempt to explore and adopt existing machine learning variants for aiding the process of clinical decision making has been made. In [8], Edwin et al. speaks on the assessment of a computer aided decision support system for asthma care in a health center, highlighting the following factors: (1) Adoption of decision support systems is largely influenced by subspecialist-specific factors. (2) Aspects of specific situations which are unique should be reflected when planning systems for such conditions. Tu et al. propose an intelligent medical decision support system for identification of the heart disease in the patients to ease the work of the physicians. C4.5 decision tree algorithm was deployed initially to decide this, however since the performance was low, Bagging was used in combination with C4.5 as well as Naïve Bayes algorithms to improve the same. [1]. A similar work where patients were classified into one of the three categories low medium and high to indicate the degree of criticality of the disease from data collected using questionnaires was carried out by Rahman et al. [3]. Prasadl et al. in [2] carried out a performance study in terms of accuracy on the expert system for asthma developed using machine learning approaches including Context sensitive auto-associative memory neural network model (CSAMM), C4.5 algorithm, Bayesian Network, Backpropagation model, Particle Swarm Optimization and concluded that Context sensitive auto-associative memory neural network model for the same outperformed the others.

A Clinical decision support system implemented using machine learning (ML) algorithms was developed by Amaral et al. to help the diagnosis of chronic respiratory diseaes using forced oscillation (FO) measurements by deploying classification algorithms based on Linear Bayes Normal Classifier, Decision Trees, Support Vector Machines (SVM), Artificial Neural Networks (ANN), K nearest neighbor (KNN) and a comparative performance analysis was done in order to the search for the best classifier. KNN, SVM and ANN classifiers were the most adequate, and were concluded as classifiers that can be relied on for very accurate clinical diagnosis [4, 9, 28]. These classifiers exhibited a minimum sensitivity of 87% and a minimum specificity of 94% while Random forests (RF) was used to study relations and interdependencies between features for analysis on genomic data [6].

A variety of approaches for classifying patients into a particular phenotype exists and a diverse collection of techniques and sources of data are used, and good performance is reported on datasets as per the reviews conducted in [10, 27, 28]. However, there is no single system that exploits electronic medical records with all its weaknesses. Out of the techniques deployed commonly, the most common ones include

language processing (NLP)-based techniques, rule based systems, statistical analysis, data mining and machine learning [26, 27]. Computer aided technologies in the form of decision support systems and diagnostic systems are widely employed in these days to make important clinical decisions and assess the outcome of healthcare delivery. These systems necessitates one to have both modeling knowledge and expert experience. Liang et al. in [17] proposed deep learning models that revealed unknown concepts and performed far better compared to conventional learning models.

3   Methodologies

3.1   Data Collection and Sampling

Dryad is a worldwide warehouse of data underlying scientific and medical publications. It is one of the well-organized general-purpose repository. In this paper, we have taken ENFAMS survey data that is made available in the Dryad digital repository [29]. The data was gathered as part of a survey, which was basically a cross-sectional study of sheltered homeless families in the Greater Paris Area in 2013. The dataset includes information on parental characteristics, age, sex, health status, housing characteristics, healthcare utilization, children characteristics, asthma comorbidities, frequency of respiratory symptoms besides general attributes such as employment status, administrative status and education level. Basically the dataset is a collection of information of 801 subjects (including asthmatic and non-asthmatic subjects) with respect to 41 variables.

Data sampling is a technique that lets one study the nature of the data under investigation by looking into a subset of the data that represents the entire data and deduce significant information about the data. Among the various techniques available for sampling we have chosen Bootstrap, as it assures the representative sample to exactly represent the entire dataset because of its nature of resampling. We adopt a deterministic/replicable sampling approach that ensures the quality of the sample drawn at any point of time. This sampling technique has shown better performance when compared to the other sampling variants as can be observed by experimental analysis.

3.2   Pre-processing

The general attributes present in the dataset including employment status, administrative status and education level are discarded as they are irrelevant for the study. The missing attribute values for the samples are imputed by averaging the most frequent values for the same. Further, features are normalized by centering and scaling. The centering is done by median and scaling by standard deviation. Normalizing features is very important whenever we are to compare attributes that have different measurements. It is also a basic requirement for most of the machine learning algorithms including machine learning classifiers that are deployed in this work.

Feature ranking

Chi-square test is deployed for feature ranking as it can test the dependency between individual features and the target feature/attribute. About 8 features from the original feature set constituting nearly 20% of the total feature set size ranked in the order of top priority are chosen as the features in the feature subset. Feature ranking is a very important task as most of the features in the original feature set may be less productive as they do not give sufficient information to proceed with further information mining tasks such as classification or clustering. Also it is most helpful to eliminate redundant features.

Features constituting the feature subset include previous consultation with medical practitioners for respiratory symptoms, use of emergency department/hospitalization for respiratory symptoms, itchy rashes, poor housing condition, and dampness, sanitation in the place of living, length of residence in the same place. These features reveal that asthma symptoms seen to occur in early stages and poor housing conditions contribute strongly to the disease.

Outlier detection

Outliers are data instances that show a marked deviation from the rest of the data, which when not eliminated can possibly lead to wrong interpretation of the data and even result in erroneous predictions of the data that is unseen. Hence it is at most important to see that such data is eliminated from the data that is under study for a cause. It is quite possible that the medical datasets have in them some proportion of outlier data as the data is usually collected manually for questionnaires posed to the interested potential population and documented. Also the data that is maintained as part of electronically entered records may contain errors due to wrong entries. Commonly employed techniques for outlier detection are Single Class Support Vector machine and Covariance estimator. Of the two Single class SVM with RBF is deployed. Generally classifiers can easily go through the process of learning with high dimensional feature space as data constituting the target population may be highly distinguishable in the high dimensional space but not so in the original feature space. Since the support vector machines ideally do not scale aptly for large number of training samples or features, several approximations to the RBF kernel are introduced of which one is deployed here where the function takes a form which makes a mapping of a single vector to a high dimensional vector.

3.3   Classifiers

Logistic regression

The technique that is most opted when the dependent variable is a categorical variable which is binary in nature, i.e., assuming either 0 or 1 at any point of time. Logistic regression is used to interpret data and to infer the degree of dependency between one dependent binary variable and one or more variables which can be basically ordinal, nominal or continuous data. However the technique is most applicable for data that is free of outliers. Since asthma is a disease that is basically heterogeneous in nature

encompassing a variety of symptoms and risk factors, the degree to which the independent variables (asthma risk factors) affect the asthma outcome in the potential population can be easily studied by employing Logistic regression.

Random forest

Random forest technique when employed for classification works by constructing a multitude of decision trees and projects its results as that class which is the mode of all the classes. It usually uses an improvised learning phase where at each split at the candidate level in the learning stage, a random set of features are chosen often called "feature bagging". Features that are selected in many of the B trees represent those that are strong predictors of the target attribute. We have chosen the total number of trees to be 5 and the number of attributes at each split 2. The maximum size of the subset such that no more splits happen is fixed as 5 and the depth of the tree is taken to be 3.

Adaboost

Adaboost, often referred to as adaptive boosting is a technique that is employed in order to strengthen the weak classifiers or boost the performance of the existing classifiers. The technique often encompasses many iterations in order to arrive at a strong learner finally. In every iteration, all attempts are made to rectify the mistakes committed in the previous iterations by classifying the misclassified samples correctly. This leads to higher accuracy of the classifier ultimately. In this paper, Adaboost is deployed as an individual classifier as well as the one which can boost the performance of the weak classifier such as Random forest. Random forest classifier has shown a better performance when used with Adaboost when compared to being used as an individual classifier.

Naive Bayes

Naïve Bayes algorithm works by applying Bayes' theorem with the assumption that the features are totally independent. The assumption however is clearly "naïve" and has no prerequisite or prior information. Naive Bayes model following Bernoulli process has been deployed for the dataset under consideration as the dataset is clearly a collection of binary feature vectors. Every feature is basically a 1 or 0 indicating the presence or absence of the characteristic clearly.

K-NN classifier

K-NN classifier is basically an instance based learning technique which aims at finding the nearest neighbors that are closer to the known instance under study. The parameter 'k' is used to decide the number of neighbors that will be present in the instance space. Here we have chosen Euclidean distance as the metric to evaluate the similarity between the target instance and the other instances. The choice of 'k' is a design decision when deploying nearest neighbor algorithm for classification. The value of 'k' is chosen to be 28, a number that is an approximation to the square root of the number of samples (a total of 801 subjects) available in the dataset, which has resulted in good performance as observed by experimentation, when compared to the other values. Uniform weights are assigned to the instances in the process of exploring the similar instances.

Support Vector Machine

Support vector machines are found to be the best classifiers when the problem under study has data that has to be classified into one of the two groups, i.e., when there are two classes. The classifier tries to draw a clear hyper plane that separates the data points into classes by plotting them on either sides of the plane. The data are however not classified clearly in the original feature space sometimes though in the higher dimensions at feature space they show a clear classification and hence, radial basis function as one of the kernel function is adopted for the support vector machine to achieve a better classification result.

Figure 1 gives the workflow of the entire study. The asthma data is first preprocessed, followed by the application of feature ranking and outlier detection which result in data that is free from noise and consisting of a reduced feature set. The sampled data is then provided as input to the various classifiers, following which we estimate the performance of each one of them and analyze the results through studying ROC plots. In the work flow described below (Fig. 1) for the overall system, the various classifiers are represented as: C1- Naïve Bayes, C2 - SVM, C3 - Random forest, C4 – Adaboost, C5- Logistic regression, C6 - KNN, C3-C4 Random forest-Adaboost.
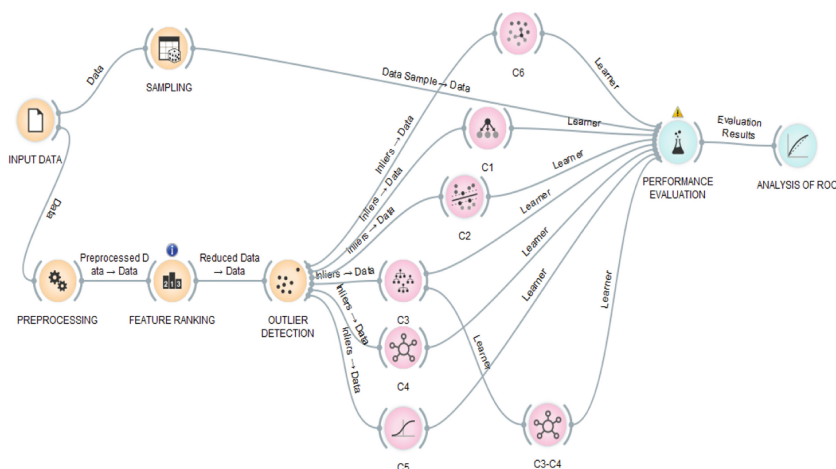
Fig. 1. Work flow of the overall system under study

4   Experimental Results and Discussion

The performance of the classifier is indicated through the metrics such as Precision, Recall, AUC (Area under Curve) and Classification Accuracy which are explained below. At this stage too we choose random sampling with training set sizes as 30%, 60% and 90% of the total data set size and the results obtained are presented below, from which it is clear that good results are expected when there are huge number of samples in the training set. When 90% of the data is used for training optimal results are achieved. The results obtained for variable training sizes under Random sampling is

presented in Tables 1, 2 and 3. Stratified sampling is adopted whereby the entire data in the training set is divided into equal number of groups such that the constituent members in each group share more or less the same properties with respect to their features. A sample from each group is drawn to make up the final random sample. This ensures that the entire population is covered in the process of training. The entire process of training and testing the system is repeated for 10 times. The process is repeated for 10 times and the results obtained is averaged. Leave One Out technique is not deployed because it incurs lot of execution time and results are not generated instantaneously. Cross validation is also employed for testing the performance with the number of folds as 10 following stratified sampling. The results for the same is presented in the Table 4 below. In the tables below as can be observed, the performance of the random trees which is comparatively low compared to the other classifiers, is boosted using Adaboost algorithm. Random forest classifier when used with Adaboost exhibits optimal classification performance, thus demonstrating the concept that ensemble approach can always lead to improved performance. In all the cases, SVM classifier used with RBF kernel function outperforms the other classifiers followed by Adaboost, Logistic Regression and Naïve Bayes techniques which shows good performance, as can be observed from the results presented in the tables below.

Of the several options available for feature ranking, we have depicted the results adopting Chi-square test as the ranking technique since it obtained an optimal feature subset that resulted in a better performance in comparison to the other feature ranking variants namely Information gain, Gain ratio, Gini decrease and Feature Based Collaborative Filtering (FBCF) techniques.

Table 1. Results obtained for Random Sampling with 30% of the total data in the training set

Method	AUC	CA	F1	Precision	Recall
Random forest	0.774	0.822	0.902	0.794	0.822
k-NN	0.567	0.821	0.902	0.674	0.821
SVM	0.666	0.867	0.924	0.859	0.867
Logistic regression	0.906	0.855	0.913	0.848	0.855
Naive Bayes	0.892	0.656	0.736	0.881	0.656
Adaboost	0.777	0.865	0.917	0.867	0.865
RF-adaboost	0.911	0.878	0.928	0.870	0.878

Table 2. Results obtained for Random Sampling with 60% of the total data in the training set

Method	AUC	CA	F1	Precision	Recall
Random forest	0.850	0.823	0.903	0.854	0.823
k-NN	0.551	0.822	0.902	0.712	0.822
SVM	0.812	0.911	0.947	0.907	0.911
Logistic regression	0.928	0.872	0.922	0.872	0.872
Naive Bayes	0.894	0.694	0.772	0.888	0.694
Adaboost	0.858	0.903	0.941	0.909	0.903
RF- adaboost	0.945	0.910	0.946	0.907	0.910

Table 3. Results obtained for Random Sampling with 90% of the total data in the training set

Method	AUC	CA	F1	Precision	Recall
Random forest	0.882	0.823	0.903	0.677	0.823
k-NN	0.614	0.828	0.905	0.858	0.828
SVM	0.897	0.941	0.964	0.940	0.941
Logistic regression	0.949	0.901	0.940	0.902	0.901
Naive Bayes	0.897	0.700	0.777	0.889	0.700
Adaboost	0.896	0.930	0.957	0.933	0.930
RF-adaboost	0.975	0.941	0.964	0.940	0.941

Table 4. Results obtained with 10-fold Cross validation

Method	AUC	CA	F1	Precision	Recall
Random forest	0.859	0.823	0.903	0.854	0.823
k-NN	0.580	0.825	0.904	0.856	0.825
SVM	0.864	0.937	0.962	0.935	0.937
Logistic regression	0.938	0.881	0.929	0.875	0.881
Naive Bayes	0.900	0.714	0.789	0.890	0.714
Adaboost	0.905	0.940	0.964	0.941	0.940
RF-adaboost	0.975	0.933	0.959	0.931	0.933

Study of Performance Plots

Here, we study the performance of the various classifiers through their respective plots namely the ROC plot and Lift Curve. Receiver Operating Characteristic Curve (ROC) is basically a graphical representation of the ability of a binary classifier that gives information on true positive rate and false positive rate. In other words it is used to interpret the sensitivity and specificity of the classifier as threshold varies. Along the x-axis is the false positive rate i.e., 1-specificty, while y-axis indicates the true positive rate i.e., sensitivity. When the data originates by performing multiple iterations of training and testing as in K-fold cross validation, the results are ideally averaged and various averaging options exist of which we have adopted the following two options: 1. Averaged TP and FP at fixed threshold which averages the positions of curves and shows horizontal and vertical confidence intervals. 2. Predictions merged from folds where in, the entire test data is taken as if it comes from a single iteration. Figs. 2 and 3 show the respective plots for the options 1 and 2. Both the figures show plots along with convex hull, where a hull (convex in shape) is plotted combining all the classifiers.

Lift curve is yet another way of visualizing the performance of a classification model. It shows a correlation between the number of instances that are predicted positive and those that are really positive, thus depicting the performance of a target classifier against a random classifier. The same is depicted in Fig. 4 for all the classifiers under study.
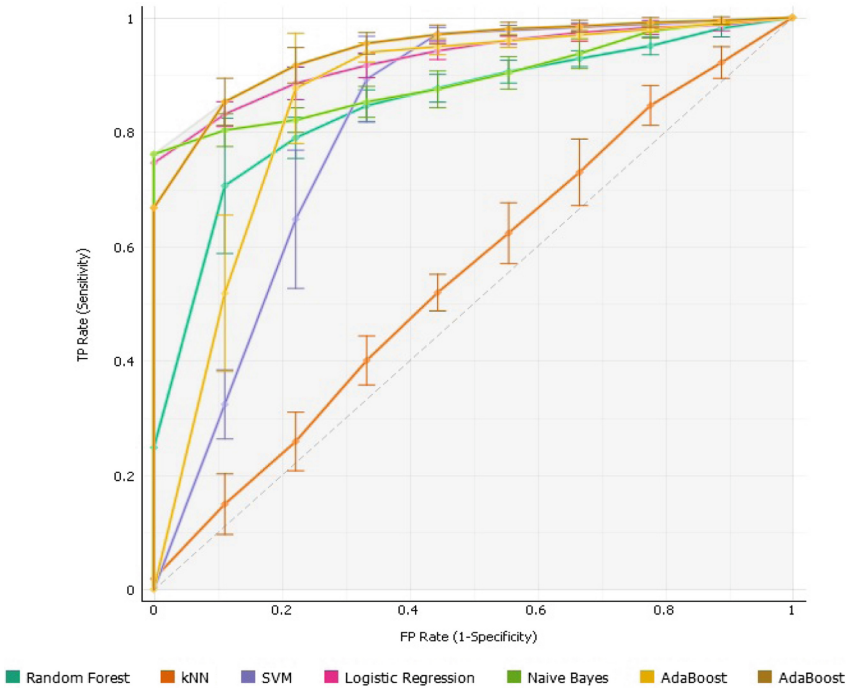
Fig. 2. ROC Curves at Mean TP and FP for various classifiers along with the convex hull
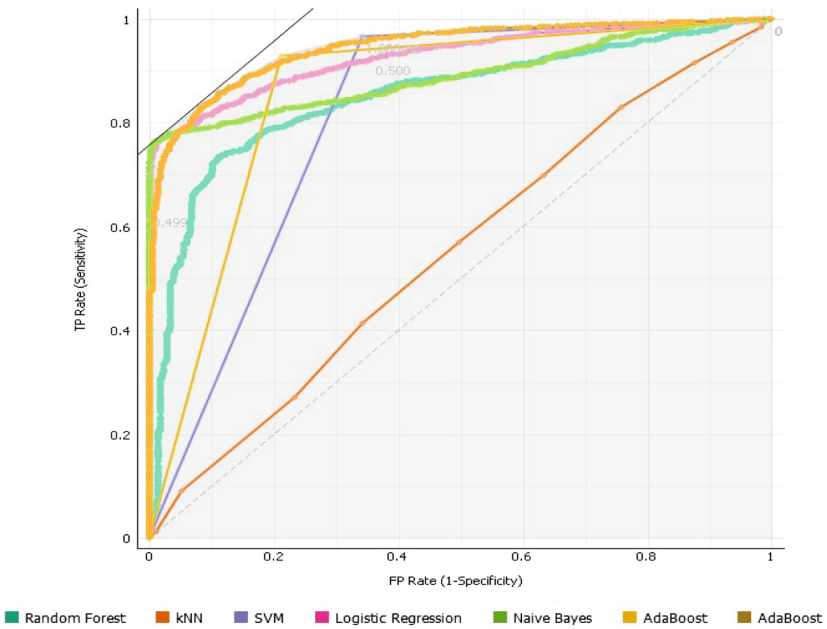
Fig. 3. ROC Curve depicting merged predictions from folds
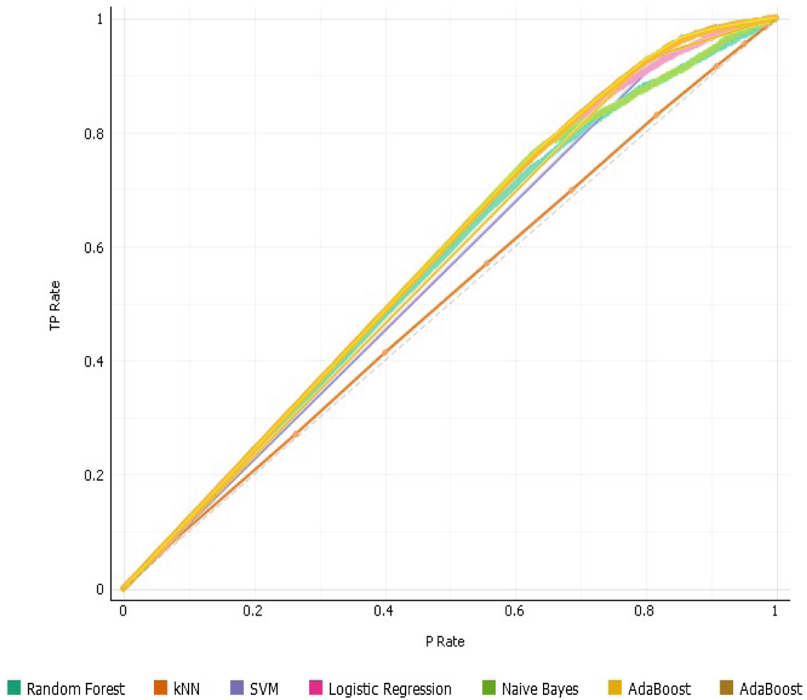
Fig. 4. Lift Curve for various classifiers under study

5   Conclusion

The methods adopted to preprocess the data combined with a suitable choice of the
design parameters for the different classifiers has shown good performance, resulting in
considerably good precision and recall as well as classification accuracy. The perfor-
mance of the Random tree classifier was boosted by employing Adaboost algorithm as
the performance was comparatively low compared to the other classifiers. The tech-
niques adopted for ranking the features and detecting the outliers in the data have
proved to yield good results when classifiers are applied. The classifiers have been used
in an effective way by deploying the learning parameters suitably after thorough
experimental investigations. The performance of the classifiers is estimated via testing
schemes, random sampling and cross validation. Random Sampling with considerably
large proportion of the data in the training set gives optimal results. Overall the wok
can be used as an effective tool to make critical decisions regarding the important
features that attribute to asthma complications and can be integrated into respiratory
care pathways to arrive at suitable conclusions regarding the outcome of the disease.

References

1. Tu, M.C., Shin, D., Shin, D.: A comparative study of medical data classification methods based on decision tree and bagging algorithms. In: Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing. IEEE (2009)
2. Prasadl, B.D.C.N., Krishna Prasad, P.N., Sagar, Y.: An approach to develop expert systems in medical diagnosis using machine learning algorithms (Asthma) and a performance study. Int. J. Soft Comput. (IJSC) **2**, 26–33 (2011)
3. Rahman, R.M., Hasan, F.R.M.: Using and comparing different decision tree classification techniques for mining ICDDR, B Hospital surveillance data. Expert Syst. Appl. **38**(9), 11421–11436 (2011)
4. Jensen, P.B., Jensen, L.J., Brunak, S.: Mining electronic health records: towards better research applications and clinical care. Nat. Rev. Genet. **13**, 395–405 (2012)
5. Saunes, M., et al.: Early eczema and the risk of childhood asthma a prospective population-based study. BMC Pediatr. **12**, 168 (2012)
6. Chen, X., Ishwaran, H.: Random forests for genomic data analysis. Genomics **99**, 323–329 (2012)
7. Bright, T.J., et al.: Effect of clinical decision-support systems systematic review. Ann. Intern. Med. **157**, 29–43 (2012)
8. Lomotan, E.A., et al.: Evaluating the use of a computerized clinical decision support system for asthma by pediatric pulmonologists. Int. J. Med. Inf. **81**, 157–165 (2012)
9. Amaral, J.L.M., et al.: Machine learning algorithms and forced oscillation measurements applied to the automatic identification of chronic obstructive pulmonary disease. Comput. Methods Programs Biomed. **105**, 183–193 (2012)
10. Pathak, J., Kho, A.N., Denny, J.C.: Electronic health records-driven phenotyping: challenges, recent advances, and perspectives. J. Am. Med. Inf. Assoc. **20**, 206–211 (2013)
11. Shivade, C., et al.: A review of approaches to identifying patient phenotype cohorts using electronic health record. J. Am. Med. Inf. Assoc. **21**, 221–230 (2013)
12. Tomar, D., Agarwal, S.: A survey on data mining approaches for healthcare. Int. J. Bio-Sci. Bio-Technol. **5**, 241–266 (2013)
13. Nambiar, R., et al.: A look at challenges and opportunities of big data analytics in healthcare. In: 2013 IEEE International Conference on Big Data (2013)
14. Zięba, M., et al.: Boosted SVM for extracting rules from imbalanced data in application to prediction of the post-operative life expectancy in the lung cancer patients. Appl. Soft Comput. **14**, 99–108 (2014)
15. Velickovski, F., et al.: Clinical decision support systems (CDSS) for preventive management of COPD patients. J. Transl. Med. **12**, S9 (2014)
16. Emanet, N., et al.: A comparative analysis of machine learning methods for classification type decision problems in healthcare. Decis. Anal. **1**, 6 (2014)
17. Wu, W., et al.: Unsupervised phenotyping of severe asthma research program participants using expanded lung data. J. Allergy Clin. Immunol. **133**(5), 1280–1288 (2014)
18. Wang, X., Sontag, D., Wang, F.: Unsupervised learning of disease progression models. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM (2014)
19. Schneeweiss, S.: Learning from big health care data. New Engl. J. Med. **370**, 2161–2163 (2014)
20. Prosperi, M.C.F., et al.: Predicting phenotypes of asthma and eczema with machine learning. BMC Med. Genom. **7**, S7 (2014)

21. Liang, Z., et al.: Deep learning for healthcare decision making with EMRs. In: IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (2014)
22. Luo, G.: MLBCD: a machine learning tool for big clinical data. Health Inf. Sci. Syst. **3**, 3 (2015)
23. Caban, J.J., Gotz, D.: Visual analytics in healthcare–opportunities and research challenges, pp. 260–262 (2015)
24. Amaral, J.L.M., et al.: Machine learning algorithms and forced oscillation measurements to categorize the airway obstruction severity in chronic obstructive pulmonary disease. Comput. Methods Programs Biomed. **11**, 186–197 (2015)
25. Bateman, E.D., et al.: Development and validation of a novel risk score for asthma exacerbations: the risk score for exacerbations. J. Allergy Clin. Immunol. **135**, 1457–1464 (2015)
26. Deo, R.C.: Machine learning in medicine. Circulation **132**, 1920–1930 (2015)
27. Howard, R., et al.: Distinguishing asthma phenotypes using machine learning approaches. Curr. Allergy Asthma Reports, 1–10 (2015)
28. Berner, Eta S., La Lande, Tonya J.: Overview of clinical decision support systems. In: Berner, Eta S. (ed.) Clinical Decision Support Systems. HI, pp. 1–17. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-31913-1_1
29. Lefeuvre, D., Delmas, M., Marguet, C., Chauvin, P., Vandentorren, S.: Asthma-like symptoms in homeless children in the greater Paris area in, prevalence, associated factors and utilization of healthcare services in the ENFAMS survey. PLOS ONE **11**(4), e0153872 (2013). https://doi.org/10.1371/journal.pone.0153872

Web Based Blood Donation Management System (BDMS) and Notifications

B. M. Shashikala[1]([✉]) [ID], M. P. Pushpalatha[1] [ID], and B. Vijaya[2] [ID]

[1] Sri Jayachamarajendra College of Engineering, Mysuru, India
shashibmk@gmail.com, mppvin@gmail.com
[2] JSS Medical College, Mysuru, India
vijayabasavaraj@gmail.com

Abstract. Only 40 lakh units of blood is available out of about 4 crore units of blood in each year in the nation. The direct contact donor and the recipient is not offered by many of the blood banks. This is critical when there is an emergency in need of blood. This aims to provide a direct link between the donor and the recipient. In this paper the blood donation management system (BDMS) which a web based application, which helps to manage the record of donors. Through web interface a blood bank database is created. In a central server the collected data is stored. Mobile application is used for sending notification via text message by using bulk message dispatch technique for providing the communication between the donor and the recipient. The recipient can also contact the donor with the help of the contact details available. Blood donors can also be searched from the mobile application, it is accessible only for registered members. The focus is to find the blood donors in an emergency situation and to provide direct link between the donor and the recipient.

Keywords: Donor · Recipient · Blood donation management system (BDMS)
Notification · Communication

1 Introduction

Blood saves millions of lives across the world in a variety of situations. Millions of blood units are collected from donors and still many millions are required to be collected to meet the requirement of the world [1]. By donating the blood, helps a needy and save a precious life. The information of the blood donor is required when there is need of blood.

The blood bank receives large number of blood from donors. The information about the blood donor needs to be updated. Whenever the information about the donor is required it is difficult for the blood bank authority's people to search the records manually. To solve this problem, we proposed a blood donation management system in which electronic information about the in the blood bank is created for the donor. The main objective is to automate the complete process of the blood donation in the blood bank. At present online databases are available, they did not offer direct link between the donor and the recipient. This is a limitation in many cases when there is a need of blood and this communication is overcome by providing a notification through automated text message technique.

Hospital Information systems provide doctors and support staff timely service. Hospital Information systems need to be developed based on the requirement of the hospital [2, 19]. Most of the existing applications of blood donation system are manual which cannot update and download the latest update of the donor. Retrieval of information takes a lot of time and accuracy is less. Lot of manpower is required [15]. The coordination between the different application and users is less. It takes lot of time to generate reports. Mismanagement in handling the data which gives less security to the system. To debug the existing system a proper approach is required. The system we present provides a lot of information about the donors. The developed system is user friendly and makes the blood donation management much easier and flexible. It provides high level security with authentication [10]. The blood donation management system we present here is built to serve the society. It provides storing, managing, retrieving and searching of blood donor information. This web application allows to access the whole information about blood donation management system and it is adaptable to meet the requirement of blood donation management system in the blood bank. Aim of this paper is to provide direct link between the donor and the recipient via web interface and mobile application short message service.

2   Literature Survey

The system we proposed here is suitable for storing and searching the blood donor information and its saves the time and money. All the blood donation details are automated and with computer system it can be more fast and accurate. This application serves the user by providing the necessary information about the donor in case of emergency. User can effortlessly access the information from the system. The features of the blood donation management system are:

- Centralized database architecture
- Information of blood bank is secured by login and password
- Donor registration
- User access control
- Detailed donor database
- Search facility for finding blood donors
- Easy adding and updating of donors details
- Thorough report formats and registers
- Sends SMS alerts to the donor
- Security to protect blood donor's potential information

The blood donor data gathered from the blood bank located in the city. The administer coordinates and manages various activities involved in the blood bank. Most of the blood bank staff seek information from the blood bank information system other than manually stored records. All the blood donation and transfusion services of acquiring, storing, validating and circulating data is done electronically by a computerized blood bank information system [12–14, 17]. The system provides the top security for blood transfusion service.

Most of the system issues are information consistency, data integrity and reliability of the system. The decisions in blood bank are based on the data. A decision support system was developed for the effective management of blood bank data [3].

To provide a direct link between the donor and the recipient in online blood bank management, algorithm defined. With other parameters algorithm also includes donor willingness and the proximity of the donor. The most eligible donor is found based on the algorithm. The developed system proved most effectiveness in accessibility and security [4, 16]. In health care and blood banking, blood collection is forced to provide efficient operations at a minimum cost [11].

Due to lack of efficient operation, the blood collection team need a working method which demonstrates the efficiency level of cost incurred in the blood collection. The decision support system addresses the existing situation at a blood bank. Decision support system is a method to examine operational efficiency. It facilitates the analysis of data and presents reasonable decision making skill [5]. Inventory management provides adequate information about the availability of blood group. It plays a vital role in the blood bank management information system. It also gives the entire information of the blood bank. In today's world to get the information about blood donation a web based blood bank management tool is required to serve the mankind [6]. All the work carried out in blood bank requires being safe and error free. Blood bank use automation to eliminate errors and provides security in delivering safe blood for transfusion. For blood bank operations a formal workflow model and automated tools are used to verify the safety properties [7].

In short we can say that blood donation management system is an online web application which helps the blood bank and hospitals to look for the blood donor information and to provide direct link between the donor and recipient. This system helps Administrator to check the database. It is very flexible for the blood banks, hospital management and users to retrieve the information and they can have the data according to the query given by them.

3 Methodology

Concept

When donor arrives at the blood bank the staff of the blood bank will collect the information about the donor like name, age, sex contact number and address details of the donor [8]. The staff conducts physical examination of the donor for the eligibility of blood donation. The registration is easy. All the information about the donor and the results of physical examination are uploaded into the system by the administrator. The system provides the unique identification number for every donor at the time of registration. The administrator can add or delete the information of the donor. The administrator can view the total account information of the donor. The search option for donor information is provided. Every blood bank also has a username and password security. In this way this system helps the blood bank to find the information about the registered blood donor.

Every donor has their own e-mail address and password with which they will login into the system. From this they can update their data. If any user request for blood, he can view the willing donor information. A blood donor database is created by collection of detail entered by the donor through interface. In a central server the data is stored.

The most eligible donor is found by algorithm. Willingness of the donor, Location, Last donation date and blood group are the major four components considered. The architecture of the donation process is shown in Fig. 1 given below:
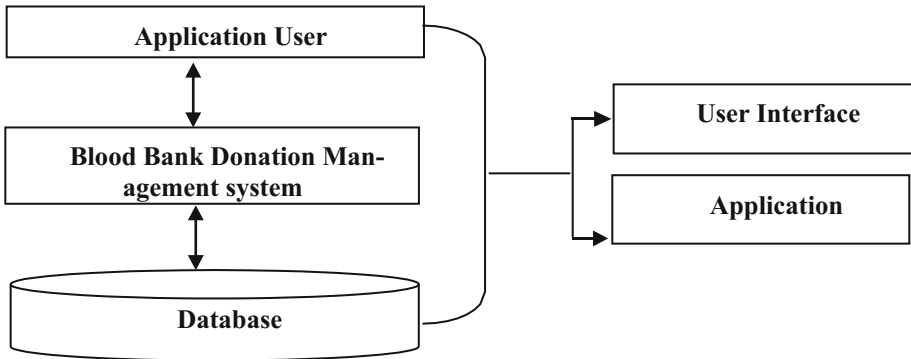
Fig. 1. System architecture for blood donation management system

Method

The proposed system is to create information about the donor and the blood bank related to the blood donation. The storage of the data for the application has been planned. The internal server has been selected as SQL server. SQL server provides the high reliability and security. The user interfaces are designed using JSP technologies. The connectivity to the database is provided by using SQL connection method [18]. The data storage formats are shown in Tables 1, 2 and 3.

The new system manages the present and upcoming requirements of the user. The new system will operate efficiently. The information of all the users stored in a database. The users access the application from any computer that has internet browsing capabilities and an internet connection. Only the available donors are displayed to the Requester. Users with correct usernames and passwords can enter into the system.

4 Analysis

Blood donation management system provides the following characteristics of information.

Information credibility: the donation information of the donor can be obtained to improve the eligibility of blood donation.

Information integrity: The system provides the integrity of the donor's information. Integral information can be built from blood donor, pre-screening results and final blood recipient.

Information interaction: The system promotes the working efficiency of blood bank. Many of operations in the blood bank depend on the blood donor information. Therefore, information interaction plays a vital role in this field.

Information security: Information security plays a vital role in the safety of blood donation and transfusion service. It prevents the unauthorized data operations and provides information recovery.

In this system, donors can create their profile and can update the information. Users can search the donors and can make request for the blood. The request for the blood is sent to the donor via short message service. The donor can view the details of the request. The management of the system is done by the admin. The activity diagram and sequence diagram for admin, donor and requester is show in the Figs. 2, 3, 4, 5, 6 and 7.
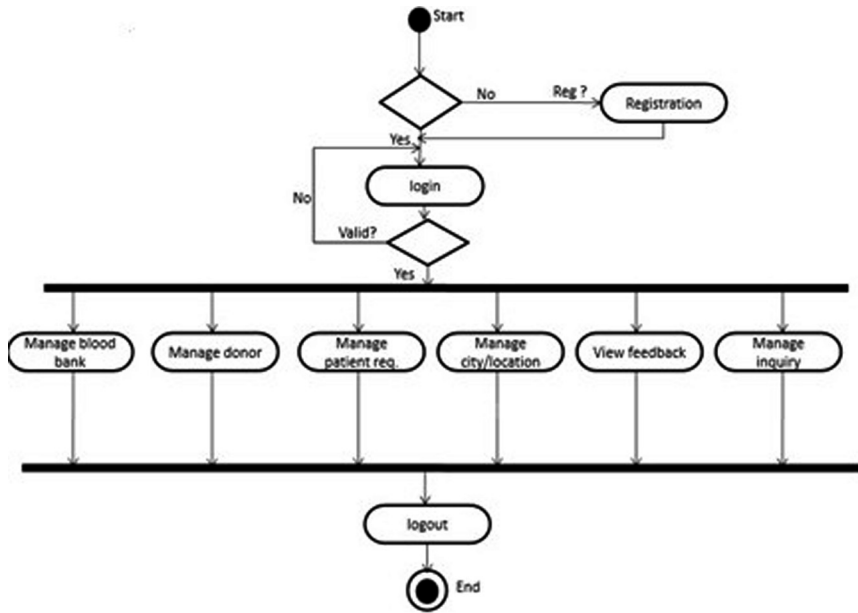
Fig. 2. Activity diagram for ADMIN
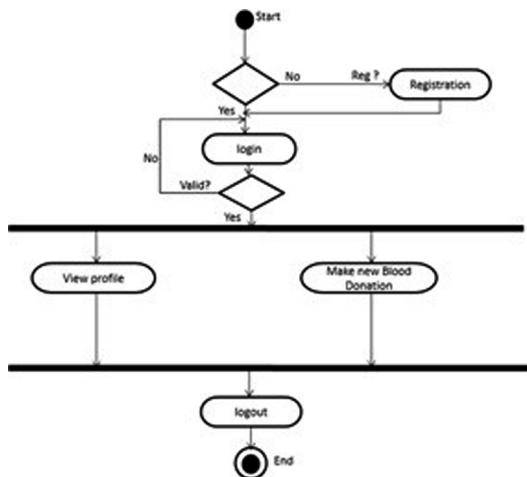
Fig. 3. Activity diagram for DONOR
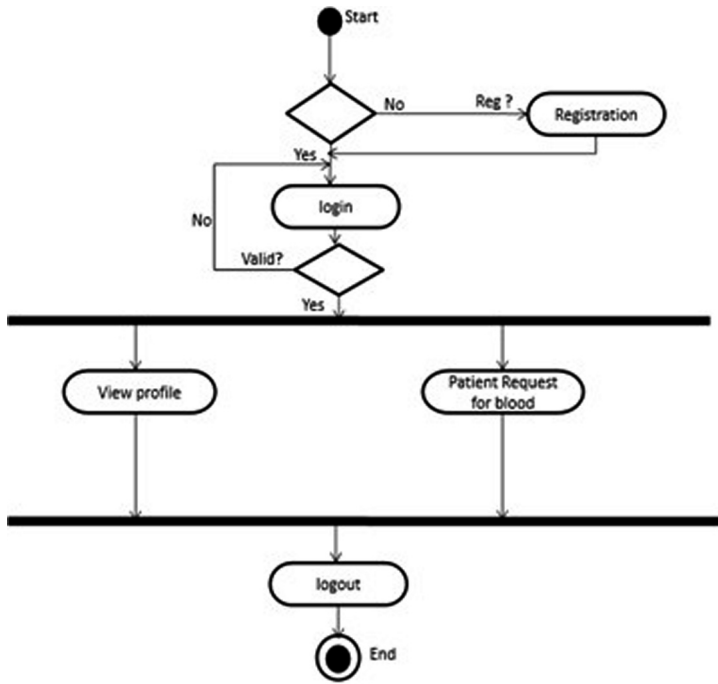
Fig. 4. Activity diagram for Requester/Patient
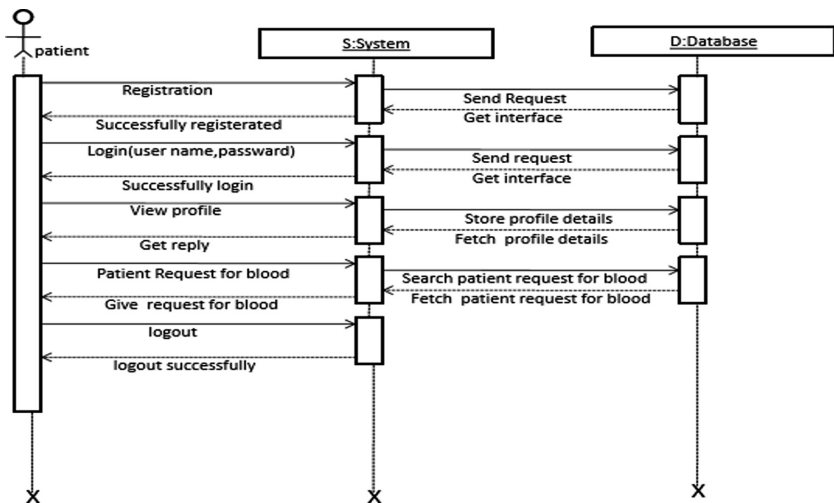
Fig. 5. Sequence diagram for ADMIN
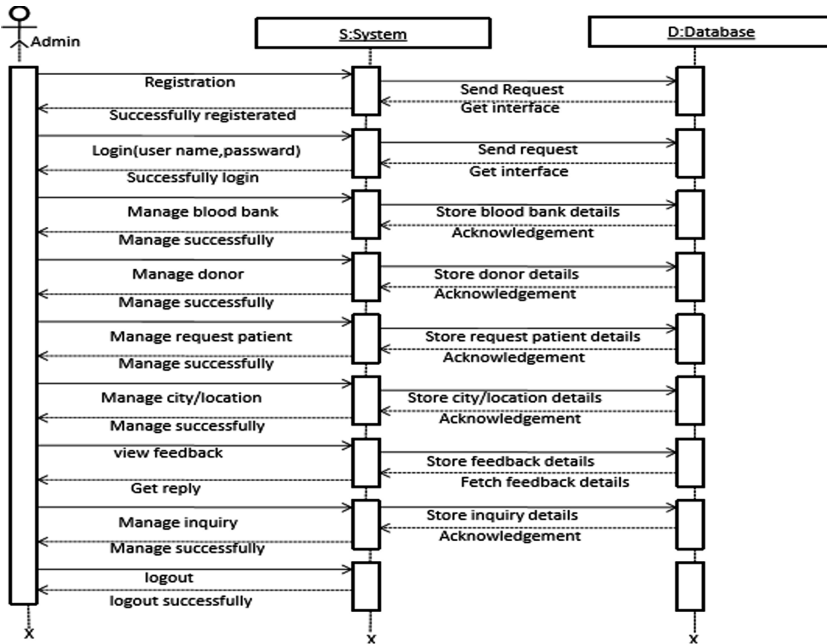
Fig. 6. Sequence diagram for DONOR
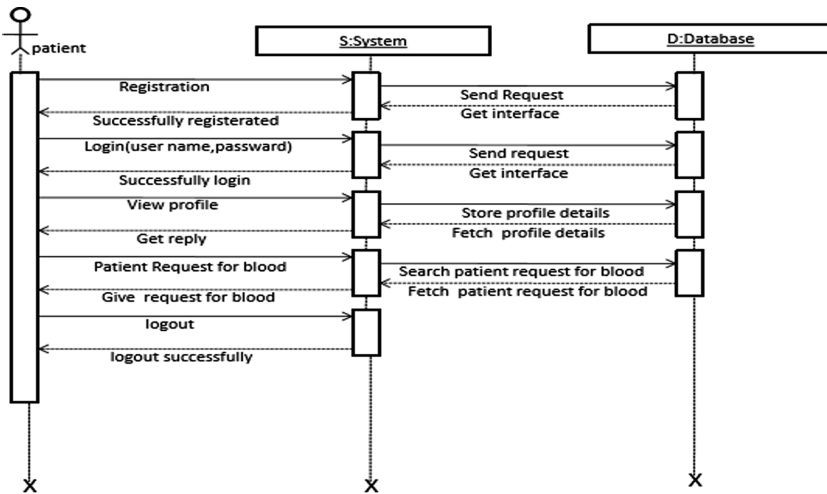
Fig. 7. Sequence diagram for Patient/Requester

Table 1. User registration information

Attribute name	Data type	Constraints
Reg_id	varchar(15)	Primary key
U_id	varchar(15)	Foreign key
Name	varchar(50)	Not null
Contact_number	varchar(15)	Not null
Address	varchar(50)	Not null
State	varchar(15)	Not null
City	varchar(15)	Not null
Location	varchar(15)	Not null
Pin_code	varchar(15)	Not null
Email	varchar(30)	Not null

Table 2. Donor information.

Attribute name	Data type	Constraints
Donation_id	varchar(15)	Primary key
Reg_id	varchar(15)	Foreign key
Name	varchar(50)	Not null
Bloodbank_id	varchar(15)	Foreign key
Blood group	varchar(15)	Not null
Donation-date	varchar(50)	Not null
State	varchar(15)	Not null
City	varchar(15)	Not null
Pin_code	varchar(15)	Not null
Email	varchar(30)	Not null
Phone_number	varchar(15)	Not null

Table 3. User request.

Attribute name	Data type	Constraints
Bloodbank_id	varchar(15)	Primary key
Reg_id	varchar(15)	Foreign key
Location_id	varchar(15)	Foreign key
City_id	varchar(50)	Foreign Key
Conatact_number	varchar(15)	Not null
Blood_group	varchar(15)	Not null
Status	varchar(15)	Not null
Request_date	varchar(15)	Not null

5 Module Stages

Screen shots of various modules are given in Figs. 8, 9, 10, 11, 12, 13, 14 and 15.

Admin Login

Enter Admin Name

password

LOGIN

Fig. 8. Admin login

DONOR

E-mail

Password

LOGIN NOW

Not registered? Create an account

Fig. 9. Donor login page.

REQUESTER

E-mail

Password
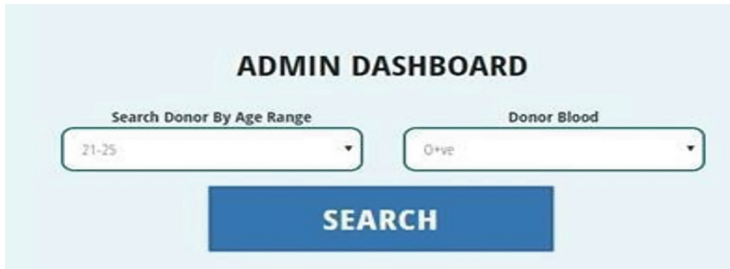
LOGIN NOW

Fig. 10. Requester login page.

Donor Details

Donor Id	Donor Name	Email	D.O.B	Mobile Number	Gender	Blood Group	Place
1	Shareef	shareef@gmail.com	23/05/1990	7406680955	Male	O+ve	Mysore
2	Shivanand	Shivanand@gmail.com	12/2/1991	9743216864	Male	AB+ve	Mysore
3	Roopa	roopa@gmail.com	12/7/1990	7899480966	Female	O+ve	Mandya
4	GG	gg@gmail.com	14/8/1990	7895640123	Female	O+ve	Mangalore
5	Asif	asif@gmail.com	4/8/1990	5698741230	Male	B-ve	Mysore
6	bb	bb@gmail.com	22/6/1991	4564564567	Female	AB-ve	Bangalore
7	cc	cvc@gmail.com	8/8/1988	7894562031	Male	B+ve	HYD

Fig. 11. Donor details.

Fig. 12. Admin search.

Fig. 13. Request blood.

Fig. 14. Profile update.

Fig. 15. Previous donation details.

6   Conclusion

The blood donation management system (BDMS) is a 24 × 7 system provides services to the hospitals and other users. The system is easy to maintain all the information about the blood donor. Proposed work provides services to persons who pursue donors who are willing to donate blood. Since blood donors lack confidence regarding the confidentiality of their private information, the blood donation management system provides a secure and confidential atmosphere. This system facilitates to get the information in case of emergency. The system helps to make the scheme paper less. The proposed web application provides interaction between blood donors, blood requester and blood bank. This system allows the blood donors to update their information and allow the requester to show their requests to the donor. The benefits that can reached are (1) reducing human error (2) Improving monitoring process (3) Increasing information sharing (4) reducing operational costs. In future, we will develop the application with emerging technologies by adding more features.

References

1. Abderrahman, B.H., Saleh, M.Y.N.: Investigating knowledge and attitudes of blood donors and barriers concerning blood donation in Jordan. Procedia-Soc. Behav. Sci. **116**, 2146–2154 (2014). https://doi.org/10.1016/j.sbspro.2014.01.535
2. Balaraman, P., Kosalram, K.: E-hospital management and hospital information systems – changing trends. Int. J. Inf. Eng. Electron. Bus. **5**, 50–58 (2013)
3. Li, B.N., Dong, M.C., Chao, S.: On decision making support in blood bank information systems. Expert Syst. Appl. **34**, 1522–1532 (2008). https://doi.org/10.1016/j.eswa.2007.01.016
4. Selvamani, K., Rai, A.K.: A novel technique for online blood bank management. Procedia Comput. Sci. **48**, 568–573 (2015). https://doi.org/10.1016/j.procs.2015.04.137

5. Kros, J.F., Pang, R.Y.: A decision support system for quantitative measurement of operational efficiency in a blood collection facility. Comput. Methods Programs Biomed. **74**, 7–89 (2004). https://doi.org/10.1016/s0169/2607(03)00072-5
6. Kulshrestha, V., Maheshwari, S.: Blood bank management system in India. Int. J. Eng. Res. Appl. (IJERA) **1**(2), 260–263. ISSN 22
7. Hazzari, D., Wijesekera, D., Hindawai, S.: Formalizing and verifying work flows used in blood banks. Procedia Technol. **16**, 1271–1280 (2014). https://doi.org/10.1016/j.protcy.2014.10.143
8. Sugijaro, D.P., Safie, N., Mukhtar, M., Sulaiman, R.: Telehealth model information flow: a case study on laboratory information system. Procedia Technol. **11**, 740–747 (2013). https://doi.org/10.1016/j.protcy.2013.12.253
9. Sulaiman, S., Hamid, A.A.K.A., Yusri, N.A.N.: Development of blood bank management system. Procedia-Soc. Behav. Sci. **195**, 2008–2013 (2015)
10. Bhowmik, A., Nabila, N.A., Imran, M.A., Rahaman, M.A.U., Karmaker, D.: An extended research on the blood donor community as a mobile application. Int. J. Wirel. Microw. Technol. **6**, 26–34 (2015). https://doi.org/10.5815/ijwmt.2015.06.04
11. Kalem, G.: Mobile technology in the healthcare industry for disease management and wellness. Procedia-Soc. Behav. Sci. **195**, 2014–2018 (2015). https://doi.org/10.1016/j.sbpro.2015.06.216
12. Chinnaswamy, A., Gopalakrishnan, G., Pandala, K.K., Venkat, S.N.: A study on automation of blood donor classification and notification techniques. Int. J. Appl. Eng. Res. **10**, 18503–18514 (2015)
13. Snaigdha, Anabhavane, V., Lokhande, P., Kasar, S., More, P.: Android blood bank. Int. J. Adv. Res. Comput. Commun. Eng. **4** (2015)
14. Jenipha, T.H., Backiyalakshmi, R.: Android blood donor life saving application in cloud computing. Am. J. Eng. Res. (AJER) **3**, 105–108 (2014)
15. Kulshrestha, V., Maheshwari, S.: Benefits of management information system in blood bank. Int. J. Eng. Sci. **1**, 05–07 (2012)
16. Meshram, A., Bramhe, P., Jaronde, S., Kamble, A., Ambataker, P.: Central blood bank data base with anti GPS mobile system. Int. J. Emerg. Technol. Adv. Eng. **3** (2013)
17. Rupsanth, J.A., Marikkannu, P.: Automated blood bank management system using direct call routing technique. Int. J. Novel Res. Comput. Sci. Softw. Eng. **3**, 107–111 (2016)
18. Ali, K.M.A., Jahan, I., Islam, M.A., Parvez, M.S.: Blood donation management system. Am. J. Eng. Res. (AGER) **4**, 123–136 (2015)
19. Nzoka, M., Ananda, F.: Blood bank management information system a case study of Kenya national blood transfusion services. In: Proceedings of Sustainable Research and Innovation Conference, pp. 146–149. ISSN2079-6226

Silent Speech Recognition

Amaresh P. Kandagal[1]([⊠]) [ID], V. Udayashankara[2] [ID],
and M. A. Anusuya[2] [ID]

[1] Sri Siddhartha Institute of Technology, Tumkur, India
amareshpk@gmail.com
[2] Sri Jayachamarajendra College of Engineering, Mysuru, India
{v_udayashankara, anusuya_ma}@sjce.ac.in

Abstract. Speech is essential to exchange information. Speech recognition is one of the interfaces for man-machine interaction. However, the performance of these systems is restricted to noisy acoustic conditions. Silent speech i.e. visual dynamic features of speech have more potential information for Human-Computer Interaction. This paper presents lip localization and segmentation by Otsu algorithm. The height and width parameters of lip movements are captured as visual cues for silent speech recognition. We develop stochastic visual word models with an in-house database of 20 subjects. Performance evaluation these models are measured by word error rate. The accuracy of the system recorded for speaker dependent female subjects is 84.6%, and 65.8% as an overall result.

Keywords: Lip reading · Otsu · Speech recognition · HMM

1 Introduction

The principle of Automatic Lip Reading (ALR) is built on the abilities of expert human lip-readers, who can accomplish almost perfect speech comprehension. The initial work on ALR appeared in 1984 when Petajan investigated the use of lip features to enhance speech recognition by creating a bimodal (audio-visual) speech recognition system [1, 2]. Visual dynamics of speech i.e. visemes are applied in various applications, such as audio-video synchronization, in biometrics for speaker identification, visual only and audiovisual speech recognition and much more [3]. Extraction of efficient visual speech features is the most complicated component to build silent speech recognition system. It is quite challenging task to extract visual features from given image or sequence of frames (video) which is captured in various conditions such as dynamic illumination, different pose, and skin tones. On lip localization and viseme extraction several algorithms are proposed and investigated by various researchers is briefly discussed in further sections of this paper.

Figure 1 represents video segmentation (extraction of keyframes from video), segmentation of a region of interest (ROI of only lip portion), feature extraction, and search algorithm or pattern recognition modules, which are building blocks of silent speech recognition system. In this particular work, the main focus is to develop visual speech recognition using geometric lip parameter statistical models for the Kannada language.

Sections 2 and 3 explains in brief lip segmentation approaches and database creation. Section 4 describes methodology and system performance is presented in Sect. 5. Experiments and its results are discussed in Sect. 6. Section 7 discusses conclusions and future enhancements of the proposed work.
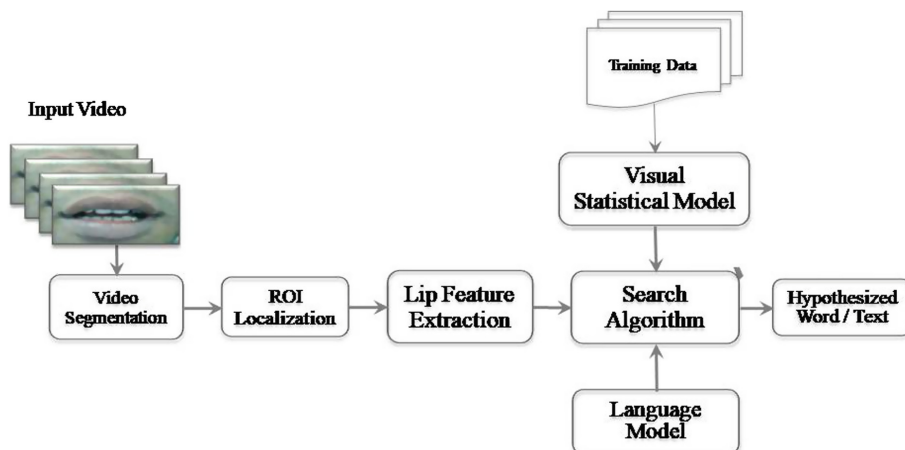
Fig. 1. Architecture of visual speech recognition system

2 Lip Detection Methods

The difficult part in "machine lip-reading" is lip segmentation. It is performed by image processing algorithms and is a challenging task to distinguish between the lip and non-lip pixels. The fundamental challenges exists because of various speaker profiles and background conditions [4] such as lip and skin color (race), lip shape (width and height), facial hair (makeup), amount of lip movement during speech, distance from the camera (scale), orientation relative to the camera (pose), Illumination conditions (intensity, shadows, glare), and Background activity.

Lip segmentation techniques are broadly categorized into color based techniques and model-based techniques.

2.1 Model-Based Lip Detection Methods

Many model-based lip detection approaches exist namely Active contour is (AC), Active Appearance (AA) and Active Shape (AS) or Point Distribution Models (PDMs) are model-based lip detection approaches. In these methods, a set of image frames trained to build lip models. With these models, lip position and shape is estimated in new images.

Active Contour Models often referred to as "snakes". These are broadly applied in lip segmentation process because of their ability to smoothing and elasticity [5]. As per Hassanat [6], active contour models are reported a good result but these models have some of the drawbacks.

Point Distribution Models (PDMs) and AAMs [7] are used for labeling and classifying the objects. Principal component analysis (PCA) is a widely used technique to analyze the shape and appearance of the objects. A collection of a labeled point object in images are trained to construct statistical contour model by PCA.

2.2   Color-Based Approaches

The color transform features are used in this In general RGB, HSV and YCbCr methods applied for color based lip segmentation. Among all these YCbCr Approach is commonly used.

In the RGB Red, green and blue pixels are used to form the RGB space. This is transformed into binary image considering only the lip portion. In HSV, Hue component can be applied to discriminate between skin and lip, as hue value for lip pixels is lesser compared to face pixels [8]. In YCbCr color structure Y corresponds to the luminous, Cb and Cr are chrominance components. YCbCr color space is extensively applied in digital videos. Lips are redder comparatively skin color and have dominant Cr and less Cb components, which is a great advantage for lip segmentation method.

3   Database Creation

Figure 2 represents the in-house database creation apparatus, in this work, a digital camera with the 5 MP camera is considered to capturing input videos of resolution 320×240 with 30 frames per second. Visual recordings were captured from 20 subjects, 10 male, and 10 female. The database is created for 10 Kannada words. Video recordings conducted in two different sessions. On every session, each word is captured 2 occurrences from respective speakers. So there will be 40 samples of visual recordings from each speaker.
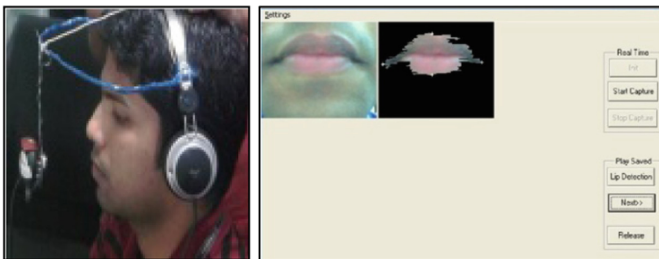
Fig. 2. Database generation and test setup

4   Methodology

4.1   Lip Geometric Feature Selection

Figure 3 describes the lip feature extraction stages. The video object file is a set of continues frames, it is not necessary to process all image frames. We extract the keyframes and these are processed further for feature extraction computation. The keyframes is estimated by considering absolute difference between present and preceding frame with reference to threshold value.

$$\text{Key frame} = \text{absolute (frame}[i] - \text{frame } [i-1]) > \text{threshold}$$

The threshold value is set empirically; in this experiment threshold value is set to 0.5.

Lip Map Generation and Color Analysis. In the process of productive feature extraction, selection of proper color space plays a significant role. To explore the status of each and every color model, histograms of basic color components are constructed to discretize the image colors. Histograms are constructed for the complete image and for the regions of interest (ROI) is depicted in Fig. 4.
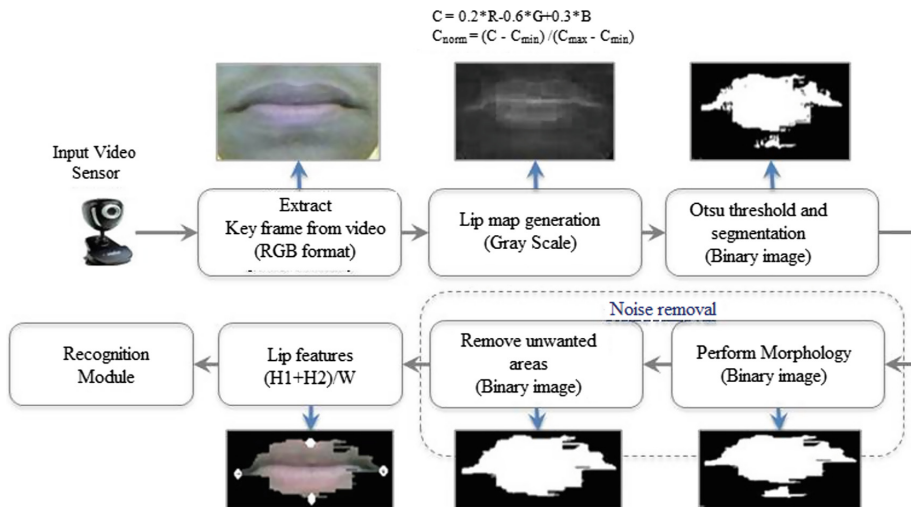
Fig. 3. Lip geometric feature extraction module
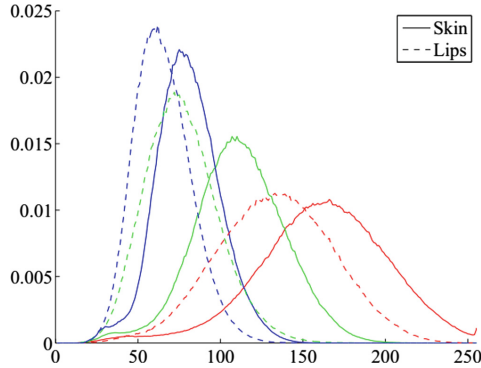
Fig. 4. R, G and B components for lip and skin (Color figure online)

The following observations are pointed out from the above figure.

(i)   The color elements (R, G, and B) produce peaks in the lip region. The color distribution at lip area is narrow with moderate uniform color.

(ii) The hue component has less similarity than histogram at the region of interest.

(iii) The distribution probability of RGB and Cb, Cr differs from subject to subject, whereas hue component relatively remains unchanged in dynamic lighting conditions, for various speakers. As a result, for lip segmentation hue component is an apt measure. Figure 4 presents histograms lip and skin color components. This distinctly shows that deficiency of RGB color components for lip object separation in all three color space. Colored images (composed of RGB) are computationally expensive compared to grayscaled image frames. Therefore these key frame images are converted to a grayscale format to reduce the computation complexity. Greyscale transformation for lip map generation lm is computed as

$$lm = a * R - b * G + c * B \tag{1}$$

Where $a = 0.2, b = 0.6$, and $c = 0.3$

$$lm_{norm} = \frac{(lm - lm_{min})}{lm_{max} - lm_{min}} \tag{2}$$

The Otsu's algorithm implementation is a gray threshold function in MATLAB® is given by [9].

$$\sigma_{Intra}^2 = x_1 \sigma_1^2 + x_2 \sigma_2^2 \tag{3}$$

$$\sigma_{Inter}^2 = x_1 x_2 (m_2 - m_1)^2 \tag{4}$$

$$\sigma_T^2 = \sigma_{Inter}^2 + \sigma_{Intra}^2 \tag{5}$$

$$\eta = \frac{\sigma_{Inter}^2}{\sigma_T^2} \quad 0 \leq \eta \leq 1 \tag{6}$$

Where σ_i^2 and σ_T^2 are a variance of ith class, x_i probability of a pixel and m_i mean of class i.

Morphological Processing. The morphological process is a used to extract image elements boundaries [10]. Dilation and Erosion are essential algorithms in morphological image processing. These methods are applied for structuring element to expand and/or shrink i.e. objects of binary image to make "thicken" and/or "thin".

Feature Selection. We measure the angle of lip movement for silent speech recognition. The ratio of lip height by width estimates lip motion angle $\tan\theta$ as illustrated in Fig. 5. As lip shape is symmetrical in nature by measuring the midpoint of width, we estimate the upper and lower lip height.

$$\tan\theta = (H1 + H2)/W \tag{7}$$

Where w width of the lip, H1 and H2 are upper and lower lip height respectively
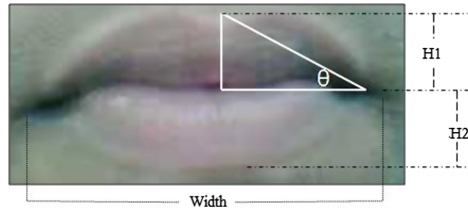
Fig. 5. Example of the extracted lip geometric features

4.2 Lip Reading by Maximum Likelihood Approach (HMM)

Many researchers used HTK mainly for acoustic speech recognition. But in our work, we use HTK API's HInit and HRest to build statistical visual models and HVite for visual observation recognition by Viterbi algorithm [11]. Figure 6 presents a schematic diagram of automatic lip pattern recognition system by HTK toolkit.

Visual HMM Model. There many conventional methods for lip reading in the visual domain. There are, various challenges come across to build visemes in ALR systems due to less number of visemes with respect to phonemes. Therefore, it causes the system performance issues. Whereas the projected "visual words" expect the characteristic cues of the complete word instead of segments lip motion sequences. A visual word HMM model is used to estimate the performance of unknown observations.
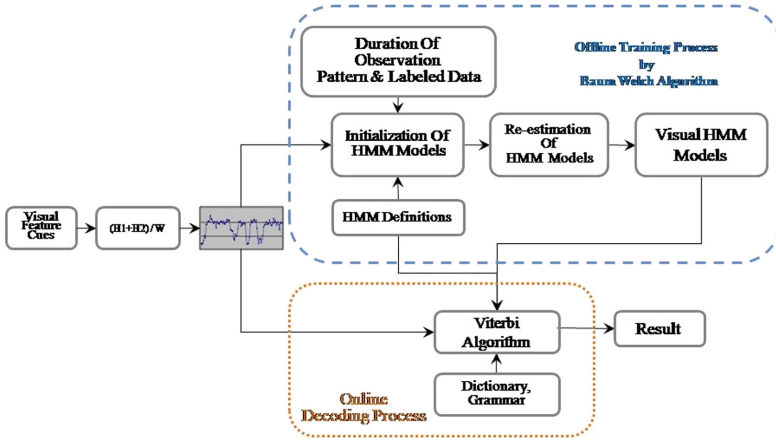
Fig. 6. Schematic diagram automatic lip pattern recognition

Visual words models are appropriate for lesser vocabulary size. For each word in vocabulary list, prototype HMM is created with dummy values. Prototype HMM consists of nine active states including two non-emitting states. For all HMMs, the same topology is followed. Each state has three Gaussian mixers expressed as mean and variance vectors. In training processes, lip feature vector (.lf files) and labeled corpus data (.lab files) are required to create initialization models by HInit API from HTK tool [11]. For each observation features HMM parameters such as transition probability, mean and variances are re-estimated iteratively by HRest API. For each HMM to train this is repeated several times. These embedded re-estimations point out the convergence factor by the change of measure in training data. The training process is repeated until convergence test is matched.

Recognition or Decoder Module. This module converts unknown observation sequences of visual parameters to text. Considering that each word corresponds to observation cues i, e lip parameter vector at time 't' L defined as:

$$L = l_1 l_2 l_3 l_4 l_5 \ldots l_T \tag{8}$$

Where l_t is the lip observation sequence vector at time t. Decoder module estimates the maximum likelihood of lip features and is given by

$$\hat{X} = \text{argmax}_X P(X|L) \tag{9}$$

Using conditional probability $P(X|L)$ is calculated using:

$$P(X|L) = \frac{P(L|X)P(X)}{P(L)} \tag{10}$$

Given a sequence of lip features L, the maximum probability is computed. The term P(X) is approximated by the language model it is a prior probability of the word. P(L|X) is computed using the visual HMM model keeping P(L) constant. The observation sequences are computed using Viterbi algorithm [11].

5 Performance Estimation

Accuracy of the system is measured by two parameters that are Word Error Rate (WER) and Percentage of Accuracy (POA) and these are computed as:

$$WER = \frac{S+I+D}{N} * 100 = 100 - (POA) \tag{11}$$

$$POA = \frac{N-D-S-I}{N} * 100 = \frac{H-I}{N} * 100 \tag{12}$$

Where N–words considered for the validation set, and D, S, I and H are a number of deletions, number of substitutions, number of insertions and number of correct labels respectively.

6 Results and Discussion

Various simulations were conducted to estimate the performance of the system using our own video database.

6.1 Speaker Dependent

This experiment is carried out on each speaker, all the training and testing samples correspond to the same speaker. As discussed in Sect. 3, from each speaker forty samples are captured, in which training is done by 60% and remaining samples are used for testing. The accuracy of the system is computed by using Eq. 12 and tabulated in Table 1. The objective is to validate, the mode of speaking style of each speaker. Each subject is capable to generate a visual signal that can simply read. To perform this cross-validation protocol is used.

6.2 Speaker Independent

In this experimentation, the system performance is evaluated on each speaker sample against the rest of the speaker's sample. One speaker sample from the training set is validated over other samples. The accuracy of the system is computed by cross validation method. These tests avoid differences in individual appearance and speaking style. The results are tabulated in Table 1.

Observations from experimental results the overall recognition rate is 65.8% and 45.6% for speaker-dependent and speaker-independent experiments respectively. Performance analysis plot, Fig. 7 shows that the accuracy of speaker-dependent is much

higher with respect to speaker-independent; this affirms by many researchers such as (Jun and Hua) [12]. From speaker to speaker for a same spoken word it differs in the way of talking and mouth appearance. This shows that visual speech recognition problem is a speaker-dependent problem. It's observed that female speaker's performance is good compared to male speakers. The performance of the system is degraded in male speakers due to incorrect lip feature extraction (because of mustaches and beards).

Table 1. Experimental results of ALR system

Female speaker	Speaker dependent	Speaker independent	Male speaker	Speaker dependent	Speaker independent
F1	92%	67%	M1	54%	27%
F2	81%	56%	M2	47%	30%
F3	85%	69%	M3	56%	33%
F4	81%	66%	M4	52%	31%
F5	88%	58%	M5	40%	29%
F6	83%	67%	M6	53%	27%
F7	80%	58%	M7	55%	32%
F8	90%	65%	M8	38%	18%
F9	82%	68%	M9	30%	27%
F10	84%	59%	M10	45%	25%
All				65.8%	45.6%
Females				84.6%	63.3%
Males				47%	27.9%
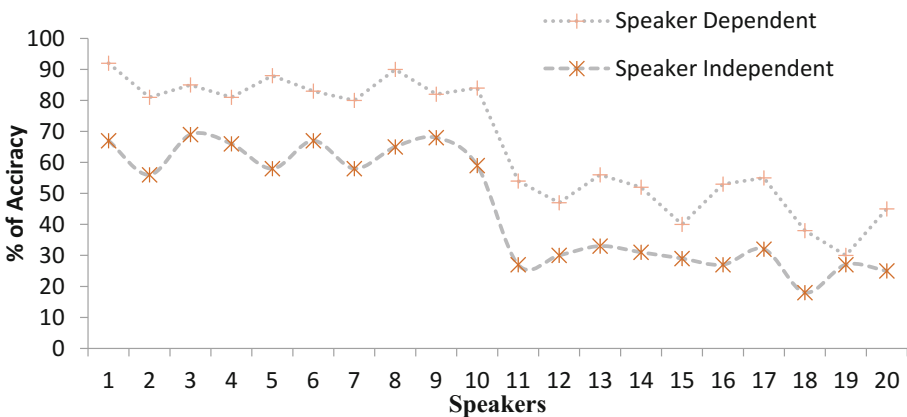
Fig. 7. Performance analysis of ALR system

7   Conclusion and Future Scope

This paper reviewed major technological perspective on lip localization methods. We presented the lip feature localization and segmentation by Otsu threshold method. Geometric parameters such height and width are used for lip tracking. Statistical visual word models and pattern recognition is performed by HMM toolkit. The performance evaluation of the system is computed for word visual models. The overall accuracy of the system is 65.8% is recorded for the in-house database. Female speaker performance is higher as compared to male speakers. The future work could be carried out to have a robust feature extraction module should be capable to handle illuminations (lighting conditions), facial hairs and tongue movements for speaker independent and large vocabulary VSR system, and also to build speech recognition system for noisy acoustic environmental conditions using multi modality.

References

1. Petajan, E.: Automatic lip reading to enhance speech recognition. In: IEEE Proceedings of Global Telecommunications Conference, Atlanta, GA, pp. 265–272 (1984)
2. Kandagal, A.P., Udayashankara, V.: Automatic bimodal audiovisual speech recognition a review. In: IEEE International Conference on Contemporary Computing and Informatics, Mysore, India, pp. 940–945 (2014). https://doi.org/10.1109/ic3i.2014.7019673
3. Tareque, M.H., Al Hasan, A.S.: Human lips-contour recognition and tracing. Int. J. Adv. Res. Artif. Intell. **3**, 47–51 (2014)
4. Luettin, J., Thacker, N.A., Beet, S.W.: Speech reading using shape and intensity information. In: 4th International Conference on Speech and Language Processing, vol. 1, pp. 58–61 (1996)
5. Kass, M., Witkin, A., Terzopoulos, D.: Snakes active contour models. Int. J. Comput. Vis. **1**(4), 321–331 (1988)
6. Hassanat, A.B.A., Jassim, S.: Color-based lip localization method. In: Proceedings of SPIE - The International Society for Optical Engineering (2010). https://doi.org/10.1117/12.850629
7. Matthews, I., Cootes, T.F., Bangham, J.A., Cox, S., Harvey, R.: Extraction of visual features for lip-reading. Trans. Pattern Anal. Mach. Intell. **24**, 198–213 (2002)
8. Eveno, N., Caplier, A., Coulon, P.Y.: New color transformation for lips segmentation. In: 4th IEEE Workshop on Multimedia Signal Processing, pp. 3–8 (2001)
9. The MathWorks Inc.: MATLAB User Guide, vol. 4 (1998)
10. Gonzalez, R.C., Woods, R.E., Eddins, S.L.: Digital image processing using MATLAB, vol. 2. Gatesmark Publishing, Knoxville (2009)
11. Young, S., Evermann, G., Kershaw, D., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V., Woodland, P.: The HTK Book. Cambridge University Engineering Department, Cambridge (2009)
12. Jun, H., Hua, Z.: Research on visual speech feature extraction. In: Proceedings of the International Conference on Computer Engineering and Technology, vol. 2, pp. 499–502 (2009)

Cognitive Computing and Its Applications

Swarm Intelligent Approaches for Solving Shortest Path Problems with Multiple Objectives

Jinil Persis Devarajan[1(✉)] and T. Paul Robert[2]

[1] National Institute of Industrial Engineering, Mumbai, India
jinilpersis@gmail.com
[2] College of Engineering Guindy, Anna University, Chennai, India
prpaul@annauniv.edu

Abstract. Finding shortest path considering multiple objectives is a widely studied graph problem which yields multiple optimal paths called the pareto optimal path set by applying dominance principle. Literature reveals that solving such problems within polynomial time is difficult even for smaller instances using traditional algorithms. This study investigates the convergence of swarm intelligent algorithms and compares them with performance metrics such as the pareto coverage, divergence rate, convergence time and cost for different sample networks.

Keywords: Ant colony optimization · Multi-objective shortest path problem
Bee colony optimization · Firefly algorithm · Swarm intelligence
Pareto optimal front · Non-dominated set

1 Introduction

To find an optimal path connecting a source and a destination in a network with minimum cost is essential in transportation networks, flow networks and data networks. Researchers are showing interest to solve this problem and to obtain routing decisions using various methods including exact algorithms, approximation methods, heuristic methods and evolutionary methods [1]. This is investigated as an optimization problem with multiple objectives since there many situations where there is a need to consider multiple decision making criteria. The multiple objectives considered in this network optimization problem would yield multiple optimal solutions leading to a non-dominated path set or pareto optimal set. Every solution pair (p1, p2) satisfies the following dominance principle. A solution is non-dominated if there does not exist a solution across the feasible region having a better value for atleast one objective. Based on the network size and problem size, determining the pareto optimal path set becomes difficult. A shortest path problem with multiple objectives contains pareto optimal path set that leads to the destination from the source vertex minimizing (in minimization problem instance) every objective considered in the problem. It finds application in vehicle routing, military path planning, robot path planning, material handling, assembly line planning, project management and network routing in communication networks [1–3].

2   Literature Review

Many routing strategies are used in the algorithms given in the literature that could solve this shortest path problem with multiple objectives. These strategies are broadly categorized into classes, viz., the classical approaches and non-classical approaches.

The various classical methods employed are exact algorithms, approximation methods and heuristics [1]. The exact algorithm searches the entire feasible region and uncovers non-dominated paths available in the solution space with exponential time complexity [4]. Many strategies are used to uncover the Pareto optimal set from the feasible region that include dynamic programming, depth first search, breadth first search, brute force search and branch and bound method. The crucial shortcoming of these exact algorithms is that the entire path set may be intractably larger to generate in polynomial time. Therefore this problem is often called as an NP-hard and NP-complete problem in the literature [1, 5, 6]. Several k-shortest path algorithms are presented that generate k shortest paths between a node pair however has exponential increase in computational complexity with the increase in k [7].

This kind of problems existing in real world challenges and the complete Pareto optimal frontier is difficult to derive. Approximation methods such as method of meta-criterion function, lexicographic method, $\in$-constraint method and weighted metric method can tackle these problems [1]. For graphs having multiplicative or additive weight vector on edges, weighted metric method or metacriterion method can be applied to transform the problem with multi-objectives to a problem with single objective which can then be solved using standard algorithms like Djikstra's and Bellman's algorithms. The Djikstra's algorithms is applicable to solve shortest path problem with multiple objectives under two conditions - The individual objectives are non-negative and additive [1]. However for problems having non-additive multiple objective values for edges, such as probability of reaching destination, maximal/ minimal of visits so far and maximal/minimal of commissions involved in this route (penalties), it is infeasible to employ weighted metric method. Such problems can be solved using dynamic programming, goal programming, data envelopment analysis, label correcting/setting algorithms [8].

In order to accelerate the search process, several heuristic methods like A* search, branch pruning, bi-directional search, hierarchical search, lexicographic search, greedy search, coloring schemes, labelling methods, binomic algorithm, scatter search and local search are proposed. Multi-objective A* search uses more precise heuristic estimators that determines the accuracy and quality of the path set [9]. The de-facto shortest path algorithm viz., Djikstra's algorithm executes greedy search method to solve any given problem. However, for problems having both maximization and minimization objectives, conflicting objectives, the meta-criterion objective function lead to negative edges, which is not solvable by Djikstra's algorithm. Modified multi-objective version of Djikstra's algorithm called Martin's algorithm and A* search are developed which generate optimal solution. Extended versions of this algorithm, viz., a Multi-Objective A* (MOA*) search and a Multi-Objective A* search to yield $\in$-Pareto set, $(MOA^*_\in)$ are proposed to determine solutions in multi-objective shortest path problem [9]. The experimental analysis on the scalability of these algorithms revealed that these

algorithms could handle larger graphs (upto 3000 nodes) and could efficiently solve multi objective problems (with upto 10 objectives). However these algorithms require large memory leading to high storage complexity [10]. Also, these heuristics do not guarantee global optimum especially in a multi-objective and multi-stage graph problems since they tend to fall into local optima [4, 11]. The complexity of computing the overall objective function value by the method of meta-criterion is $\Theta(VE)$ [1].

Following are the observations made from the review of classical algorithms:

- The shortest path problems with multiple objectives are identified as a NP-hard problem [12], the complete enumeration of non-dominated path set from the feasible area is not feasible in polynomial time.
- Approximation techniques to transform the problem with multiple objectives to a problem having single objective using method of meta-criterion and weighted metric method are developed when each individual in the objective vector satisfies the approximation conditions.
- Heuristic techniques can be used under suitable conditions however at the cost of less accurate solution set.

Several non-traditional algorithms that yield quasi-optimal solutions to complex optimization problems are reported in the literature. These algorithms lead to sub pareto optimal front when applied to this kind of problems. Multi-objective versions of GA, Simulated Annealing (SA), PSO are found in literature. In an attempt to compare the performance of ACO and GA in solving a network flow problem, ACO algorithm is identified to be more efficient when compared to GA in terms of minimizing the cost and computational time. ACO variants like ACO with local search [13], ACO with Tabu search [14] have been introduced to minimize them when compared to the traditional ACO algorithm. Multi-Objective Ant Colony System (MOACS), Pareto Ant Colony Optimization (P-ACO) algorithm, Multi-Objective Ant Colony Optimization (MO-ACO) algorithm are few multi-objective extensions of the traditional ACO algorithm. A new modified version of the ant colony optimization combined with insert, swap and 2-opt algorithm solved the 'multiple traveling salesman problem'. The proposed algorithm uses a global updating policy, state transition rule and an efficient candidate list that improve the efficiency of this algorithm from the standard ACO algorithm. This algorithm could efficiently solve problem instances having number of cities ranging between 76 and 1002 with minimum convergence time and high solution quality [15]. The review of the variants in ACO algorithms for vehicle routing problem has been carried out and these algorithms are found to be efficient and able to substantially improve the solution costs [16]. A new algorithm combining beam search with max-min ant system solved weighted vehicle routing problem in cargo transportation [17]. The effectiveness of this algorithm when examined against general ACO and max-min ant system for larger benchmarking instances. The results showed lower operation costs and increased cost savings even when there is a spreaded customer distribution and huge weight.

Multi-objective numerical optimization is solved using BCO and MO-BCO proposed has yielded better convergence and diversity than NSGA II. The problem of path planning is solved using artificial bee colony optimization (ABC) algorithm combined with a local trajectory planning scheme in mobile robotics. Experiments revealed that average total path deviation and average uncovered target distance are much better in

the proposed approach when it is examined against differential evolution and PSO algorithms. ABC for solving economic/environmental load shipment problem having multiple objectives is presented to minimize fuel charge, system loss and pollution. The experimental results revealed that this approach generate widely distributed Pareto optimal solutions when compared to NSGA, NPGA, SPEA, multi-objective differential evolution and multi-objective PSO. A robust version of BCO algorithm to solve travelling salesman problem incorporated pattern reduction scheme and local search [18]. This algorithm improved solution quality and computational cost when compared to the classical BCO algorithm. A 110 node network is solved in transit network design problem using BCO algorithm which yielded high quality solutions [19]. The initial path set is obtained by applying greedy method and final path set is obtained using the proposed BCO algorithm which is found to be fast and efficient.

In this study, the performance of ACO, BCO and Firefly algorithms for determining pareto optimal path set for shortest path problem having multiple objectives is investigated when compared with the exact algorithm.

3   Methodology

The variable used are,

The variable used are,
G - a graph
V - vertices
E - edges connection (i,j) between a node pair
W - a weight vector represented as, $W_{ij} = (d_{ij}, c_{ij}, de_{ij}, l_{ij}, r_{ij})$
d_{ij} - distance of (i,j)
c_{ij} - cost of (i,j)
de_{ij} - delay of (i,j)
l_{ij} - load of (i,j)
r_{ij} - reliability of (i,j)
T_f - Feasible Set
T_{ND} -Non-dominated set
NH_i - Neighbor list of node i
N - Number of nodes
current - The node lastly added to path p
τ_{ij} - Pheromone matrix containing intensity on each link (i,j)
P_{ij}- Probability of selecting the link (i,j)
α - Relative importance given to pheromone value
β - Relative importance given to the objective function
ρ - Evaporation rate $\in (0,1)$
m - Population size of ant/bee/firefly swarm

ψ_{ij} - Combined objective function value of link (i,j)
$$\psi_{ij} = \overline{d_{ij}} + \overline{de_{ij}} + \overline{l_{ij}} + \overline{c_{ij}} - \overline{r_{ij}}$$
F_i - Overall cost of the path taken by ant i
F_{best} - Overall cost of the best path obtained so far
$current\_pos_i$ - Current Node position of ant i
$reached_i$ Binary variable denoting whether the ant reached destination node or not
$next\_pos_i$ - Next node position determined for ant i
P - Probability of accepting the best route
B_i - Brightness of the firefly i
B_i^0 - Initial Brightness of the firefly i
At_{ij} - Attractiveness of the firefly's current node position i to its neighborhood node position j
γ - Absorption coefficient of light
α' - Randomization parameter

Exact algorithm is a straightforward approach that generates the entire pareto optimal front by examining all candidate solutions existing in the considered problem. The exact algorithm given in this study as in Fig. 1 uses an enumerative approach that generates all the feasible paths present in a given network to connect source and destination nodes. Every pair of paths in the feasible set is applied with the dominance principle to obtain pareto optimal set of non-dominated paths, TND. This algorithm yielded enumerated solutions from the dominant region of the solution space. However this algorithm can be applied only for small-scale networks as the algorithm is intractable for larger networks.

```
For every node i ∈ N,
  obtain the list of neighbour, NH_i
For every node n in NH_s,
  T_f={S,n}//add an incomplete path
For n=3:N
  For every r ∈ T_f
   this_node=last(p)
   If this_node ≠ D
   For every j ∈ NH_this_node
      If j ∉ r
          concat r,j
    Delete r from T_f
  For every r∈ T_f ,
   Find (D_r,C_r,De_r,L_r,R_r)
  T_ND=∅
  For every r ∈  T_f
   For every r' ∈  T_f
       If r' is dominating r, r ∉ T_ND
  T_ND= T_ND ∪ r
```

Fig. 1. Exact algorithm applying dominance principle

```
Initialize τ_ij for all links
For each ant i in m
 Initialize current_pos_i=source
 While(current_pos_i!=destination)
```

$$P_{current\_pos_i,j} = \frac{(\tau_{current\_pos_i,j})^\alpha (\psi_{current\_pos_i,j})^\beta}{\sum_{\forall g \in NH_{current\_pos_i}} (\tau_{current\_pos_i,g})^\alpha (\psi_{current\_pos_i,g})^\beta}$$

For each j in NH_current_pos,

```
next_pos_i=maximum(P_current_pos_i,j) for all j
F_i= F_i+ψ_current_pos_i,next_pos_i
             if F_best > F_i , F_best = F_i
T^i_ND = T^i_ND ∪ {next_pos_i};
//Pheromone update
For all node i' in N
   For all node j' in N
      If  i' = next_pos_i and j' = current_pos_i
```

$$\Delta\tau = \frac{1}{F_{best}} \quad ; \quad \tau_{i'j'} = \tau_{i'j'}+\Delta\tau$$

```
      Else τ_i'j' = (1- ρ) τ_i'j'
  current_pos_i = next_pos_i
```

Fig. 2. ACO algorithm for shortest path routing

```
For each bee i in m
  Initialize current_pos_i=source
  While(current_pos_i!=destination)
  next_pos_i=choose  j in NH_current_pos_i with probability P
```
$F_i= F_i + \psi_{current\_pos_i.next\_pos_i}$
$\quad\quad T^i_{ND} = T^i_{ND} \cup \{next\_pos_i\};$
$\quad\quad if\ F_{best} > F_i\ ,\ F_{best} = F_i;\quad\quad p_{best}=T^i_{ND}\ ;\ best=i$
$\quad\quad current\_pos_i = next\_pos_i$
```
For each bee i∈ m
   Update T^i_ND with p_best with a probability
```
$\quad\quad\quad F_i = F_{best}$
$\quad\quad\quad current\_pos_i = current\_pos_{best}$

Fig. 3. BCO algorithm for shortest path routing

Now this exhaustive method can establish quality benchmark solutions that can be compared with the pareto optimal front yielded by other algorithms. This method yielded the entire pareto optimal path set but increasing the network size, the computational effort increases which ultimately leads to intractability. Ant colony, bee colony and firefly based algorithms for this problem are given in Figs. 2, 3, 4. These algorithms can be applied to obtain approximate pareto optimal set of paths in any given network.

A sample network is shown in Fig. 5 in which Mobile Adhoc Network (MANET) is considered and the algorithms discussed in this study are applied to get pareto optimal routes from the source to destination nodes minimizing the delay, load, distance, cost and maximizing the reliability.

```
For each firelfy i in m
  Initialize initial brightness B^i
  Initialize current_pos_i=source
  While(current_pos_i!=destination)
  For each j in NH_current_pos_i ∉ T^i_ND ,
```
$\quad\quad Estimate\ At_{current\_pos_i,j}\ and\ B^i_{current\_pos_i}$
$\quad next\_pos_i = \max_j At_{current\_pos_i,j}$
$\quad F_i= F_i + \psi_{current\_pos_i.next\_pos_i}$
$\quad T^i_{ND} = T^i_{ND} \cup \{next\_pos_i\}$
$\quad current\_pos_i = next\_pos_i$
$\quad F_{best} = \min_i F_i$
$\quad B^i_{current\_pos_i} *= exp^{-\gamma F^2_{best}}$

Fig. 4. Firefly algorithm for shortest path routing
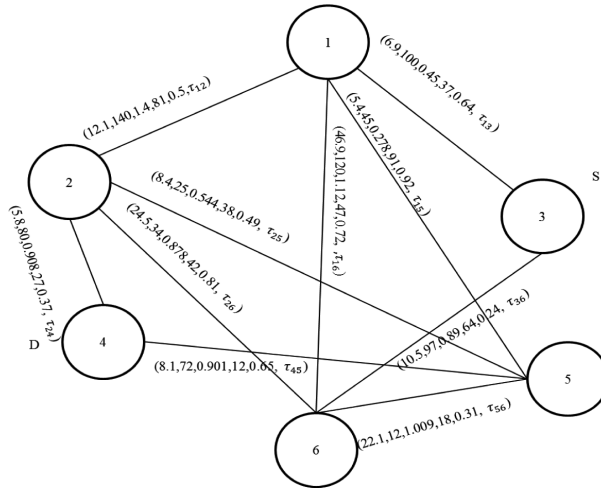
Fig. 5. Sample 6-node network

The average values of the distance, cost, time delay, load and reliability of the pareto optimal path set obtained using exact algorithm is compared with the sub optimal paths obtained using ACO, BCO and Firefly algorithms as shown in Fig. 6.
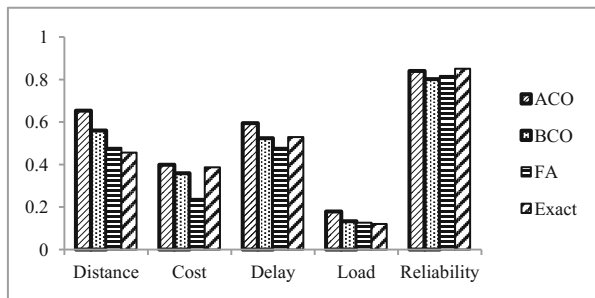
Fig. 6. Performance comparison

4   Performance Comparison

The application of the algorithms proposed in this study are used to solve the shortest path routing in MANET using MATLAB and the convergence characteristics of the algorithms are obtained. Exhaustive simulation experiments are carried out to derive the appropriate range of parameters at which the convergence is obtained at less computational cost and computational time. So as to compute the sub optimal path set quality, divergence and pareto coverage (the ratio of the total solutions yielded by a

swarm intelligent algorithm to that of the exact pareto optimal frontier) are used. The approximated/sub pareto optimal path set obtained by these algorithms for a 5-node, 8-edge network are presented in Table 1.

Table 1. Route metrics of pareto optimal path set for the 5-node network

Solution set		Distance	Cost	Delay	Load	Reliability
Exact algorithm	3-2-4	118	0.6713	1.2103	12.5805	0.8509
	3-5-4	104	6.1947	2.5095	39.3982	0.5196
	3-1-2-4	144	11.2775	2.3064	7.9407	0.3387
ACO algorithm	3-5-4	104	6.1947	2.5095	39.3982	0.5196
(population size = 5)	3-2-4	118	0.6713	1.2103	12.5805	0.851
	3-1-2-4	144	11.2775	2.3064	7.9407	0.3387
BCO algorithm	3-5-4	104	6.1947	2.5095	39.3982	0.5196
(population size = 5)	3-2-4	118	0.6713	1.2103	12.5805	0.851
	3-1-2-4	144	11.2775	2.3064	7.9407	0.3387
Firefly algorithm	3-5-4	104	6.1947	2.5095	39.3982	0.5196
(population size = 5)	3-1-2-4	144	11.2775	2.3064	7.9407	0.3387

Several simulation trials are done with varying values of relative importance given to pheromone value, α and relative importance given to the objective function, β. The convergence of ACO algorithm is presented in Table 2.

Table 2. Convergence characteristics of ACO algorithm

Parameter setting	Pareto coverage	Convergence time	Convergence cost
$\alpha = 0.1; \beta = 0.9$	0.3333	0.0192	−0.0096
$\alpha = 0.3; \beta = 0.7$	0.2222	0.1049	0.0139
$\alpha = 0.5; \beta = 0.5$	0.22222	0.0236	0.0034
$\alpha = 0.7; \beta = 0.3$	0.33333	0.0196	−0.0027
$\alpha = 0.9; \beta = 0.1$	0.44	0.0196	−0.0118

The simulation results revealed that at $\alpha = 0.9; \beta = 0.1$, the pareto coverage is maximum and the convergence cost is the minimum. However the convergence time is minimum at $\alpha = 0.1; \beta = 0.9$ but there is only a slight difference. So the optimum parameter setting is determined as $\alpha = 0.9; \beta = 0.1$ and evaporation rate $\rho = 0.2$. The study revealed that the convergence characteristics of the ACO algorithm is found to be good at population size = 5 and with the considered α, β and ρ values. The parameter of the BCO algorithm is the probability with which the bee decides to follow the bee which has discovered the best partial path so far. Simulation trials are done with various probability of acceptance, P values. The convergence of BCO algorithm is shown in Table 3 that aid in deriving the optimal setting.

Table 3. Convergence characteristics of BCO algorithm

Parameter setting	Pareto coverage	Convergence time	Convergence cost
$P = 0.2$	0.22	0.0277	−0.0126
$P = 0.4$	0.44	0.0278	−0.00296
$P = 0.5$	0.33	0.0418	−0.00964
$P = 0.6$	0.11	0.0415	0.010603
$P = 0.7$	0.44	0.0419	−0.00296
$P = 0.8$	0.22	0.0416	−0.0126
$P = 0.9$	0.22	0.0417	−0.03581

Table 4. Convergence characteristics of Firefly algorithm

Parameter setting	Pareto coverage	Convergence time	Convergence cost
$\alpha' = 1; \gamma = 0.1$	0.3333	0.0754	−0.0096
$\alpha' = 1; \gamma = 0.3$	0.22222	0.08220	−0.01260
$\alpha' = 1; \gamma = 0.5$	0.33333	0.07780	−0.0096
$\alpha' = 1; \gamma = 0.7$	0.33333	0.08320	−0.0096
$\alpha' = 1; \gamma = 0.9$	0.33333	0.07850	−0.0096
$\alpha' = 2; \gamma = 0.9$	0.2222	0.0799	0.0034
$\alpha' = 2; \gamma = 0.7$	0.3333	0.0832	−0.0096
$\alpha' = 2; \gamma = 0.5$	0.3333	0.0865	−0.0096
$\alpha' = 2; \gamma = 0.3$	0.2222	0.0804	0.0034
$\alpha' = 2; \gamma = 0.1$	0.3333	0.0784	−0.0096

It is found from table that the pareto coverage is maximum at $P = 0.4$, and the convergence cost and the convergence time are the minimum. Therefore, the optimum parameter value is decided as $P = 0.4$. Better convergence of BCO algorithm is obtained at population size 5 with $P = 0.4$. So as to determine the parameters of Firefly Algorithm viz., the randomization parameter, α' and absorption coefficient of light, γ, several simulation runs are executed with varying values of α', γ. The convergence characteristics of FA are presented in Table 4. The parameter values $\alpha' = 1; \gamma = 0.5$, are found to be the optimum parameter setting. The impact of the population size on the algorithm's convergence when examined shows better characteristics at population size 5 with the obtained α' and γ values. This process of parameter setting is repeated for all the networks considered in the study. A network with 10 nodes and 21 edges is considered and optimization on routing objectives yielded the Pareto optimal set of paths. Several simulation runs are carried out to derive the optimal parameter setting for this network configuration. The outcomes are shown in Table 5.

Table 5. Parameter setting for the 10-node network

Algorithm	Parameters	Convergence characteristics		
		Pareto coverage	Convergence time	Convergence cost
ACO	$\alpha = 0.6; \beta = 0.4; \rho = 0.8;$ $m = 50$	0.163636	0.029	0.473988
BCO	$P = 0.3; m = 50$	0.1	0.031	0.488656
FA	$\alpha' = 2; \gamma = 0.1; m = 25$	0.136364	0.1508	0.2516

It is found from the study that with 25 fireflies considered during the implementation of FA yielded maximum Pareto covering while ACO and BCO algorithms required a swarm size of 50 to obtain good solutions.

A 15 node and 28 edge network is considered and it's pareto optimal path set is found out. Several simulation trials are conducted to derive the optimal parameter setting for this network. The outcomes are shown in Table 6.

Table 6. Parameter setting of 15-node network

Algorithm	Parameter values	Convergence characteristics		
		Pareto coverage	Convergence time	Convergence cost
ACO algorithm	$\alpha = 0.6; \beta = 0.4; \rho = 0.6;$ $m = 150$	0.478022	0.1885	1.312839
	$\alpha = 0.6; \beta = 0.4; \rho = 0.6;$ $m = 250$	0.631868	0.255	1.370367
	$\alpha = 0.6; \beta = 0.4; \rho = 0.6;$ $m = 375$	0.906593	0.4844	1.534548
BCO algorithm	$P = 0.3; m = 100$	0.120879	0.031	1.028493
Firefly algorithm	$\alpha' = 1; \gamma = 0.2; m = 100$	0.291209	0.3475	0.300549
	$\alpha' = 1; \gamma = 0.9; m = 100$	0.175824	0.2683	0.279171
	$\alpha' = 2; \gamma = 0.5; m = 100$	0.197802	0.3256	0.221356
	$\alpha' = 1; \gamma = 0.2; m = 125$	0.302198	0.3899	0.307046
	$\alpha' = 1; \gamma = 0.9; m = 150$	0.340659	0.3368	0.317717

The study revealed that, upto 90% of the Pareto set can be covered in ACO algorithm at the expense of computational time and cost. In Firefly Algorithm, 35% coverage was achieved with less cost and time when compared to ACO algorithm with a swarm size of 150, beyond which the algorithm showed almost no improvement. BCO algorithm converged at a swarm size of 100 but showed 18% Pareto covering.

A 20 node and 130 edge network is considered. It is observed from the study the exact algorithm when applied to this problem could not converge since the number of feasible paths in the solution space is beyond numerical limits. Thereby extracting the non-dominated set becomes intractable. The swarm intelligent algorithms are applied. The simulation study shows that ACO algorithm performed well when a swarm of 275 ants are employed at which maximum divergence is obtained in the path set with minimum convergence cost and time. The convergence characteristics are presented in Fig. 7. Further from the study the optimal parameter setting is derived as $\alpha = 0.8;$ $\beta = 0.2; \rho = 0.8$. More diversified Pareto front is obtained by increasing the swarm size but beyond 300 ants the convergence cost gradually increases, which indicates that the sub optimal front obtained gets away from the pareto optimal front and hence the solution quality is declined. Also it can be observed from the graph that, the convergence time shows exponential increase with the increase in swarm size. Considering all these insights, the optimal swarm size is determined as 300 for this instance considered.
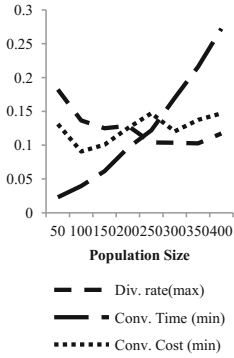
Fig. 7. Convergence characteristics of ACO algorithm
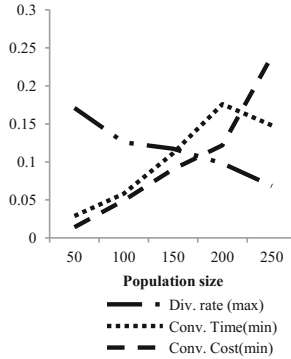
Fig. 8. Convergence characteristics of BCO algorithm
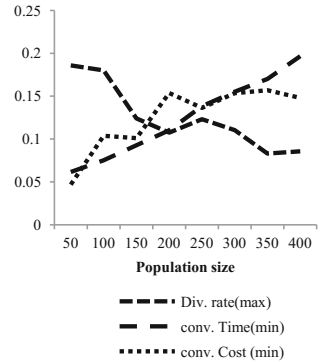
Fig. 9. Convergence characteristics of Firefly algorithm

When BCO algorithm is applied for this problem, it is observed that better results are obtained with the probability value 0.3. From Fig. 8 the convergence of BCO algorithm can be observed. It is seen from Fig. 8 that the total solutions increases with the increase in swarm size and attains maximum at 200. With further increase in swarm size both the solution quality and the strength of the path set decrease. However the convergence time decreases at the cost of the strength of the path set which is not desirable. So the optimal population size is found to be 200.

More diversified solutions are obtained from Firefly algorithm when randomization parameter, $\alpha' = 1$ and absorption coefficient, $\gamma = 0.9$ are used. The results obtained from Firefly Algorithm with this parameter setting for different population size are shown in Fig. 9. The simulation study revealed that the convergence time and cost are minimum and the divergence is maximum when the population size is 250.

The exact algorithm explored the entire pareto optimal path set from the feasible area of the search space. But the simulation experiments showed that, the algorithm could handle upto 15 nodes beyond which it becomes intractable. The nearness of the approximated Pareto set yielded by swarm intelligent algorithms with the actual pareto optimal front generated by the exact algorithm is analyzed. The algorithms' convergence time and cost is presented in Table 7.

Table 7. Cost and time analysis

Convergence time					Mean convergence cost				
Nodes	Exact algorithm	ACO	BCO	FA	Nodes	Exact algorithm	ACO	BCO	FA
5	0.1482	0.0212	0.0379	0.1156	5	0.0040	−0.0126	−0.0027	−0.0198
10	0.4006	0.0435	0.0376	0.1505	10	0.331306	0.0297	0.1977	0.2552
15	0.9465	0.1252	0.0459	0.2761	15	0.1905	0.3490	0.1984	0.1046

It is observed from the pair wise comparison test that the variation in the convergence time among other swarm intelligent algorithms is insignificant. As far as the classical approach is concerned, the convergence time augments as the number of nodes and edges are gradually increased.

The convergence cost of the algorithms varies significantly from exact solutions yielded by the exact algorithm. The correlation analysis revealed that the mean cost of the path set yielded by ACO is weakly correlated with that of the exact solution set. The mean cost of the path set given by BCO and FA are strongly correlated with that of the exact set.

The main limitation seen in the exact algorithm remains intractability. This algorithm runs indefinitely when confronted with larger networks. Simulation study revealed that the algorithm could solve networks with upto 15 nodes.

The scalability of the algorithms namely ACO, BCO and Firefly algorithms is established by implementing them on larger networks with 25, 50, 75 and 100 nodes. These networks are randomly generated by Waxman model in a 2-dimensional simulation grid. The probability of any two nodes being connected is considered as 0.5.

Table 8. Mean computational cost of paths

Algorithm	Number of nodes			
	25	50	75	100
ACO	0.00803	0.002	0.0012	0.0007
BCO	0.0136	0.005	0.0026	0.001
FA	0.0067	0.003	0.00108	0.00070

Table 9. Convergence time analysis

Algorithm	Number of nodes			
	25	50	75	100
ACO	0.0616	0.53834	0.40861	2.1623
BCO	0.0572	0.1257	0.09536	0.200
FA	0.455	3.9764	4.220	2.1475

The mean computational cost yielded by these algorithms is shown in Table 8. The computational time taken by each algorithm is shown in Table 9.

It is observed that ACO and BCO algorithms converge faster than Firefly algorithm. The swarm intelligent algorithms generate multiple paths in single run and the time complexity depends on the swarm size and the network size. The study also shows that ACO and BCO algorithms converge faster than the firefly algorithm. The divergence of the solution space for the swarm intelligent algorithms is shown in Fig. 10.
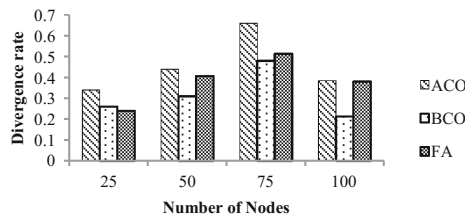
Fig. 10. Comparison of divergence rate

It is observed from Fig. 10 that the divergence rate increases as the nodes are increased upto 75, after which it declines. Also ACO algorithm yielded more diversified path set when compared to BCO and Firefly algorithm. Firefly algorithm performs well than BCO algorithm especially when the size of the network is increased more than 25 nodes.

5   Conclusion

The proposed routing algorithms presented in this study for solving shortest path problem with multiple objectives are simulated under a static simulation environment. The algorithms are proposed with a limitation to single source node and a single destination node graph problem. The convergence characteristics of the algorithms are studied in terms of the pareto coverage, divergence, computational cost and computational time. The nearness of the obtained path set with that the exact set is obtained for smaller instances of the problem by which the accuracy of the solutions yielded by the swarm intelligent algorithms are obtained. For larger problem instances, the divergence of the path set is used to acquire the strength of approximated pareto optimal path set yielded by swarm intelligent algorithms.

References

1. Tarapata, Z.: Selected multicriteria shortest path problems: an analysis of complexity, models and adaptation of standard algorithms. Int. J. Appl. Math. Comput. Sci. **17**, 269–287 (2007). https://doi.org/10.2478/v10006-007-0023-2
2. Haddad, O.B., Mirmomeni, M., Mehrizi, M.Z., Marino, M.A.: Finding the shortest path with honey-bee mating optimization algorithm in project management problems with constrained/unconstrained resources. Comput. Optim. Appl. **47**, 97–128 (2008). https://doi.org/10.1007/s10589-008-9210-9
3. Tarapata, Z.: Military route planning in battlefield simulation: effectiveness problems and potential solutions. J. Telecommun. Inf. Technol. **4**, 47–56 (2003)
4. Hui, L., Yonghui, C.: Study of heuristic search and exhaustive search in search algorithms of the structural learning. In: IEEE Second International Conference on Multimedia and Information Technology, pp. 169–171. IEEE (2010). https://doi.org/10.1109/mmit.2010.163
5. Häckel, S., Fischer, M., Teich, T., Zechel, D.: A multi-objective ant colony approach for pareto-optimization using dynamic programming categories and subject descriptors. In: Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation, GECCO 2008, pp. 33–40 (2008)
6. Liu, L., Mu, H., Luo, H., Li, X.: A simulated annealing for multi-criteria network path problems. Comput. Oper. Res. **39**, 3119–3135 (2012). https://doi.org/10.1016/j.cor.2012.03.013
7. Mohanta, K.: Comprehensive study on computational methods for k-shortest paths problem. Int. J. Comput. Appl. **40**, 22–26 (2012)
8. Reinhardt, L.B., Pisinger, D.: Multi-objective and multi-constrained non-additive shortest path problems. Comput. Oper. Res. **38**, 605–616 (2011). https://doi.org/10.1016/j.cor.2010.08.003

9. Perny, P., Spanjaard, O.: Near admissible algorithms for multiobjective search. In: European Conference on Artificial Intelligence (ECAI), pp. 490–494 (2008)
10. Aljazzar, H., Leue, S.: K*: a heuristic search algorithm for finding the k shortest paths. Artif. Intell. **175**, 2129–2154 (2011). https://doi.org/10.1016/j.artint.2011.07.003
11. Beg, S., Khan, A., Nauman, U., Mohsin, S.: Performance evaluation of bionomic algorithm (BA) in comparison with genetic algorithm (GA) for shortest path finding problem. Int. J. Comput. Sci. Issues **8**, 238–242 (2011)
12. Demeyer, S., Goedgebeur, J., Audenaert, P., Pickavet, M., Demeester, P.: Speeding up Martins' algorithm for multiple objective shortest path problems. 4OR **11**, 323–348 (2013). https://doi.org/10.1007/s10288-013-0232-5
13. Monteiro, M.S.R., Fontes, D.B.M.M., Fontes, F.A.C.C.: An ant colony optimization algorithm to solve the minimum cost network flow problem with concave cost functions. In: Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation, GECCO 2011, pp. 139–145. ACM Press, New York. https://doi.org/10.1145/2001576.2001596
14. Yoshikawa, M., Otani, K.: Ant colony optimization routing algorithm with Tabu search. In: Proceedings International Multi Conference Engineers Computer Science, vol. III, pp. 17–20 (2010)
15. Yousefikhoshbakht, M., Didehvar, F., Rahmati, F.: Modification of the ant colony optimization for solving the multiple traveling salesman problem. Rom. J. Inf. Sci. Technol. **16**, 65–80 (2013)
16. Naqvi, N.Z., Matheru, H.K., Chadha, K.: Review of ant colony optimization algorithms onvehicle routing problems and introduction toestimation-based ACO. In: Proceedings of International Conference on Environment Science and Engineering, vol. 8, pp. 161–166 (2011). http://cpfd.cnki.com.cn/Article/CPFDTOTAL-CDYA201104003036.htm
17. Tang, J., Guan, J., Yu, Y., Chen, J.: Beam search combined with MAX-MIN ant systems and benchmarking data tests for weighted vehicle routing problem. IEEE Trans. Autom. Sci. Eng. 1–13 (2013). http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6705646
18. Girsang, A.S., Tsai, C.-W., Yang, C.-S.: A fast bee colony optimization for traveling salesman problem. In: 2012 Third International Conference on Innovations in Bio-Inspired Computing and Applications. IEEE. pp. 7–12 (2012). https://doi.org/10.1109/ibica.2012.44
19. Nikolić, M., Teodorović, D.: Transit network design by bee colony optimization. Expert Syst. Appl. **40**, 5945–5955 (2013). https://doi.org/10.1016/j.eswa.2013.05.002

Improved Directionally Driven
Self-regulating Particle Swarm Optimizer

Saumya Jariwala(✉)

Dhirubhai Ambani Institute of Information and Communication Technology,
Gandhinagar, India
saumya259@gmail.com

Abstract. Particle Swarm Optimization (PSO) and many of its variants tend to suffer from premature convergence on strongly multimodal test problems. It is known that maintenance of diversity in the swarm is very important for preventing the swarm from converging prematurely for multimodal function problems. In this paper, an improved version of the Directionally Driven Self-Regulating Particle Swarm Optimization (DD-SRPSO) algorithm is proposed with a mechanism to maintain diversity without incurring any significant computational cost referred to as Improved DD-SRPSO. An *attractive* and *repulsive* swarm update strategy inspired from ARPSO is incorporated in DD-SRPSO. The proposed method is tested using the shifted, rotated and complex benchmark functions from CEC2013. These results are compared with eight PSO variants including DD-SRPSO. The results indicate that the proposed method improves results on some of the benchmark functions and gives comparable results for others.

Keywords: Particle swarm optimization
Directional update strategy · Rotational invariant strategy
Diversity guided search

1 Introduction

Many nature inspired algorithms have become a tool of choice for optimization problems due to their high convergence speed and robustness. Many engineering problems are complex in nature with multidimensional and multimodal attributes where finding global optimum is challenging. With the increasing complexity of such problems and quick decision making requirements, these methods have gained significant research attention in the last decade [1]. Among these algorithms, Particle Swarm Optimization(PSO), developed by Eberhart and Kennedy [2], has created quite a stir in the research community due to its simplicity and computational efficiency. In this method, the behaviour of the bird flock is mathematically modeled to solve optimization problems. The position of the members of the swarm (referred to as particles) is randomly initialized in the search space for exploration and exploitation in order to converge

at the optimum solution. The particles move towards the global best solution by sharing the information about the search space and changing their velocity accordingly. The performance of PSO is quite good on many optimization problems [3], however, it is quite possible that some members of the swarm might lead the whole swarm to a suboptimal solution and the region of global optimum is left unexplored. Initially, the swarm may be able to escape these local optima using the momentum from the previous timestep but as the algorithm proceeds, the particles tend to lose momenta and get stuck in the sub-optimal solution from which the basic PSO and many of its variants are unable to escape [4].

This problem is more severe in high dimensional multimodal problems. To overcome this limitation of PSO, many approaches to maintain the diversity of the swarm have been proposed like - FDR-PSO [5], Regrouping PSO [4], pPSO [6], ARPSO [7], SEPSO [8], OPSO [9], MPSO [10] and some more. Some of these methods are described below.

An opposition based learning scheme has been proposed by Wang et al. [9]. This method is mainly aimed towards avoiding convergence to suboptimal solutions on multimodal functions. Van den Bergh [10] developed a multistart PSO (MPSO) which automatically restarts when stagnation is detected. After a pre-specified number of function evaluations, the best solution of all is reported as the final solution. This approach is equivalent to running the same algorithm multiple times independent of the previous runs. This might waste computations in the repetitious scouting of the same region which is not likely to contain a quality solution. As a solution to this limitation, a regrouping mechanism was proposed by Evers et al. [4]. In this method each time stagnation is detected, instead of re-initializing swarm randomly in the whole search space the swarm is regrouped in a smaller region to avoid repetitious scouting while still being able to escape the local optima. A two phase attractive-repulsive PSO has been proposed by Riget et al. [7]. In this method, the standard PSO keeps oscillating between two phases where the repulsive phase is introduced to regain the lost diversity due to possible premature convergence and keep the exploration going.

In this paper, a mechanism inspired from [7] has been incorporated in DD-SRPSO [11] to address the limitation of premature convergence resulting in a new Improved DDSRPSO model of optimization. Experimental results on CEC-2013 [12] benchmark problems show that the proposed method obtains a promising performance.

The organization of the paper is as follows. A brief review on SRPSO and DD-SRPSO algorithms is provided in Sect. 2. The proposed algorithm and its details along with the pseudo code is presented in Sect. 3. A brief illustration is presented in Sect. 4 so that the reader could get proper intuition. In Sect. 5 the results of this study and the performance analysis of the proposed method on benchmark functions is provided. The results of the proposed method are compared with other PSO variants in this section. Finally, the results of this study are summarized in Sect. 6.

2   Brief Review of SRPSO and DD-SRPSO

2.1   SRPSO

Self-Regulating Particle Swarm Optimization (SRPSO) algorithm, introduced in [13], uses two learning strategies **(i)** the self-regulating inertia weight only for the best particle and **(ii)** self-perception in the global search direction for the rest of the particles.

The self-regulating inertia weight of the SRSPO algorithm is defined as:

$$\omega_i^{t+1} = \begin{cases} \omega_i^t + \Delta\omega, & for \ best \ particle \\ \omega_i^t - \Delta\omega, & otherwise \end{cases} \tag{1}$$

where ω_i^t is the inertia weight for i^{th} particle of the swarm at timestep t. Value of $\Delta\omega$ is set according to $\frac{\omega_I - \omega_F}{N_{Iter}}$ (N_{Iter} is the max number of iterations, ω_I is the initial and ω_F is the final value of inertia weight). For allowing the best particle to accelerate in its previous search direction, the self and social-perception for it are set to zero. The velocity update equation for the best particle is given by:

$$V_{id}^{t+1} = \omega_i V_{id}^t \tag{2}$$

where V_{id}^t denotes the velocity of the ith particle and d^{th} dimension at timestep t.

The self-perception in global search direction is introduced for the rest of the population to keep them from blindly following the global best solution. It can be described by equation given below:

$$p_{jd}^{so} = \begin{cases} 1, & \text{if a} > 0.5 \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

where p_{jd}^{so} is the self perception in the global search direction for the jth particle and d^{th} dimension. Here $a \sim U([0,1])$. The partial social cognition keeps the population from getting biased towards the global best. The value $a > 0.5$ is set empirically. While a very small value of a can make the population follow global best blindly, a too large value can also harm the convergence by making the motion too random. Using the above cognition approach the velocity update equation can be given by:

$$V_{jd}^{t+1} = \omega_j V_{jd}^t + (c_1 r_1 (P_{jd}^t - X_{jd}^t) + c_2 r_2 p_{jd}^{so} (P_{gd}^t - X_{jd}^t)) \tag{4}$$

where P_{jd}^t is the personal best for particle j in the dth dimension and P_{gd}^t is the global best position. c_1 and c_2 represent the acceleration coefficients. r_1 and r_2 are random numbers which are uniformly distributed within the range [0, 1]. The equation for position update is defined as

$$X_{id}^{t+1} = X_{id}^t + V_{id}^{t+1}, i = 1, ..., n; d = 1, ..., D \tag{5}$$

Overall, these strategies provide faster and closer convergence to the global best solution [13]. Further details about the algorithm can be found in [13].

2.2   DD-SRPSO

It has been observed in SRPSO that the poor performers of the swarms are unable to learn on their own and they need help from the global best as well as other elite learners [11]. Also, the algorithm performs poorly for the rotated functions. To address these limitations two new velocity update strategies have been proposed in DD-SRPSO [11] along with two existing strategies of SRPSO [13] described above. So the particles in DD-SRPSO follow one of these 4 update strategies:

1. Self Regulating Inertia Weight for best particle
2. Self-perception in the Global Search Direction Strategy
3. Directional Update Strategy
4. Rotational Invariant Strategy

Particles following directional update strategy follow the Eq. 6 for velocity update [11]. This strategy addresses the limitation of poor learning particles not being able to improve on their own. It improves the exploitation ability of the poorly performing particles by guiding them closer to the optimum value defined by elite particles. The size of the group of elite performers has been decided as the 5% of the total swarm size. This size cannot be very large because if too many particles participate in guiding the poor particles, it will make the exploitation unstable.

$$V_{jd}^{t+1} = \omega_j V_{jd}^t + (c_1 r_1 (lP_{jd}^t - X_{jd}^t) + c_2 r_2 (P_{gd}^t - X_{jd}^t)) \tag{6}$$

where lP_{jd}^t is the median of personal best of the 3 particles which are selected from the group of elite performers of the swarm and k is a poorly performing particle.

The rotational invariant strategy using the global best location for center of gravity has been introduced using the Eqs. 7, 8, 9 and 10 [11]. The rotational invariance strategy for the ith particle is defined by a center of gravity denoted by G_i. For a given particle i, G_i is defined by three points: X_i, a point slightly beyond P_i represented as p_i and a point slightly beyond p_g represented as p_g. The calculation of G_i is dependent on the self-perception of particles for global search directions p_i^{so} as given in Eq. 9.

$$p_i^t = X_i^t + c_1 r_1 \otimes \left(P_i^t - X_i^t \right) \tag{7}$$

$$p_g^t = X_i^t + c_2 r_2 p_i^{so} \left(P_g^t - X_i^t \right) \tag{8}$$

$$G_i^t = \begin{cases} \frac{p_i^t + p_g^t + X_i^t}{3}, & \text{if } p_i^{so} = 1 \\ \frac{p_i^t + X_i^t}{2}, & \text{if } p_i^{so} = 0 \end{cases} \tag{9}$$

$$V_i^{t+1} = \omega V_i^t + H_i \left(G_i^t, \|G_i^t - X_i^t\| \right) - X_i^t \tag{10}$$

where $H_i \left(G_i^t, \|G_i^t - X_i^t\| \right)$ denotes a random point in the hypersphere centered at G_i^t with radius $\|G_i^t - X_i^t\|$.

To sum up, the best particle uses self-regulated inertia weight strategy and continues moving in the previous direction without using the social cognition. The poorly performing particles use the directional update strategy. Now the remaining particles are divided into two groups by assigning each of them a random number. A threshold value (β) is decided and according to the number assigned to a particle is greater than or less than the threshold, they are assigned the rotational invariant strategy or the self-perception in the global search direction strategy respectively. The value of threshold β is set to 0.6 as mentioned in [11]. For more details about the algorithm refer to [11].

3 Improved DD-SRPSO

In DD-SRPSO two new strategies: directional update strategy and rotational invariant strategy have been introduced which improves its performance significantly than SR-PSO but it has been noted that on strongly multimodal test problems, like many PSO variant, the DD-SRPSO too faces the problem of premature convergence. This is because of the loss of diversity of the swarm in search space which in turn leads to stagnation at a suboptimal solution.

Among these, we have selected the strategy proposed in ARPSO to introduce diversity in DD-SRPSO yielding better results for some of the multimodal functions. The proposed algorithm works in two phase, namely **(i)** Attraction Phase **(ii)** Repulsion Phase

The attraction phase is simply the normal DD-SRPSO algorithm. As the algorithm proceeds, the particles will attract each other and the swarm starts to converge. In the repulsion phase the particles which follow the self-perception in global search direction strategy (strategy (ii) mentioned in Sect. 2), instead of being attracted towards better-known positions(global best and the personal best) start being repelled by the same as given below:

$$V_{jd}^{t+1} = \omega_j V_{jd}^t - (c_1 r_1 (P_{jd}^t - X_{jd}^t) + c_2 r_2 p_{jd}^{so} (P_{gd}^t - X_{jd}^t)) \tag{11}$$

The switching between two phases is performed based on the diversity of the swarm, measured as below [7].

$$diversity(t) = \frac{1}{|S||L|} \sum_{i=1}^{|S|} \|X_i^t - \bar{X}^t\| \tag{12}$$

where $|S|$ denotes the swarmsize, $|L|$ denotes the length of longest the diagonal in the search space, X_i^t is the position of the ith particle at timestep t and $\bar{X}^t$ stands for the average position of the swarm at timestep t. $\|X_i^t - \bar{X}^t\|$ is the euclidean distance between i^{th} particle and the average swarm position at timestep t.

During the attraction phase the particles come closer to each other and results in decreased diversity. If the diversity drops below a certain threshold, say d_low, the swarm switches to the repulsion phase, in which particles travel according

to Eq. 11 and the diversity increases. When the swarm attains the diversity d_high, the algorithm goes again in the attraction phase. As a result of this, the algorithm keeps alternating between two phases. An important difference from [7] is that during the repulsion phase not all the particles of the swarm are repelled. The best particle, the poorly performing particles and the particles following rotational invariant strategy keep updating their positions as proposed in DDSRPSO [11]. This takes care of the exploitation of the best region found so far. Also after $\gamma\%$ of the total function evaluations, the algorithm stops oscillating between two phases and focuses on the exploitation of the best region found so far. The exact flow of the algorithm is described in Algorithm 2.

To sum up the velocity update formula of Eq. 4 was modified by multiplying last two terms with *dir* giving the equation below:

$$V_{jd}^{t+1} = \omega_j V_{jd}^t + dir(c_1 r_1 (P_{jd}^t - X_{jd}^t) + c_2 r_2 p_{jd}^{so}(P_{gd}^t - X_{jd}^t)) \tag{13}$$

where value of *dir* is determined according to Algorithm 1

Algorithm 1. setDirection

1 **if** $dir > 0$ *&& diversity < d_low* **then**
2 $\quad$ $dir = -1$;
3 **else if** $dir < 0$ *&& diversity > d_high* **then**
4 $\quad$ $dir = 1$

---

4   Illustration

This section illustrates how the two phase switching introduced in the proposed method improves the chances of finding the global optimum. A well-known optimization function-Schwefel's Function is used for the illustration purposes. The details about the function can be found in [12]. It is 14th of the proposed 28 functions in [12]. Figures 1, 2, 3, 4 and 5 show the position of the swarm at different timesteps and Fig. 6 compares how the diversity of the swarm varies in DDSRPSO [11] and the proposed modification. The problem dimension is set to 2 to make the illustration more intuitive. Fig. 1 shows the initial(at t = 1) position of the swarm.

Figure 6 shows that by the 400 timesteps the swarm has converged to some optimum and the diversity is below the threshold *d_low*. This triggers the switch to the repulsion phase. When diversity touches the upper threshold *d_high* the swarm is again switched to the attraction phase. Figure 6 shows that the swarm keeps oscillating between the two phases to maintain its diversity until the $\gamma\%$ of the total function evaluations is used. The value of γ is set to 80%. These oscillations improve the chances of the swarm to escape the possible suboptimal solution and start exploring for better possible solutions. Figures 2, 3, 4 and 5 show how this strategy improves the performance of the normal DD-SRPO algorithm. In Fig. 2 the swarm has almost converged to a solution which is not the

---

Algorithm 2. Improved DD-SRPSO

---

1 initialization;
2 **for** *each particle i of population* **do**
3   Initialize the position X_i randomly in search range (X_{min}, X_{max});
4   Initialize velocity V_i randomly;

5 Using the fitness function calculate the fitness value;
6 Initialize personal best position by X_i for each particle;
7 Calculate the *diversity* of the swarm using Eq. 12;
8 **while** *success=0 and t≤ max_iter* **do**
9   *setDirection()* ;
10   **if** *iterations>γmax_iter* **then**
11   *dir = 1*;

12   *Let the position of the particle with the best fitness value be the global best;*
13   *Sort all particles w.r.t the fitness value;*
14   *Group top 5% particles as elite particles;*
15   *Group bottom 5% particles as poor particles;*
16   *Update inertia weight and velocity of the best particle using Eq. 1 and 2;*
17   **for** *poorly performing particles* **do**
18   *randomly select 3 particles from the group of elite particles;*
19   **for** *j=1:Dimension* **do**
20   *Calculate the median of the personal best of the 3 particles;*
21   *Update the velocity using Eq. 6;*

22   **for** *the remaining particles* **do**
23   *Generate the random number δ and set β = 0.6;*
24   **for** *j=1:Dimension* **do**
25   *Generate a using uniform random distribution;*
26   **if** *a > 0.5* **then**
27   *Global best dimension is selected;*
28   **else**
29   *The dimension is rejected;*
30   **if** *δ > β* **then**
31   *Calculate G_{id} using Eq. 9;*
32   *Use Eq. 10 to update velocity;*
33   **else**
34   *Use Eq. 13 to update velocity;*

35   *Position-update using Eq. 5 for all particles;*
36   *Recalculate diversity of the swarm using Eq. 12;*
37 *Report the global_best as the minimum*

---

global best. Now due to the attraction-repulsion switching introduced is normal DD-SRPSO, the swarm is able to escape this local optimum and has found a region for a better possible solution as shown in Fig. 4. Finally, the swarm has converged to the global best solution as shown in Fig. 5.
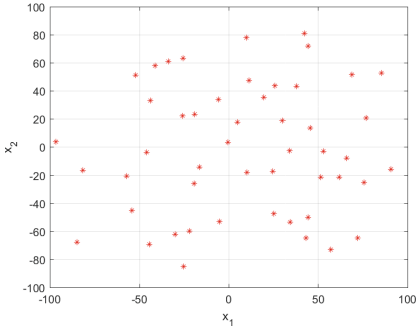
Fig. 1. t = 1
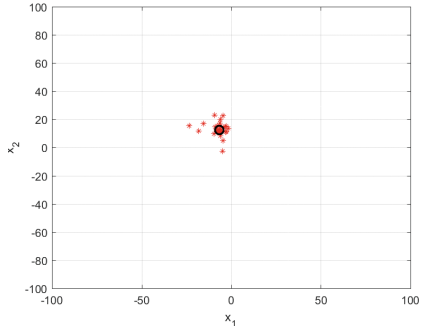
Fig. 2. t = 378
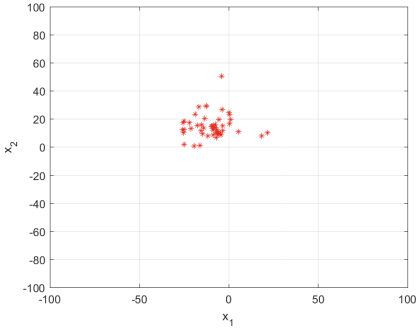
Fig. 3. t = 527
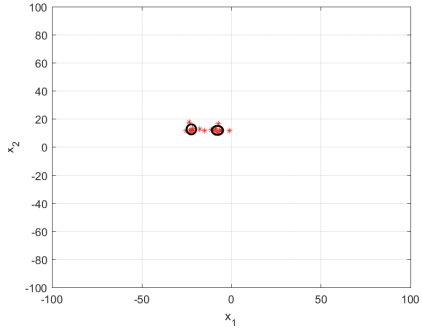
Fig. 4. t = 725
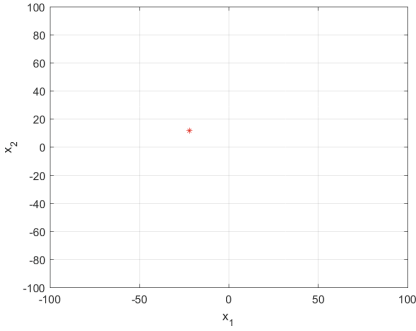
Fig. 5. t = 1000

Fig. 6. Diversity over 1000 timesteps

5 Experimental Setup and Results

This section provides the results for all the benchmark functions described in CEC2013 of the Proposed-Improved DDSRPSO and DDSRPSO [11] along with seven other PSO variants namely SLPSO [3], DMSPSO [14], LPSO [15], SRPSO [13], GPSO [16], CLPSO [17] and FIPS [18]. The experiments have been performed for 50 Dimensions and swarm size of 100. The experimental

Table 1. Mean performance on 50D CEC2013 benchmark functions

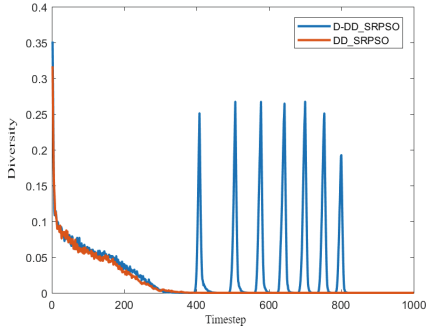| Function	Results for CEC2013 Benchmark Functions								
	SLPSO	LPSO	GPSO	DMSPSO	FIPS	CLPSO	SRPSO	DDSRPSO	Imp-DDSRPSO
F_1	0.00×10^{00}	2.34×10^{02}	5.29×10^{03}	8.27×10^{-06}	0.00×10^{00}	0.00×10^{00}	0.00×10^{00}	0.00×10^{00}	0.00×10^{00}
F_2	2.22×10^{06}	5.12×10^{07}	8.34×10^{07}	1.11×10^{07}	2.71×10^{07}	6.55×10^{07}	8.31×10^{06}	2.96×10^{06}	2.70×10^{06}
F_3	1.31×10^{07}	4.96×10^{08}	9.81×10^{10}	3.58×10^{08}	6.61×10^{08}	5.92×10^{09}	2.71×10^{08}	1.80×10^{08}	1.02×10^{08}
F_4	3.82×10^{04}	2.44×10^{04}	1.58×10^{04}	3.21×10^{04}	4.26×10^{04}	6.47×10^{04}	1.82×10^{04}	1.70×10^{04}	1.88×10^{04}
F_5	0.00×10^{00}	4.88×10^{02}	1.59×10^{03}	5.32×10^{-04}	0.00×10^{00}	0.00×10^{00}	3.15×10^{-08}	0.00×10^{00}	0.00×10^{00}
F_6	4.52×10^{01}	1.41×10^{02}	4.88×10^{02}	4.74×10^{01}	4.55×10^{01}	4.79×10^{01}	4.88×10^{01}	4.41×10^{01}	4.31×10^{01}
F_7	7.49×10^{00}	1.83×10^{02}	1.63×10^{02}	5.63×10^{01}	9.38×10^{01}	1.14×10^{02}	3.50×10^{01}	3.17×10^{01}	2.96×10^{01}
F_8	2.12×10^{01}	2.11×10^{01}	2.12×10^{01}	2.12×10^{01}	2.12×10^{01}	2.12×10^{01}	2.11×10^{01}	2.11×10^{01}	2.11×10^{01}
F_9	1.82×10^{01}	5.03×10^{01}	4.43×10^{01}	4.72×10^{01}	5.93×10^{01}	5.71×10^{01}	2.82×10^{01}	2.77×10^{01}	2.59×10^{01}
F_{10}	2.41×10^{-01}	3.43×10^{02}	1.45×10^{03}	6.99×10^{00}	2.97×10^{-01}	2.95×10^{01}	1.34×10^{00}	2.06×10^{-01}	1.56×10^{-01}
F_{11}	2.65×10^{01}	1.24×10^{02}	1.56×10^{02}	5.92×10^{00}	1.32×10^{02}	7.97×10^{-05}	6.82×10^{01}	6.94×10^{01}	6.92×10^{01}
F_{12}	3.39×10^{02}	1.89×10^{02}	3.73×10^{02}	1.26×10^{02}	4.09×10^{02}	3.31×10^{02}	1.09×10^{02}	8.76×10^{01}	8.20×10^{01}
F_{13}	3.43×10^{02}	3.25×10^{02}	5.87×10^{02}	2.36×10^{02}	4.18×10^{02}	4.07×10^{02}	2.70×10^{02}	1.99×10^{02}	1.93×10^{02}
F_{14}	1.08×10^{03}	6.53×10^{03}	2.59×10^{03}	1.87×10^{01}	1.07×10^{04}	1.44×10^{02}	3.00×10^{03}	3.44×10^{03}	2.83×10^{03}
F_{15}	1.23×10^{04}	8.11×10^{03}	7.76×10^{03}	9.22×10^{03}	1.41×10^{04}	1.04×10^{04}	1.13×10^{04}	6.74×10^{03}	6.58×10^{03}
F_{16}	3.33×10^{00}	2.52×10^{00}	2.09×10^{00}	2.01×10^{00}	3.50×10^{00}	2.82×10^{00}	2.71×10^{00}	2.76×10^{00}	2.76×10^{00}
F_{17}	3.70×10^{02}	1.18×10^{02}	3.46×10^{02}	6.47×10^{01}	3.56×10^{02}	6.28×10^{01}	1.12×10^{02}	1.12×10^{02}	1.04×10^{02}
F_{18}	3.97×10^{02}	2.57×10^{02}	3.45×10^{02}	2.08×10^{02}	4.44×10^{02}	4.52×10^{02}	3.92×10^{02}	3.13×10^{02}	3.07×10^{02}
F_{19}	9.18×10^{00}	6.44×10^{01}	4.36×10^{04}	4.15×10^{00}	2.94×10^{01}	3.92×10^{00}	5.54×10^{00}	5.99×10^{00}	5.39×10^{00}
F_{20}	2.24×10^{01}	2.12×10^{01}	2.22×10^{01}	2.35×10^{01}	2.17×10^{01}	2.38×10^{01}	2.15×10^{01}	2.02×10^{01}	2.02×10^{01}
F_{21}	7.60×10^{02}	1.19×10^{03}	9.33×10^{02}	7.88×10^{02}	3.62×10^{02}	4.91×10^{02}	8.18×10^{02}	8.13×10^{02}	7.73×10^{02}
F_{22}	1.14×10^{03}	6.79×10^{03}	4.25×10^{03}	5.54×10^{01}	1.05×10^{04}	6.07×10^{02}	3.59×10^{03}	3.72×10^{03}	3.47×10^{03}
F_{23}	1.10×10^{04}	9.56×10^{03}	1.06×10^{04}	7.95×10^{03}	1.46×10^{04}	1.15×10^{04}	9.69×10^{03}	7.71×10^{03}	7.60×10^{03}
F_{24}	2.46×10^{02}	3.62×10^{02}	3.37×10^{02}	2.71×10^{02}	3.37×10^{02}	3.55×10^{02}	2.88×10^{02}	2.78×10^{02}	2.72×10^{02}
F_{25}	2.94×10^{02}	3.80×10^{02}	4.81×10^{02}	3.29×10^{02}	3.73×10^{02}	3.94×10^{02}	3.45×10^{02}	3.39×10^{02}	3.39×10^{02}
F_{26}	3.35×10^{02}	4.45×10^{02}	4.17×10^{02}	3.88×10^{02}	2.50×10^{02}	2.07×10^{02}	3.57×10^{02}	2.64×10^{02}	3.53×10^{02}
F_{27}	7.29×10^{02}	1.81×10^{03}	1.68×10^{03}	1.19×10^{03}	1.74×10^{03}	1.87×10^{03}	1.13×10^{03}	1.08×10^{03}	1.08×10^{03}
F_{28}	4.00×10^{02}	3.17×10^{03}	4.33×10^{03}	1.68×10^{03}	4.00×10^{02}	4.00×10^{02}	7.02×10^{02}	4.00×10^{02}	4.00×10^{02}

setup has been kept same as [11]. The performances of SLPSO, GPSO, LPSO, FIPS, DMSPSO, CLPSO have been compared in [3]. For fair comparison, all the parameter settings are kept same as mentioned in [3] since the results of those algorithms are taken from the same. The results of DDSRPSO and Improved DD-SRPSO are compared with all the methods described in [3]. The parameter settings is as follows: inertia weight $\omega = [0.4, 0.9]$ for GPSO and LPSO, $\omega = [0.7, 0.9]$ for CLPSO and $\omega = 0.729$ for FIPS and DMSPSO. The settings for acceleration coefficients $c_1 = c_2 = 2$ for GPSO and LPSO, $c1 = c2 = 1.49445$ for SLPSO, DMSPSO and CLPSO. Further, the settings for SRPSO, DD-SRPSO and Improved-DDSRPSO are $\omega = 1.05 - 0.5$ and $c_1 = c_2 = 1.49445$. The diversity parameters d_high and d_low are set to 0.25 and $5.0 \cdot 10^{-6}$ respectively as mentioned in [7]. The algorithms were performed for 30 runs and for each run, function evaluation budget was set to 250,000 as given in [3].

The mean performance comparison for all the above-mentioned methods is provided in Table 1. The first five functions (F_1 to F_5) are unimodal benchmark

Table 2. Rank analysis of mean performance

Function	Ranking for CEC2013 Benchmark Functions								
	SLPSO	GPSO	LPSO	FIPS	DMSPSO	CLPSO	SRPSO	DDSRPSO	Imp. DDSRPSO
F_1	1	9	8	1	7	1	1	1	1
F_2	1	9	7	6	5	8	4	3	2
F_3	1	9	6	7	5	8	4	3	2
F_4	7	1	5	8	6	9	3	2	4
F_5	1	9	8	1	7	1	6	1	1
F_6	3	9	8	4	5	6	7	2	1
F_7	1	8	9	6	5	7	4	3	2
F_8	5	5	1	5	5	5	1	1	1
F_9	1	5	7	9	6	8	4	3	2
F_{10}	3	9	8	4	6	7	5	2	1
F_{11}	3	9	7	8	2	1	4	6	5
F_{12}	7	8	5	9	4	6	3	2	1
F_{13}	6	9	5	8	3	7	4	2	1
F_{14}	3	4	8	9	1	2	6	7	5
F_{15}	8	3	4	9	5	6	7	2	1
F_{16}	8	2	3	9	1	7	4	5	5
F_{17}	9	7	6	8	2	1	4	4	3
F_{18}	7	5	2	8	1	9	6	4	3
F_{19}	6	9	8	7	2	1	4	5	3
F_{20}	7	6	3	5	8	9	4	1	1
F_{21}	3	8	9	1	5	2	7	6	4
F_{22}	3	7	8	9	1	2	5	6	4
F_{23}	7	6	4	9	3	8	5	2	1
F_{24}	1	6	9	6	2	8	5	4	3
F_{25}	1	9	7	6	2	8	5	3	3
F_{26}	4	8	9	2	7	1	6	3	5
F_{27}	1	6	8	7	5	9	4	2	2
F_{28}	1	9	8	1	7	4	6	1	1
Average Rank	3.892857143	6.928571429	6.428571429	6.142857143	4.214285714	5.285714286	4.571428571	3.071428571	2.392857143
Overall Rank	3	9	8	7	4	6	5	2	1

functions. For these functions, the performance of Improved DD-SRPSO is similar to the DD-SRPSO algorithm because on unimodal functions the swarm does not get stuck in sub-optimal regions. Functions F_6 to F_{20} are multimodal functions. From Table 1 it is evident that Improved DD-SRPSO performs quite well as it has produced solutions closer to the optimum for 7 out of 15 functions. Except for F_{11} where CLPSO is able to generate a solution several order better than all the other algorithms, all the algorithms are mostly able to generate the results of the same order. The remaining functions (F_{21} to F_{28}) are composition functions. On these functions, SLPSO generates the best solution for 4 out of 8 functions whereas Improved DD-SRPSO generates best performance for 2 functions but it is able to generate the solutions of the same order in all the cases except F_{27} where SLPSO generates better solution by one order. As compared to DD-SRPSO, Improved DD-SRPSO improves the performance on almost all composition functions other than F_{26}. In Table 2, all the algorithms have been ranked according to their performance on benchmark functions. The average rank and the overall rank mentioned in Table 2 show the superiority of the proposed algorithm over other algorithms in consideration. Table 3 provides the average rank achieved by all the algorithms in different groups which provides better insight about the performance of the algorithm. It is observed that for the groups of unimodal and multimodal functions the Improved DD-SRPSO algorithm has the best average rank and in the composition function group it stands second. SLPSO algorithm performs a little better for composition function group and stands first. Nevertheless, the significance of any algorithm over

Table 3. Rank based analysis of mean performances in each group

	Avg. rank in unimodal	Avg. rank in multimodal	Avg. rank in composition
SLPSO	2.2	5.13	2.63
GPSO	7.4	6.53	7.38
LPSO	6.8	5.81	7.75
FIPS	4.6	7.2	5.12
DMSPSO	6	3.733	4
CLPSO	5.4	5.467	4.88
SRPSO	3.6	4.467	5.38
DDSRPSO	2	3.27	3.38
Imp. DDSRPSO	2	2.33	2.87

Table 4. Bonferroni-Dunn test results with 95% confidence level

Algorithm	Average rank difference w.r.t Imp. DDSRPSO
SLPSO	1.46
GPSO	4.5
LPSO	4
FIPS	3.71
DMSPSO	1.78
CLPSO	2.86
SRPSO	2.14
DDSRPSO	0.64
Critical Difference	1.82

others cannot be concluded by just looking at the ranks. [11]. Therefore a statistical analysis has also been performed to check the significance. First, the Friedman test has been conducted on the mean performances to examine the null hypothesis that whether the performance of Improved DD-SRPSO is same as all the other PSO variants. The computed p value for 95% confidence level is less than $\alpha = 0.05$, so we can reject the null hypothesis and conclude that the mean performances of the 9 algorithms in question are different. As the null hypothesis has been rejected, the Bonferroni-Dunn test [19] can now be used to find the winner. The average rank difference of proposed Improved DD-SRPSO with the eight previous algorithms is provided in Table 4. The Bonferroni-Dunn's Critical Difference at 95% confidence level in average ranks is 1.82. It is evident from the Table 4 that the difference between Improved DD-SRPSO and GPSO, LPSO, FIPS, CLPSO, SRPSO is greater than the critical value which shows the

significance of the improvement of the proposed method. For DD-SRPSO, DMS-PSO and SLPSO the difference is not significant for 95% confidence. As Diversified DD-SRPSO is an extension to DD-SRPSO method, in many cases both methods have similar results. SLPSO performs better than the proposed method for many composite functions whereas DMS-PSO gives comparable results for multimodal functions as it was designed to solve them. Overall, the Improved DD-SRPSO is significantly better than some methods where as the difference is not significant for three methods but it does improve the performance on some problems as suggested by the overall rank in Table 1.

6 Conclusion

This paper presents a hybrid algorithm combining the directionally updated and rotationally invariant update strategies presented in DD-SRPSO [11] with a strategy to address premature convergence inspired from [7], referred to as Improved DD-SRPSO. The performance of Improved DD-SRPSO is evaluated on 28 benchmark optimization problems mentioned in CEC2013 [12]. The performance is compared with 8 other PSO variants including DD-SRPSO. As the proposed method possesses the properties of DD-SRPSO as well as ARPSO it performs quite well on rotated multimodal functions because the ability to escape local minima keeps the exploration going which in turn makes it possible for the algorithm to find a better region of the solution space. The involvement of only a part of the population in the exploration phase while rest of the particles keep exploiting the current best region makes it possible for the algorithm to improve overall exploration without losing the ability to exploit effectively.

References

1. Wari, E., Zhu, W.: A survey on metaheuristics for optimization in food manufacturing industry. Appl. Soft Comput. **46**, 328–343 (2016)
2. Kennedy, J.: Particle swarm optimization. In: Sammut, C., Webb, G.I. (eds.) Encyclopedia of machine learning, pp. 760–766. Springer, Heidelberg (2011). https://doi.org/10.1007/978-0-387-30164-8
3. Cheng, R., Jin, Y.: A social learning particle swarm optimization algorithm for scalable optimization. Inf. Sci. **291**, 43–60 (2015)
4. Evers, G.I., Ghalia, M.B.: Regrouping particle swarm optimization: a new global optimization algorithm with improved performance consistency across benchmarks. In: IEEE International Conference on Systems, Man and Cybernetics. SMC 2009, pp. 3901–3908. IEEE (2009)
5. Peram, T., Veeramachaneni, K., Mohan, C.K.: Fitness-distance-ratio based particle swarm optimization. In: Proceedings of the 2003 IEEE Swarm Intelligence Symposium. SIS 2003, pp. 174–181. IEEE (2003)

6. Xinchao, Z.: A perturbed particle swarm algorithm for numerical optimization. Appl. Soft Comput. **10**(1), 119–124 (2010)
7. Riget, J., Vesterstrøm, J.S.: A diversity-guided particle swarm optimizer-the ARPSO. Department of Computer Science, University of Aarhus, Aarhus, Denmark. Technical report 2 (2002)
8. Krink, T., VesterstrOm, J.S., Riget, J.: Particle swarm optimisation with spatial particle extension. In: Proceedings of the 2002 Congress on Evolutionary Computation. CEC 2002, vol. 2, pp. 1474–1479. IEEE (2002)
9. Wang, H., Li, H., Liu, Y., Li, C., Zeng, S.: Opposition-based particle swarm algorithm with Cauchy mutation. In: IEEE Congress on Evolutionary Computation. CEC 2007, pp. 4750–4756. IEEE (2007)
10. Van Den Bergh, F.: An analysis of particle swarm optimizers. Ph.D. thesis, University of Pretoria (2007)
11. Tanweer, M.R., Auditya, R., Suresh, S., Sundararajan, N., Srikanth, N.: Directionally driven self-regulating particle swarm optimization algorithm. Swarm Evol. Comput. **28**, 98–116 (2016)
12. Liang, J., Qu, B., Suganthan, P., Hernández-Díaz, A.G.: Problem definitions and evaluation criteria for the CEC 2013 special session on real-parameter optimization. Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou, China and Nanyang Technological University, Singapore. Technical report 201212, 3–18 (2013)
13. Tanweer, M.R., Suresh, S., Sundararajan, N.: Self regulating particle swarm optimization algorithm. Inf. Sci. **294**, 182–202 (2015)
14. Liang, J.J., Suganthan, P.N.: Dynamic multi-swarm particle swarm optimizer with local search. In: The 2005 IEEE Congress on Evolutionary Computation, vol. 1, pp. 522–528. IEEE (2005)
15. Kennedy, J., Mendes, R.: Population structure and particle swarm performance. In: Proceedings of the 2002 Congress on Evolutionary Computation. CEC 2002, vol. 2, pp. 1671–1676. IEEE (2002)
16. Shi, Y., Eberhart, R.C.: Empirical study of particle swarm optimization. In: Proceedings of the 1999 Congress on Evolutionary Computation. CEC 1999, vol. 3, pp. 1945–1950. IEEE (1999)
17. Liang, J.J., Qin, A.K., Suganthan, P.N., Baskar, S.: Comprehensive learning particle swarm optimizer for global optimization of multimodal functions. IEEE Trans. Evol. Comput. **10**(3), 281–295 (2006)
18. Mendes, R., Kennedy, J., Neves, J.: The fully informed particle swarm: simpler, maybe better. IEEE Trans. Evol. Comput. **8**(3), 204–210 (2004)
19. Demšar, J.: Statistical comparisons of classifiers over multiple data sets. J. Mach. Learn. Res. **7**(Jan), 1–30 (2006)

Judgement of Learning for Metacognitive Type-2 Fuzzy Inference System

Khyati Mahajan[✉]

Dhirubhai Ambani Institute of Information and Communication Technology,
Gandhinagar, Gujarat, India
`khyati7mahajan@gmail.com`

Abstract. This paper proposes the McIT2FIS-JOL algorithm for oceanic wave data prediction. The learning algorithm is preceded by a meta-memory layer for judgement of learning (JOL). The input is divided into subsets. Each subset is learnt by the learning algorithm, which in turn is regulated by the metacognitive layer. Each model begins with zero rules. The underlying learning algorithm then employs prediction error and novelty of sample to add rules to the network. The metacognitive component of the underlying learning algorithm determines whether each incoming sample is learnt, deleted, or reserved for future use. Once the models have learnt from their subsequent subsets, the learning of the input provided to each model is judged by employing JOL. The model whose JOL measure is the best is then utilised for processing the entire input. The performance of McIT2FIS-JOL is evaluated on a real world oceanic wave data prediction problem. It is compared with MLP, SAFIS, and its predecessor McIT2FIS. The results indicate that significant learning is possible utilising a smaller sample of incoming data, improving learning time.

Keywords: Interval Type-2 Fuzzy systems · Meta-cognition
Meta-memory · Judgement of learning · Oceanic wave data prediction

1 Introduction

Human learning of everything around us is greatly characterised by neuro-plasticity. What we learn, how we learn, and when we learn is directly linked to what we know. If new information is provided, we learn it. If information we knew has a new aspect to it, we update our knowledge. If we already know information that is being provided to us, we need not relearn it. This is termed as the control-monitoring version of learning, in Nelson and Narens' study [11].

Nelson and Narens further extend the idea of study into meta-memory, which were initially referred to by Hart [9] as the "memory management system". The concept states that for the control mechanisms to regulate the system effectively, the current state of the system needs to be known. Nelson and Narens break this knowledge of learning down into three parts, to form a distinction between retrospective and prospective monitoring:

1. *Ease of Learning Judgement* - before learning, to gain an understanding of what will be easier to learn.
2. *Judgement of Learning* - during or after learning, to check for how well future predictions will be based on currently learned knowledge of the learner
3. *Feeling of Knowing Judgements* - during or after learning, to get a feeling of how well knowledge gleaned will be retained.

Machine learning has come a long way in emulating aspects of how to learn knowledge in a way similar to humans. But there is still a major gap with how the knowledge learned by a human and a computer differs - a human can learn how something works with much less input than is required by a machine to reach similar levels of accuracy while testing. A striking example is how a bot might learn winning strategies required during gameplay - deep neural networks often applied to such tasks require billions of bytes to learn what a human might learn from a few games (or even invent some new strategies along the way).

Learning efficiently, therefore, has become a task of paramount importance. To combat overfitting of learning models learnt computationally, and make the learning process more streamlined, the metacognitive algorithm was proposed. Emulating the control-monitoring concept for machines, the meta-cognitive model for sequential learning was proposed in [14]. It has been shown to out-perform quite a few models, and has been developed further for computational efficiency, showcased in [1,2,4–6,12,13].

Motivated by trying to bridge the gap between human and machine learning, this paper proposes an extension to the learning methods derived by the pre-ceding publications. It aims to supplement the learning process so as to require less data for learning; and yet produce results similar, if not better, to the ones obtained earlier. It aims to do so by introducing the concept of meta-memory, which means that the algorithm has information about its memory. In essence, it aims to develop the algorithm so as to evaluate how well the learning algorithm is trained by the input it receives.

The meta-cognitive algorithms simulated the concept of cognition about cog-nition. In effect, the learning algorithm is aware of the knowledge it already possesses, and monitors further learning based on this knowledge. It therefore consists of two levels of learning: the meta-cognitive component which contains a dynamic model of the cognitive component, self-regulated thresholds and knowl-edge measures; and the cognitive component which consists of the Interval Type-2 Fuzzy Inference System [6].

The IT2FIS has been chosen because it has proven abilities to handle noise and uncertainties in data that needs to be processed serially [4,5]. It also demon-strates good performance over known sequential processing algorithms such as OS-ELM. The structure employs interval type-2 sets in rule antecedents. The learning algorithm begins with no rules. Rules are added when a new sam-ple arrives. The structure and parameters of the network of rules is updated based on the knowledge content of the sample. The meta-cognitive component is responsible for determining the knowledge content of the sample.

The proposed algorithm adds a layer of meta-memory over the aforementioned layers of learning. It takes in input data and divides it into a given number n of subsets. It then trains n models using these subsets. Then, it compares the performance of the n trained models using the Goodman-Kruskal gamma coefficient [8], determined by Nelson [10] to be the best measure of the judgement of learning of an agent of learning. The algorithm compares both the training as well as the testing gamma coefficient, to gain knowledge of how well the models have learnt the data and how well they are able to predict data based on their training. It then picks the models which demonstrates the best gamma coefficient, and then runs the entire dataset on this model. The performance comparison results for this have been presented in a separate section.

The algorithm is utilised to predict real world oceanic wave data. The performance comparison with the Multi-Layer Perceptron (MLP), the Sequential Adaptive Fuzzy Inference System (SAFIS), and the Meta-cognitive Interval Type-2 Fuzzy Inference System (McIT2FIS) indicates efficient performance of the proposed algorithm.

The rest of the paper is organised as follows. In Sect. 2, the architecture of the McIT2FIS-JOL is presented. In Sect. 3, the learning algorithm of McIT2FIS-JOL is discussed. In Sect. 4, the performance evaluations and comparisons are presented. Section 5 is the conclusion.

2   Architecture of McIT2FIS-JOL

The underlying learning algorithm, which realises the IT2FIS algorithm, consists of 5 layers which process data from the input to the output. The aim is to determine a functional relationship from input $\mathbf{x} \in \mathfrak{R}^{1 \times m}$ to output $\mathbf{y} \in \mathfrak{R}^{1 \times n}$. We assume that the input consists of m features and the output consists of n features. Further assuming that the system has processed $t - 1$ samples and added K rules, the inference mechanism of the tth sample is showcased below:

Layer 1—*Input layer*: Consists of m nodes. The output of the i-th node is given by

$$u_i(t) = x_i(t) \tag{1}$$

where i ranges from 0 to m.

Layer 2—*Fuzzification layer*: Fuzzifies inputs by employing an Interval Type-2 Gaussian membership function. The lower and upper bounds of the footprint of uncertainty of the membership function are given by:

$$\phi_{ki}^{up}(t) = \begin{cases} \phi(\mu_{ki}^l, \sigma_{ki}, u_i(t)) & u_i(t) < \mu_{ki}^l \\ 1 & \mu_{ki}^l \leq u_i(t) \leq \mu_{ki}^r \\ \phi(\mu_{ki}^r, \sigma_{ki}, u_i(t)) & u_i(t) > \mu_{ki}^r \end{cases} \tag{2}$$

$$\phi_{ki}^{lo}(t) = \begin{cases} \phi(\mu_{ki}^r, \sigma_{ki}, u_i(t)) & u_i(t) \leq \frac{\mu_{ki}^l + \mu_{ki}^r}{2} \\ \phi(\mu_{ki}^l, \sigma_{ki}, u_i(t)) & u_i(t) > \frac{\mu_{ki}^l + \mu_{ki}^r}{2} \end{cases} \tag{3}$$

where the primary Gaussian function is:

$$\phi(\mu_{ki}, \sigma_{ki}, u_i) = \exp\left(-\frac{1}{2}\left(\frac{u_i - \mu_{ki}}{\sigma_{ki}}\right)^2\right)$$

where μ_{ki} is the centre, and σ_{ki} is the width, of the ith feature in the kth rule.

Layer 3—*Firing layer*: Each node contains in the lower and upper firing strengths of a rule. The output of the layer is:

$$F_k^{lo}(t) = \prod_{i=1}^{m} \phi_{ki}^{lo}(t) \text{ and } F_k^{up}(t) = \prod_{i=1}^{m} \phi_{ki}^{up}(t) \tag{4}$$

where F_k^{lo} and F_k^{up} represent the lower and upper firing strengths of the kth rule.

Layer 4—*Output processing layer*: The output of this layer is determined by the combination of the output of layer 4 with consequent weight parameters, and adjusts the contribution of the rules using control parameters q^l and q^r. The output y^l and y^r of this layer is:

$$y_j^l(t) = \frac{(1 - q_j^l)\sum_{k=1}^{K} w_{jk}^l F_k^{lo}(t) + q_j^l \sum_{k=1}^{K} w_{jk}^l F_k^{up}(t)}{\sum_{k=1}^{K}(F_k^{lo}(t) + F_k^{up}(t))} \tag{5}$$

$$y_j^r(t) = \frac{(1 - q_j^r)\sum_{k=1}^{K} w_{jk}^r F_k^{lo}(t) + q_j^r \sum_{k=1}^{K} w_{jk}^r F_k^{up}(t)}{\sum_{k=1}^{K}(F_k^{lo}(t) + F_k^{up}(t))} \tag{6}$$

where $j = 1, \ldots, n$.

Layer 5—*Output layer*: This layer computes the final output given by:

$$\hat{y}_j(t) = y_j^l(t) + y_j^r(t) \tag{7}$$

where $j = 1, \ldots, n$.

The next section details the meta-cognitive and meta-memory concepts of the algorithm.

3   Meta-cognitive and Meta-memory Concepts for McIT2FIS-JOL

Meta-cognitive Strategy

The objective of the learning algorithm is to utilise training data $[(\mathbf{x}(1), \mathbf{y}(1)), \ldots, (\mathbf{x}(t), \mathbf{y}(t)), \ldots]$ and train a model by learning rules and associated parameters, such that the model is able to determine a functional relationship between the input and output.

The error $\mathbf{e}(t) = [e_1(t) \ldots e_n(t)]$ of the tth sample is given by:

$$e_j(t) = y_j(t) - \hat{y}_j(t) \tag{8}$$

where $\mathbf{y}(t)$ is the actual output and $\hat{\mathbf{y}}(t)$ is the predicted output.

A detailed explanation of the meta-cognitive strategies is given in [14]. A brief explanation of how the meta-cognitive component works is provided here.

To determine the knowledge content of a sample with respect to the knowledge contained in the network, the following measures are utilised:

1. **The prediction error:** For the current sample, the prediction error is determined by

$$E(t) = \frac{1}{2}\mathbf{e}^2(t) \tag{9}$$

2. **Spherical potential:** Measures novelty of the current sample by calculating average distance of sample from the existing rules in the network:

$$\psi(t) = \sum_{k=1}^{K} \frac{F_k^{lo}(t) + F_k^{up}(t)}{2K} \tag{10}$$

3. **Sample participation:** Measures the relative importance of the sample to be learnt, given by

$$\beta = \frac{E_d - E_l}{R}(epoch\_count) + E_l \tag{11}$$

   where E_d is the delete threshold, E_l is the learning threshold, and R is the number of reserved samples.

Based on these three measures, the following decisions could be made with respect to the current sample for the learning algorithm.

3.1   Sample Delete Strategy

If for a sample $E(t) < E_d$, where E_d is the delete threshold, it is deleted. The strategy prevents over-fitting to data by deleting samples that can be predicted by the network without much error. A higher value of E_d means more samples are deleted without being learnt. Therefore in general the value of E_d lies in the range [0.0001, 0.001] [5].

3.2   Sample Learning Strategy

Sample learning is for one of two things for the network - it either involves the updation of the parameters of a rule, or the addition of a new rule. Both these are explained in detail below:

– **Parameter Update:** If for a sample $E(t) > E_l$, where E_l is the learning threshold, the sample is utilised for updating the parameters of the network instead of rule addition. The learning threshold is self-adapted as follows:

$$E_l = max(\beta, E_d) \tag{12}$$

which enables the threshold to determine how close the sample is to the rules that already exist in the network. Hence if the sample lies close enough to the network, yet contains information that is new to the network, the sample contributes to parameter updation.

The value of the learning threshold is a tradeoff for efficient learning. Too low, and the network results in more rule updates, risking a generalisation error. Too high, and the network results in less rule updates, putting the network at a risk of under performance. Hence E_l is generally chosen to lie in the range [0.04, 0.2].

– **Rule Growing:** If $E(t) > E_a$ and $\psi(t) < E_s$, where E_a is the self-adaptive addition threshold and E_s is the novelty threshold, the sample is utilised for rule growth.

Higher value of E_a means the network adds fewer rules which may affect generalisation, whereas a lower value means the network structure will suffer due to less rule addition. The self-regulation allows it to globally keep adding knowledge as well as fine tune what the network learns. It is adapted according to the following equation [5]:

$$E_a = (1 - \delta)E_a + \delta E(t) \tag{13}$$

where δ is the slope parameter and decides the rate of increase of E_a (initially set to 0).

A value of E_s closer to 0 means very few rules will be added to the network, whereas a value closer to 1 means the current sample will not be judged as having much new information.

A higher E_a value and a lower E_s value shows resistance to rule addition. Therefore the value of E_a is chosen to lie in [0.10, 0.30] and the value of E_s is chosen to lie between [0.01, 0.60].

3.3 Sample Reserve Strategy

When a sample is neither deleted nor learnt, it is reserved; to be later judged upon in subsequent epochs.

Meta-memory Strategy

The aim of the meta-memory concept is to be aware of the knowledge learnt by the network on the basis of the input training samples that have already been learnt. The basic assumptions and working of the model is explained below.

3.4 The Gamma Coefficient

The Goodman-Kruskal gamma coefficient (γ) [8] was quoted by Nelson and Narens in [11] as being the best way to measure the judgement of learning for human learning. The gamma coefficient essentially takes into account the concordant pairs and reversed pairs, thus measuring what in lay man terms would

mean correct and false alarms. The equation for this calculation is explained below:

$$\gamma = \frac{N_s - N_d}{N_s + N_d} \qquad (14)$$

where N_s denotes the number of concordant observations and N_d denotes the number of reverse observations.

This, however, is modified to the following for use as a computational measures:

$$\gamma = \frac{ad - bc}{ad + bc} \qquad (15)$$

where a is the number of true positives, b is the number of false positives, c is the number of false negatives, and d is the number of true negatives. To avoid any division by zero errors, all of the above variables are initialised at 0.5.

The γ, by taking in the correct and wrong predictions made by the system, proves to be a simple method for reflecting the judgement of learning of the system. However, while it seems to be a straightforward calculation for classification problems, the calculation for approximation problems does not seem as straightforward. The method adopted for this paper is highlighted next.

3.5   Gamma Calculation for Approximation

To adapt the γ calculation for approximation problems, the following strategy is followed:

True positives: $\hat{y}_i = \mu^y \pm \sigma^y; \epsilon_i = \mu^\epsilon \pm \sigma^\epsilon$
False positives: $\hat{y}_i \neq \mu^y \pm \sigma^y; \epsilon_i = \mu^\epsilon \pm \sigma^\epsilon$
False negatives: $\hat{y}_i = \mu^y \pm \sigma^y; \epsilon_i \neq \mu^\epsilon \pm \sigma^\epsilon$
True negatives: $\hat{y}_i \neq \mu^y \pm \sigma^y; \epsilon_i \neq \mu^\epsilon \pm \sigma^\epsilon$

where

μ^y denotes the mean of the target output
μ^ϵ denotes the mean of the error
σ^y denotes the standard deviation of the target output
σ^ϵ denotes the standard deviation of the error
$\hat{y}_i$ denotes the predicted output sample
ϵ_i denotes the error of the sample

With this, the working of the JOL algorithm is shown next.

3.6   Judgement of Learning (JOL)

The JOL code takes in the training and testing output, the predicted outputs and the number of divisions for the subsets n as an input. It then involves the following steps:

1. Divide the training data into n subsets by sampling them. This is done so that the structure of the training data, if there is one, is not completely lost.

2. Run the learning algorithm for n models learning from these n subsets.
3. Calculate the training and testing γ for each of the n models.
4. Choose the model with the best training and testing γ. If there is no model with both measures the best, the higher γ is utilised for choosing the model.
5. The entire training and testing data is then run on the chosen model.

The results of McIT2FIS-JOL are presented in the next section.

4   Performance Evaluation on Oceanic Wave Data

Many renewable sources of energy have been discovered and discussed for usage. Most of these require high capital usage to set up, like solar panels; relocation, like hydro-energy; installation of huge equipment, like wind farms. However, there is another source of energy which has not yet been tapped into much, and is a point of ongoing research considering its potential - oceanic wave energy. It has been shown that oceanic wave energy is quite viable [15], and has potential. Experiments have been performed to validate this potential [7]. To improve the efficiency of these methods, wave data analysis and prediction would prove useful for researching efficient methods for extraction of energy, especially since wave data can be analysed utilising finite features. To this end, the project employs Machine Learning methods to analyse and predict wave data.

Oceanic wave data is a time series distribution, having four main features, namely:

1. *Significant Wave Height (SWH)*
2. *Peak Wave Period (PWP)*
3. *Mean Wave Period (MWP)*
4. *Peak Wave Direction (PWD)*

Oceanic wave energy is energy stored in waves, which essentially originates from solar energy. Heating of air caused by solar heating caused winds, which in turn cause waves when wind blows on large stretches of water. This is documented well in [15].

For conditions that are predictable (such as storm-free conditions) over a considerable period of time, the energy $E(W_h)$, per unit wavelength in the direction of the wave, per unit width of wave front is given by:

$$E = \frac{1}{16} \frac{\rho g^2}{\pi} (H^2 T^2) \tag{16}$$

where ρ is the density of sea water $(\frac{kg}{m^3})$, g is the acceleration due to gravity $(\frac{m^2}{s})$, H is the crest wave height (m) and T is the wave period (s^{-1}).

The power P per unit width of a wave front is given by [3]:

$$P = \frac{1}{64} \frac{\rho g^2}{\pi} (H_s^2 T_e^2) \tag{17}$$
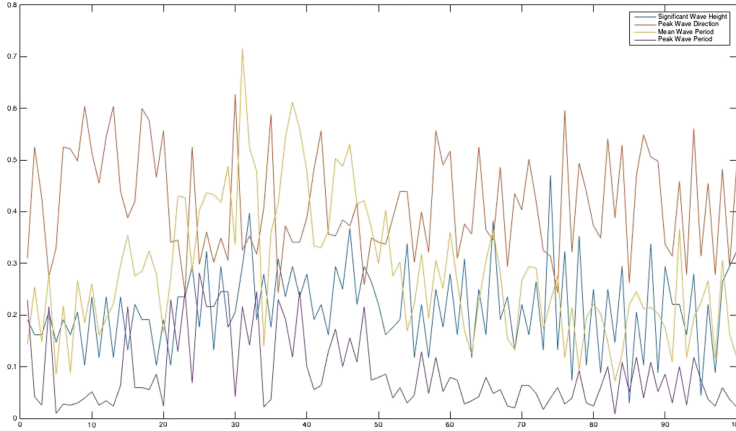
Fig. 1. A sampled subset data plot for oceanic wave data, showing all the variables associated with the wave metrics

where P is in $(\frac{W}{m})$, H_s is the significant wave height (m) and T_e is the wave period (s^{-1}). Figure 1 shows a sampled subset of the oceanic wave data.

Visualisations for the performance metrics are given next. In brief, Table 1 shows the average absolute training error exhibited when the oceanic wave data is run through the algorithms Table 2 shows the similar error for a test run.

Table 1. Oceanic wave data - training

Features in wave data	Training average error			
	MLP	*SAFIS*	*McIT2FIS*	*McIT2FIS-JOL*
SWH	0.0720	0.0455	0.0272	0.0282
PWD	0.0910	0.0775	0.0617	0.0630
MWP	0.1274	0.0720	0.0390	0.0404
PWP	0.0615	0.8555	0.0531	0.0498

Table 2. Oceanic wave data - testing

Features in wave data	Testing average error			
	MLP	*SAFIS*	*McIT2FIS*	*McIT2FIS-JOL*
SWH	0.1828	0.1772	0.0633	0.0670
PWD	0.0781	0.0703	0.0569	0.0571
MWP	0.1934	0.1215	0.0735	0.0784
PWP	0.0349	0.7360	0.0352	0.0388

Table 3. Time elapsed

Running time	Algorithms (both training and testing)			
	MLP	*SAFIS*	*McIT2FIS*	*McIT2FIS-JOL*
SWH	104.14	47.505	9.8674	5.7941
PWD	105.59	64.541	9.8541	4.2034
MWP	104.64	48.523	8.5082	4.9484
PWP	105.53	53.688	10.643	7.5620
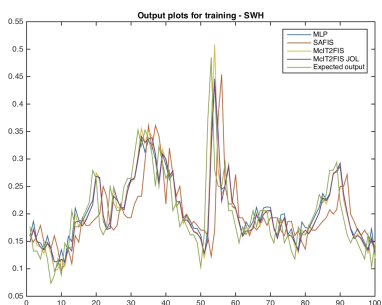
Fig. 2. Subset visualisation of output for SWH (Training)
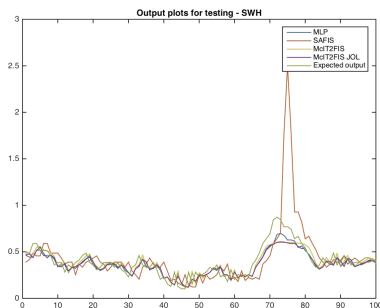
Fig. 3. Subset visualisation of output for SWH (Testing)
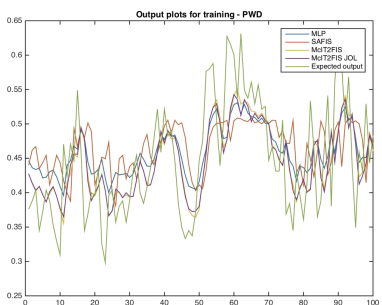
Fig. 4. Subset visualisation of output for PWD (Training)
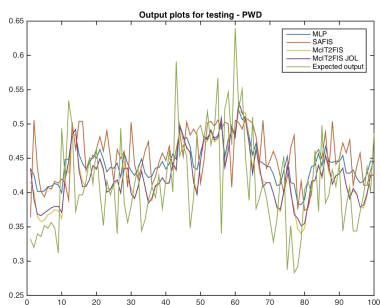
Fig. 5. Subset visualisation of output for PWD (Testing)

Table 3 shows the time efficiency achieved by JoL-McIT2FIS. Figures 2 and 3 show the subset visualisation of the Significant Wave Height outputs given by the algorithms during training and testing, respectively. Figures 4 and 5 do the same for the Peak Wave Direction, Figs. 6 and 7 do the same for Peak Wave Period, and Figs. 8 and 9 similarly do so for the Mean Wave Period.
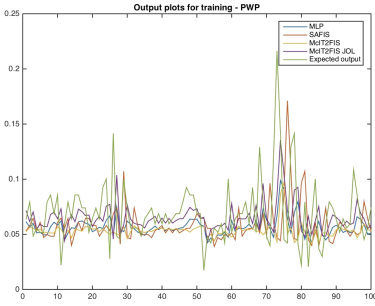
Fig. 6. Subset visualisation of output for PWP (Training)
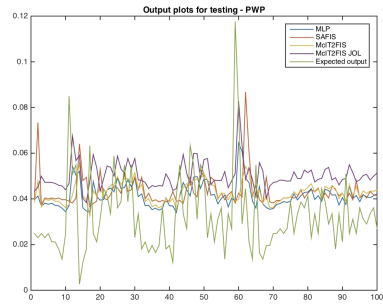
Fig. 7. Subset visualisation of output for PWP (Testing)
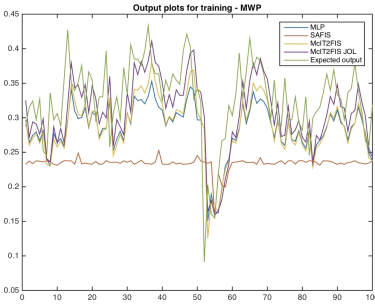
Fig. 8. Subset visualisation of output for MWP (Training)
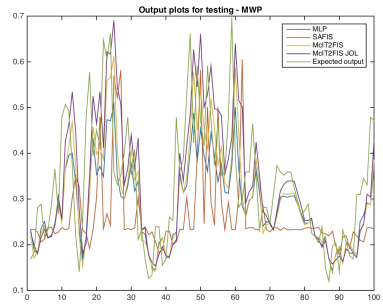
Fig. 9. Subset visualisation of output for MWP (Testing)

5   Conclusion

This paper has presented the McIT2FIS-JOL learning algorithm. The performance is evaluated on 4 different features with differently structured data and noise. The results show that by evaluating the learning of the algorithm model, one can train models on lesser data (20% of the total data in this case); thereby saving time (about 40–48% in this case) while keeping the error metrics in check.

References

1. Babu, G.S., Savitha, R., Suresh, S.: A projection based learning in meta-cognitive radial basis function network for classification problems. In: The 2012 International Joint Conference on Neural Networks (IJCNN), pp. 1–8. IEEE (2012). https://doi.org/10.1109/IJCNN.2012.6252769
2. Babu, G.S., Suresh, S.: Meta-cognitive neural network for classification problems in a sequential learning framework. Neurocomputing **81**, 86–96 (2011). https://doi.org/10.1016/j.neucomm.2011.12.001

3. Bahaj, A.S.: Generating electricity from the oceans. Renew. Sustain. Energy Rev. **15**(7), 3399–3416 (2011)
4. Das, A.K., Anh, N., Suresh, S., Srikanth, N.: An interval type-2 fuzzy inference system and its meta-cognitive learning algorithm. Evol. Syst. **7**, 95–105 (2016). https://doi.org/10.1007/s12530-016-9148-6
5. Das, A.K., Subramanian, K., Sundaram, S.: An evolving interval type-2 neuro-fuzzy inference system and its metacognitive sequential learning algorithm. IEEE Trans. Fuzzy Syst. **23**(6), 2080–2093 (2015). https://doi.org/10.1109/TFUZZ.2015.2403793
6. Das, A.K., Suresh, S., Srikanth, N.: Meta-cognitive interval type-2 neuro-fuzzy inference system for wind prediction. In: 2014 International Conference on Multi-sensor Fusion and Information Integration for Intelligent Systems (MFI), pp. 1–6 (2014)
7. Eriksson, M., Waters, R., Svensson, O., Isberg, J., Leijon, M.: Wave power absorption: experiments in open sea and simulation. J. Appl. Phys. **102**(8), 084910 (2007). https://doi.org/10.1063/1.2801002
8. Goodman, L.A., Kruskal, W.H.: Measures of association for cross classifications. J. Am. Stat. Assoc. **49**(268), 732–764 (1954). https://doi.org/10.1007/978-1-4612-9995-0_1
9. Hart, J.T.: Memory and the memory-monitoring process. J. Verbal Learn. Verbal Behav. **6**(5), 685–691 (1967)
10. Nelson, T.O.: Metamemory: a theoretical framework and new findings. Psychol. Learn. Motiv. **26**, 125–173 (1990). https://doi.org/10.1016/S0079-7421(08)60053-5
11. Nelson, T.O., Narens, L.: Why investigate metacognition? In: Metacognition: Knowing About Knowing, pp. 1–25 (1994)
12. Savitha, R., Suresh, S., Sundararajan, N.: A meta-cognitive learning algorithm for a fully complex-valued relaxation network. Neural Netw. **32**, 209–218 (2012). https://doi.org/10.1016/j.neunet.2012.02.015
13. Subramanian, K., Das, A.K., Sundaram, S., Ramasamy, S.: A meta-cognitive interval type-2 fuzzy inference system and its projection based learning algorithm. Evol. Syst. **5**(4), 219–230 (2014). https://doi.org/10.1007/s12530-013-9102-9
14. Suresh, S., Subramanian, K.: A sequential learning algorithm for meta-cognitive neuro-fuzzy inference system for classification problems. In: 2011 International Joint Conference on Neural Networks (IJCNN), pp. 2507–2512. IEEE (2011). https://doi.org/10.1109/IJCNN.2011.6033545
15. Waters, R., Stålberg, M., Danielsson, O., Svensson, O., Gustafsson, S., Strömstedt, E., Eriksson, M., Sundberg, J., Leijon, M.: Experimental results from sea trials of an offshore wave energy system. Appl. Phys. Lett. **90**(3), 034105 (2007). https://doi.org/10.1063/1.2432168

Inter Intensity and Color Channel Co-occurrence Histogram for Color Texture Classification

S. Shivashankar[1] , Madhuri R. Kagale[1(✉)] ,
and Prakash S. Hiremath[2]

[1] Karnatak University, Dharwad 580003, Karnataka, India
shivashankars@kud.ac.in, madhuri.kagale@gmail.com
[2] KLE Technological University, BVBCET, Hubballi 580031, Karnataka, India
hiremathps53@yahoo.com

Abstract. In this paper we propose a new method to analyze the color texture image based on inter intensity and color channel co-occurrence histogram, which characterizes the color texture more effectively. This corresponds to the relationships between intensity and color channel along with their neighboring pixels. The proposed color texture descriptor is experimented on VisTex texture dataset. The results are analyzed and compared with Local Binary Patterns (LBP) method and Histogram ratio method. The computational intelligence-based approach, namely, fuzzy classification is used for texture classification. The proposed descriptors achieve better classification results when compared with other two methods. The proposed color texture descriptors are sufficiently robust and precise to distinguish images of different textures even if the sample size is small. The results suggest that the proposed color texture descriptors have the potential for use in real-world applications involving recognition of patterns in digital images.

Keywords: Color texture · Texture classification
Inter co-occurrence histogram · Color-channel · Inter intensity and color

1 Introduction

In computer vision, the analysis of color and texture is widely studied topic for texture classification and segmentation. In texture analysis, an extensive work has been performed on grayscale images, namely, Co-occurrence matrix [1], Markov random fields [2], Local Binary Patterns [3], Gabor filtering [4] and Wavelet decomposition [5]. Color is an important feature of an image which gives more information than the grayscale value of an image. The classification accuracy of the grayscale images can be improved by incorporating color attribute into texture analysis. The standard gray level texture analysis methods are also extended to extract the features by considering the different color spaces [6, 7], but the computational time of these methods is usually high. Many researchers have worked on the color texture analysis in the literature. Drimbarean and Whelan [8] investigated that the classification of color images gives

better results than gray scale images. Maenpaa and Pietikainen [9] have presented processing methods of combined color and texture information as well as separate usage of color and texture attributes. Paschos and Petrou [10] used ratio features extracted from the color histogram for color texture classification and obtained improved results over the histogram integration method which proves the importance of color in texture pattern. Histogram approach ignores inter-pixel relationships, which is one of the most important properties of texture. Co-occurrence matrix approach describes the relationship between the neighboring pixel values in an image [1], used mostly for the color texture classification. In the literature, different types of color texture feature extraction methods have been exist [11–13].

The color texture classification has been dealt with three different approaches [14]: they are, sequential [15], parallel [16] and integrative [17–19]. Recent approaches focus on the utilization of both intra and inter-channel information to color texture analysis. Palm [14] has described combined color and grayscale textures using integrative co-occurrence matrices for the classification of color texture images. Rosenfled et al. [20] have introduced the concept of multichannel co-occurrence matrices but difficult in computation. However, Vadivel et al. [21] have made an attempt to analyze the similarity of colors and their distribution using an integrated color and intensity co-occurrence matrix.

The objective of this research work is to analyze the similarity of the colors and their distributions at the same time by proposing a color texture descriptor based on the inter intensity and color co-occurrence histogram. The bin values of histogram are considered as color texture descriptors. This corresponds to the relationships between intensity and color channel along with their neighboring pixels and hence characterizes the color texture more effectively. The computational intelligence-based approach, fuzzy system is used for evaluation of classification. The classification performance is tested on VisTex dataset [22] using the proposed inter intensity and color co-occurrence histogram method. The experimental results obtained by the proposed descriptor are compared with the results obtained by using the Local Binary Patterns (LBP) method [3] and Histogram ratio method [10]. The proposed method achieves better classification accuracy for different sample sizes when compared with other two methods.

This paper is organized as follows, the proposed method based on inter intensity and color channel co-occurrence histogram is described in Sect. 2. In Sect. 3, the fuzzy texture classification method is explained. In Sect. 4, the experimental results using the proposed descriptors and performance comparison with other methods are presented. Finally, Sect. 5 concludes the discussion.

2 Proposed Method Based on Inter Intensity and Color Channel Co-occurrence Histogram

The proposed method is to construct a histogram using the intensity and color channel neighborhood relationship of a color image, which captures intensity variations as a texture along with color variations. These histogram bin values are considered as color texture descriptors, which are further used for color texture classification. The proposed algorithm is illustrated diagrammatically in the Fig. 1.

Color image   Intensity image (I)

(I, R)

Blue channel (B)

Green channel (G)

Red channel (R)

Histogram (H)
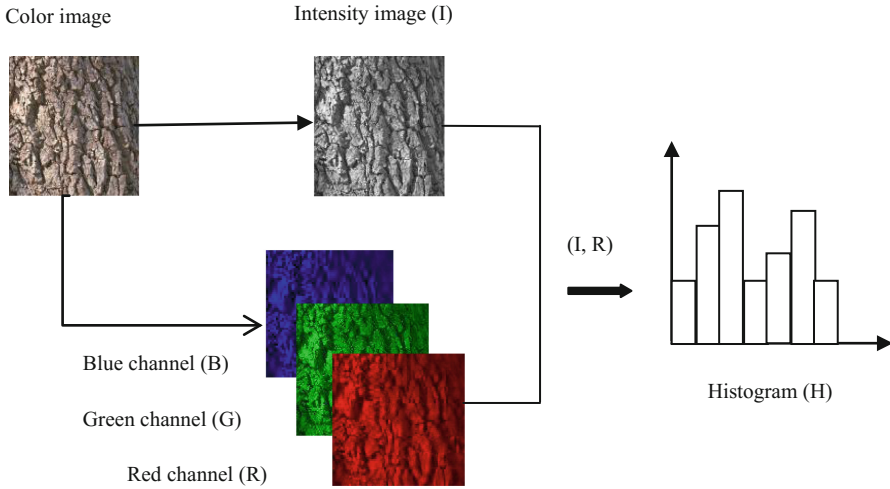
Fig. 1. Schematic diagram of the proposed algorithm. (Color figure online)

The proposed algorithm for computing the color texture descriptors is described below.

Algorithm:

1. Input the color texture image $X = (R, G, B)$
2. Compute $I = 0.299*R + 0.587*G + 0.114*B$ [3].
3. Histogram construction

- Consider the pair of images (I, R)
- Consider a pixel X_{ij} with its 8 neighbors in I and the corresponding pixel Y_{ij} with its 8 neighbors in R as shown below

a_4	a_3	a_2
a_5	X_{ij}	a_1
a_6	a_7	a_8

I

b_4	b_3	b_2
b_5	Y_{ij}	b_1
b_6	b_7	b_8

R

- Initialize vector of nb bins, $H = 0$;

```
for i = 1 to P
    for j = 1 to Q
        Let M = maximum (minimum (X_ij, b_1), minimum (Y_ij, a_1));
            If M == minimum (X_ij, b_1)
                H (X_ij) = H (X_ij) + 1;
        end
```

      end
     end
where P, Q is the size of intensity (I) image.

X_{ij} is the ij^{th} element in I.

Y_{ij} is the ij^{th} element in R.

a_1 is the neighbor of X_{ij} at distance d =1 and angle $\theta = 0^{\circ}$.

b_1 is the neighbor of Y_{ij} at distance d=1 and angle $\theta = 0^{\circ}$.

nb is the number of bins considered while histogram construction.

4. The bin values of H are considered as the texture descriptors with respect to R. The experiments are performed for different combinations of, (I, R), (I, G), (I, B), for angle $\theta = 0^{\circ} + 45^{\circ} + 90^{\circ} + 135^{\circ}$ and distance d= 1, 2, 3 using the proposed algorithm.

3 Fuzzy Texture Classification

The classification can be carried out using conventional approach and computational intelligence-based approach [23]. The computational intelligence-based approach includes fuzzy system, neural network, statistical methods [24–26]. In the present study we use the fuzzy classification method on the proposed color texture descriptors. The classification can be performed in two stages, training and testing. The 50% of randomly selected samples from each class are used for training and the remaining 50% for testing.

3.1 Training

The texture features of train samples from each class are extracted using the proposed algorithm. The class mean (μ_{ij}) and standard deviation (σ_{ij}) for all j(j = 1, 2, …, m) features of i(i = 1, 2, …, c) classes are stored in the library, which are further used for texture classification.

  Let $X = (x_1, x_2, …, x_m)$ be the feature vector, where x_j is the jth bin value of histogram of a texture image. We consider the bin values as fuzzy sets with gaussian membership function having mean μ_{ij} and standard deviation σ_{ij}, for i = 1, 2, …, c and j = 1, 2, …, m. given by (1),

$$f_{ij}(x) = exp\left[\frac{-(x - \mu_{ij})^2}{2\sigma_{ij}^2}\right] \tag{1}$$

3.2 Testing

The texture features of test samples from each image are extracted using the proposed algorithm. We apply the fuzzy classification method [27, 28] to classify the test sample. The procedure is described as follows:

1. The jth feature x_j of the test sample is computed using the gaussian membership function as in (2),

$$f_{ij}(X) = exp\left[\frac{-(x_j - \mu_{ij})^2}{2\sigma_{ij}^2}\right] \tag{2}$$

   where x_j is the jth feature of test sample X.

2. Compute $maxf_j = \max_{1 \le i \le c}\{f_{ij}\}$

3. Compute the distance D_i between feature vector of ith class $f_i = (f_{i1}, f_{i2}, \ldots, f_{im})$ and $maxf = (maxf_1, maxf_2, \ldots, maxf_m)$ given in (3),

$$D_i = \sum_j |maxf_j - f_{ij}| \tag{3}$$

   where $i = 1, 2, \ldots, c$.

4. Compute $D_{i^*} = \min_{1 \le i \le c}\{D_i\}$

   where i^* is the class to which min $\{D_i\}$ correspond to and the test sample is classified to the class i^*.

4 Experiments and Discussion

The experimental tests are performed on VisTex color dataset [22]. It has a set of 164 images with size 128×128. Each image is considered as one class. The 20 samples (overlapping) are randomly selected from each image whose size is 100×100. The 50% randomly chosen samples are used in training and remaining 50% samples are used in testing [10]. The results based on different sample sizes are presented for comparative analysis of texture classification. The experiments are conducted based on the methodology presented in the Fig. 1, the features are extracted using the algorithm illustrated in the Sect. 2 and the experimental results are presented in the Tables 1, 2, 3 and 4.

In the Tables 1, 2 and 3, first column indicates the different combination of color channels along with intensity and the other columns indicate the combination of angles/distances. The entries in the table shows the classification accuracy (%) averaged over 10 experiments for different combinations of angles/distances. The Table 1, shows the classification accuracy obtained using the proposed method for 256 bins at distance d = 1 for different combination of angles. The classification accuracy obtained using the proposed method for 100 bins at distance d = 1 for different combination of angles is shown in the Table 2. It is observed from the Tables 1 and 2 that approximately similar results are obtained with 100 bins and 256 bins. There is a reduction in the number of descriptors from 256 to 100. The Table 3, shows the classification accuracy obtained using the proposed method with 100 bins at angle $\theta = 0^\circ + 45^\circ + 90^\circ + 135^\circ$ for different combination of distances which helps to enhance the results compare to results in previous two tables. From the Tables 1, 2

and 3, it is seen clearly that for the combination ((I, R), (I, G), (I, B)), at distance d = 1, 2, and 3 and at an angle $\theta = 0^\circ + 45^\circ + 90^\circ + 135^\circ$ the classification accuracy 98.23% is obtained.

Table 1. Classification accuracy (%) of the proposed method with 256 bins at d = 1, for different combination of angles.

Intensity and color channel combination	Angle			
	0°	$0^\circ + 45^\circ$	$0^\circ + 45^\circ + 90^\circ$	$0^\circ + 45^\circ + 90^\circ + 135^\circ$
(I, R)	96.80	97.38	97.69	97.70
(I, G)	96.03	96.85	96.99	96.97
(I, B)	95.42	96.37	96.83	96.81
(I, R), (I, G)	97.97	98.43	98.47	98.45
(I, R), (I, B)	97.55	98.17	98.35	98.29
(I, G), (I, B)	96.38	97.36	97.61	97.46
(I, R), (I, G), (I, B)	97.70	98.26	98.43	98.29

Table 2. Classification accuracy (%) of the proposed method with 100 bins at d = 1, for different combination of angles.

Intensity and color channel combination	Angle			
	0°	$0^\circ + 45^\circ$	$0^\circ + 45^\circ + 90^\circ$	$0^\circ + 45^\circ + 90^\circ + 135^\circ$
(I, R)	96.48	96.84	97.20	97.21
(I, G)	96.32	96.97	97.16	97.14
(I, B)	95.27	96.15	96.54	96.44
(I, R), (I, G)	97.68	98.13	98.16	98.10
(I, R), (I, B)	97.36	97.79	97.98	97.94
(I, G), (I, B)	96.61	97.23	97.51	97.41
(I, R), (I, G), (I, B)	97.68	98.07	98.12	98.13

Table 3. Classification accuracy (%) of the proposed method with 100 bins at 4 angles ($\theta = 0^\circ + 45^\circ + 90^\circ + 135^\circ$).

Intensity and color channel combination	Distance				
	d = 1	d = 2	d = 3	d = 1, d = 2	d = 1, d = 2, d = 3
(I, R)	97.21	97.12	96.97	97.27	97.32
(I, G)	97.14	96.92	96.54	97.19	97.07
(I, B)	96.44	96.25	96.15	96.48	96.61
(I, R), (I, G)	98.10	98.31	98.25	98.23	98.29
(I, R), (I, B)	97.94	98.00	98.04	98.02	98.10
(I, G), (I, B)	97.41	97.30	97.19	97.52	97.51
(I, R), (I, G), (I, B)	98.13	98.28	98.16	98.24	98.23

Further the results of the proposed method are compared with that of Local Binary Patterns (LBP) and Histogram ratio method. The Table 4 shows the comparative analysis of proposed method along with other two methods for varying sample sizes. The proposed method gives better results than the other two methods. It is observed that, the classification accuracy decreases as the sample size decreases. Even if the 50% of the pattern image is considered, the better classification accuracy can be expected using the proposed method as compared with other two methods. Thus it establishes the existence of correlation between intensity and color channel information along with their corresponding neighbors.

Table 4. Classification accuracy (%) of the proposed method at $d = 1, 2, 3$ and $\theta = 0° + 45° + 90° + 135°$ for the combination (I, R), (I, G), (I, B). Also shown are corresponding classification accuracy when using the Local Binary Patterns (LBP) and Histogram ratio methods.

Size of sample	Proposed method	LBP method [3]	Histogram ratio method [10]
120	99.79	97.62	-
110	99.65	96.42	-
100	98.23	93.52	96.36
80	97.92	91.20	94.36
50	94.31	79.49	67.87
20	56.04	28.01	14.36
10	19.49	7.70	5.09

5   Conclusion

In the present paper a novel method, named inter intensity and color channel co-occurrence histogram, to extract texture features from color images is proposed. The approach is based on the intensity and color channel along with their neighboring pixels of an image. The performance of the proposed method has been tested on VisTex texture dataset. The fuzzy classification method is used for texture classification. The result obtained by the proposed method achieves better classification accuracy as compared with the Local Binary Patterns (LBP) and Histogram ratio method. The inter intensity and color channel descriptors are sufficiently robust and precise to distinguish images of different textures even if the sample size is small. The results suggest that the proposed color texture descriptors have the potential for use in real-world applications involving recognition of patterns in digital images.

References

1. Haralic, R.M., Shanmugam, K.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **3**(6), 610–621 (1973)
2. Wang, L., Liu, J.: Texture classification using multiresolution Markov random field models. Pattern Recogn. Lett. **20**(2), 171–182 (1999)

3. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. **24**(7), 971–987 (2002)

4. Jain, A.K., Farrokhnia, F.: Unsupervised texture segmentation using Gabor filters. Pattern Recogn. **24**(12), 1167–1186 (1991)

5. Wang, J.W., Chen, C.H., Chien, W.M., Tsai, C.M.: Texture classification using non-separable two-dimensional wavelets. Pattern Recogn. Lett. **19**(13), 1225–1234 (1998)

6. Backes, A.R., Casanova, D., Bruno, O.M.: Color texture analysis based on fractal descriptors. Pattern Recogn. **45**(5), 1984–1992 (2012)

7. Paschos, G.: Perceptually uniform color spaces for color texture analysis: an empirical evaluation. IEEE Trans. Image Process. **10**(6), 932–937 (2001)

8. Drimbarean, A., Whelan, P.F.: Experiments in colour texture analysis. Pattern Recogn. Lett. **22**(10), 1161–1167 (2001)

9. Maenpaa, T., Pietikainen, M.: Classification with color and texture: jointly or separately? Pattern Recogn. **37**(8), 1629–1640 (2004)

10. Paschos, G., Petrou, M.: Histogram ratio features for color texture classification. Pattern Recogn. Lett. **24**(1–3), 309–314 (2003)

11. Khan, F.S., Anwer, R.M., van de Weijer, J., Felsberg, M., Laaksonen, J.: Compact color-texture description for texture classification. Pattern Recogn. Lett. **51**, 16–22 (2015)

12. Bianconi, F., Harvey, R., Southam, P., Fernandez, A.: Theoretical and experimental comparision of different approaches for color texture classification. J. Electron. Imaging **20**(4), 043006 (2011)

13. Cusano, C., Napoletano, P., Schettini, R.: Combining multiple features for color texture classification. J. Electron. Imaging **25**(6), 061410 (2016)

14. Palm, C.: Color texture classification by integrative co-occurrence matrices. Pattern Recogn. **37**(5), 965–976 (2004)

15. Kukkonen, S., Kailviaiinen, H., Parkkinen, J.: Color features for quality control in ceramic tile industry. Opt. Eng. **40**(2), 170–177 (2001)

16. Paschos, G.: Fast color texture recognition using chromaticity moments. Pattern Recogn. Lett. **21**(9), 837–841 (2000)

17. Hoang, M.A., Geusebroek, J.M., Smeulders, A.W.: Color texture measurement and segmentation. Signal Process. **85**(2), 265–275 (2005)

18. Jain, A., Healey, G.: A multiscale representation including opponent color features for texture recognition. IEEE Trans. Image Process. **7**(1), 124–128 (1998)

19. Maenpaa, T., Pietikainen, M., Viertola, J.: Separating color and pattern information for color texture discrimination. In: Proceedings of the 16th IEEE International Conference on Pattern Recognition, vol. 1, pp. 668–671 (2002)

20. Rosenfeld, A., Wang, C.Y., Wu, A.Y.: Multispectral texture. IEEE Trans. Syst. Man Cybern. **SMC-12**(1), 79–84 (1982)

21. Vadivel, A., Sural, S., Majumdar, A.K.: An integrated color and intensity co-occurrence matrix. Pattern Recogn. Lett. **28**(8), 974–983 (2007)

22. VisTex: Vision texture database of MIT media lab (1995). http://www-white.media.mit.edu/vismod/imagery/VisionTexture/vistex.html

23. Cateni, S., Colla, V., Vannucci, M., Borselli, A.: Fuzzy inference systems applied to image classification in the industrial field. In: Fuzzy Inference System-Theory and Applications. InTech (2012)

24. Bezdek, J.C.: Computing with uncertainty. IEEE Commun. Mag. **30**(9), 24–36 (1992)

25. Fukunaga, K.: Introduction to Statistical Pattern Recognition. Academic Press, Cambridge (2013)

26. Haykin, S.: Neural Networks: A Comprehensive Foundation, 2nd edn. Prentice Hall PTR, Upper Saddle River (1994)
27. Wang, F.: Fuzzy supervised classification of remote sensing images. IEEE Trans. Geosci. Remote Sens. **28**(2), 194–201 (1990)
28. Keller, J.M., Gray, M.R., Givens, J.A.: A fuzzy k-nearest neighbor algorithm. IEEE Trans. Syst. Man Cybern. **15**(4), 580–585 (1985)

LDA Based Discriminant Features for Texture Classification Using WT and PDE Approach

Rohini A. Bhusnurmath[1](✉) and P. S. Hiremath[2]

[1] Department of Computer Science, Government P.U. College for Girls,
Vijayapur 586101, Karnataka, India
rohiniabmath@gmail.com
[2] Department of Computer Science (MCA), KLE Technological University,
BVBCET Campus, Hubli 580031, Karnataka, India
hiremathps53@yahoo.com

Abstract. Texture classification is an important aspect of image processing and computer vision applications. In this paper, discriminant features for texture classification using WT (wavelet transform) and PDE (partial differential equation) approach is presented. WT is used to obtain directional information of image. A PDE for anisotropic diffusion is used to extract texture component. Different statistical features are computed from texture component. The linear discriminant analysis (LDA) is used to boost the class separability. The discriminant features so obtained are robust class representatives. The robustness of proposed approach is experimented using three gray scale texture datasets: Oulu, Kylberg and VisTex. The k-NN classifier is used to evaluate the classification accuracy. The experimental results exhibit better performance as compared to the other methods in the literature in terms of computational efficiency.

Keywords: Discriminant texture features · Texture classification
PDE · k-NN · Image processing · Wavelet transform

1 Introduction

Texture classification is used in a variety of applications such as medical image analysis, object recognition, content-based image retrieval, remote sensing. There is a need to perform comparative analysis of effective discriminant texture descriptors for different datasets. The objective of the present study is to investigate the robustness of the proposed texture features on Oulu, Kylberg and VisTex gray scale datasets. These datasets differ in class homogeneity, illumination, perspective, number of classes and samples. The proposed method is computationally simple. It is based on the Haar wavelet and anisotropic diffusion, to yield better discriminant features and more accurate classification results.

The four novel contributions of the proposed method are: firstly, to evaluate the robustness of the proposed method using different benchmark datasets. Secondly, to increase the classification accuracy and reduce the computational time. Thirdly, to obtain the discriminant features for texture classification. Lastly, to study the effect of combination of feature sets on texture classification.

2   Background Literature

The wavelet co-occurrence histograms based method for feature extraction is effectively applied to script identification [1]. Texture classification using the local directional binary patterns (LDBP) and nonsubsampled contourlet transform is presented in [2]. The multiresolution LDBPs are explored to investigate the dominant LDBP for texture classification [3]. Perona and Malik [4] introduced the concept of PDE (partial differential equation) based anisotropic diffusion, wherein the inter region smoothing is done keeping the edges sharp. The anisotropic diffusion approach has become a useful tool for texture classification using local directional binary patterns [5], image segmentation [6]. The robustness of texture analysis based on PDE and LDBP is tested on different datasets in [7]. Various feature extraction methods are explored in the literature [8–15]. Several works [16–18] shows that combined texture descriptors perform better in texture classification. Texture classification method based on local vector quantization pattern, to overcome coarse conversion into binary format and noise sensitivity, is proposed in [25]. A set of features to capture discriminative information called dense micro-block difference is introduced. The method considers characteristic structure as small patches in a texture image [26]. Classifier fusion for texture classification irrespective of descriptors are introduced in [27]. LBP and non LBP features using deep convolution are explored for texture classification [28].

Motivated by the previous works in the literature, the focus of present study is to obtain the increased classification accuracy with reduced computational time, so that the proposed method can be applied to real time applications.

3   Proposed Methodology

The proposed method aims to obtain discriminant features for texture classification using WT (wavelet transform) and PDE approach. The WT is used to obtain directional information of image. The anisotropic diffusion is applied to wavelet components (D, V and H) to obtain texture component of image, from which texture features are extracted. The class separability of feature set is enhanced using LDA. The effectiveness of the proposed method is exhibited using three texture datasets. The classification is performed by using k-NN classifier.

3.1   Wavelet Transform

Wavelet transform (WT) is a powerful tool for multi-resolution analysis, plays significant role in texture analysis. The theory and implementation of wavelet based algorithms are discussed in [19, 20]. The continuous WT of a one dimensional signal g(x) is defined by the Eqs. (1) and (2):

$$\left(W_\psi g\right)(a,b) = \int g(x)\psi^*_{(a,b)}(x)dx \tag{1}$$

and

$$\psi_{a,b} = a^{-1/2}\psi((x - a)/b) \tag{2}$$

where b represents the translation parameter, a represents the scaling factor, and $\psi^*(x)$ represents transforming function. The two dimensional WT is defined as a product of 1-D WT, which yields decomposition of image into one approximation image A and three detail images D, H and V. In the present work, the Haar wavelet is employed for the wavelet transform of an image.

3.2 PDE for Anisotropic Diffusion

Perona and Malik [4] proposed scale-space definition through anisotropic diffusion (AD). It is a non-linear PDE that models the diffusion process. PDE-based algorithms conserve the information and structure of the image. The theory of PDEs contains solutions of various mathematical processes by providing stability, convergence. AD filtering successfully smoothes image by preserving the region boundaries and small structures within the image. As proposed in [4], diffusion coefficient takes the form given in Eq. (3) or Eq. (4):

$$g(x) = \frac{1}{1 + (x/k)^2} \tag{3}$$

$$g(x) = \exp\left[-(x/k)^2\right] \tag{4}$$

where g(x) provides threshold for the gradient magnitude. It is the edge preserving parameter in diffusion process.

3.3 Features Extracted

The Table 1 represents the different first order (FS1) and second order (FS2-FS9) statistical features [17, 18], obtained from the wavelet decomposed image applied to anisotropic diffusion.

Table 1. The different first and second order statistical features taken for the experimentation.

Feature set	Description
FS1	First order statistics
FS2	Haralick features [9]
FS3	Gray level difference statistics [8]
FS4	Neighborhood gray tone difference matrix [10]
FS5	Statistical feature matrix [11]
FS6	Law's texture energy measures [12, 13]
FS7	Fractal dimension texture analysis [14]
FS8	Fourier power spectrum [15]
FS9	Shape

3.4   Texture Classification

The texture classification aims to assign the unknown test sample to one of the known texture class. The overview of the proposed method for texture classification is shown in the Fig. 1. It comprises wavelet transform and anisotropic diffusion for feature extraction, LDA for obtaining discriminant features and k-NN classifier [21] for texture classification.
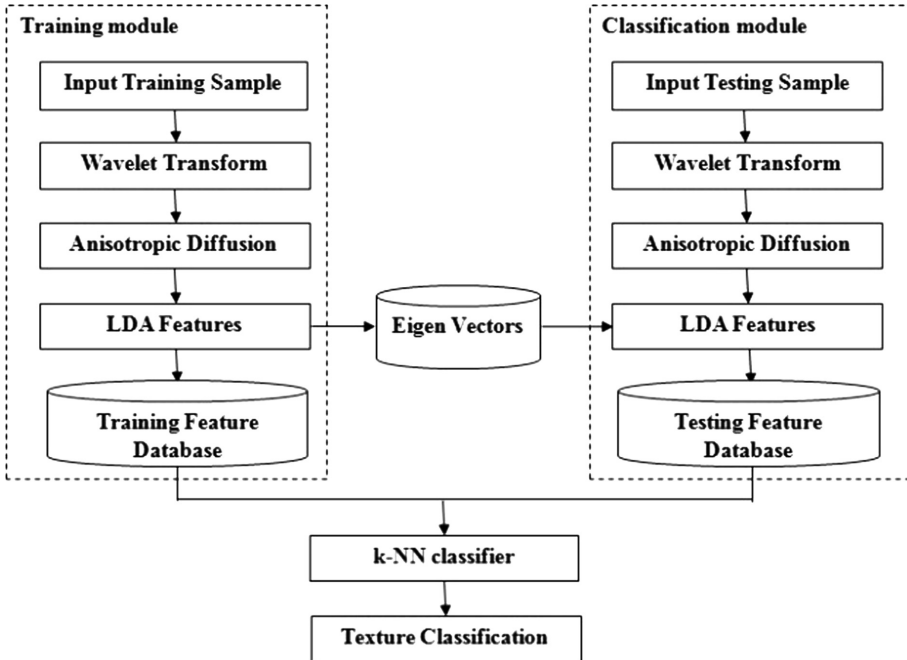
Fig. 1. The overview of proposed method using WT and PDE approach.

4   Experimental Results and Discussion

4.1   Image Datasets

The effectiveness of the proposed method is evaluated using three benchmark texture image datasets, namely, Oulu [22], Kylberg [23] and VisTex [24]. These datasets differ in class homogeneity, illumination, perspective, number of classes and samples. The Fig. 2 shows sample images of the datasets considered for experimentation.

The sixteen gray scale images of Oulu color dataset [22] are used for experimentation. Many textures vary in local color properties, resulting challenging local gray scale variations.

The twenty eight gray classes of 160 samples each, of Kylberg dataset [23] are used for experimentation. The classes have high degree of homogenity in terms of
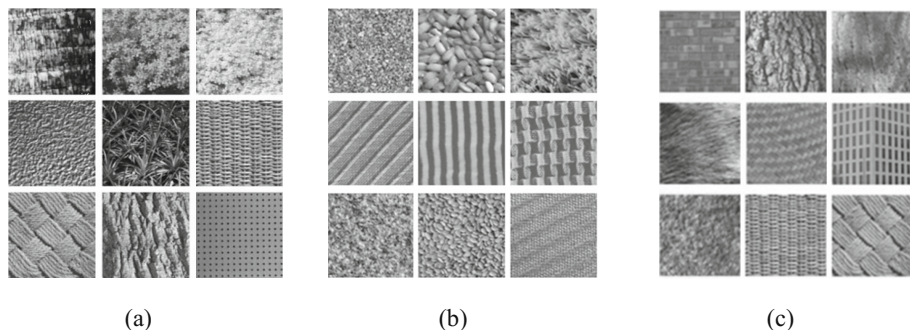
Fig. 2. Texture images from different datasets. (a) Oulu texture images (from left to right and top to bottom): Bark1, Flowers2, Flowers1, Metal, Leaves, Fabric7, Fabric2, Bark2, Tile (b) Kylberg texture images: sand1-a-p1, rice2-a-p1, rug1-a-p1, screen1-a-p1, scarf2-a-p1, scarf1-a-p1, stone1-a-p1, sesameseeds1-a-p1, seat1-a-p1, (c) VisTex texture images: Brick1, Bark9, Bark8, Fabric4, Fabric0, Buildings9, Fabric15, Fabric14, Fabric11 [7].

illumination, scale and perspective. The 0 degree orientation in all the twenty eight classes is used for the experimentation.

The nineteen gray scale texture classes from Vision Texture (VisTex) color dataset [24] are used for experimentation. The images consist of homogeneous frontal acquisition of textures. The images are acquired under uncontrolled illumination and viewing conditions.

The properties of the datasets are reported in the Table 2. The two fold experimentation is repeated five times for different training and testing sets.

Table 2. Basic properties of the datasets considered for experimentation.

Dataset name	Number of classes	Sub images for each class	Total sub images	Image size	Sub image size	Image format
Oulu [22]	16	16	256	512×512	128×128	ppm
Kylberg [23]	28	2560	71680	576×576	144×144	png
VisTex [24]	19	16	304	512×512	128×128	ppm

4.2 Experimental Results

The implementation of the proposed method is performed using MATLAB 7.9 software on Intel® Core™ i3-2330M @ 2.20 GHz with 4 GB RAM. The original input image is subjected to Haar wavelet transform. This results in diagonal (D), vertical (V), horizontal (H) and average (A) subband components. Texture approximation of the D, V and H components is obtained using anisotropic diffusion. Further, different statistical features (as listed in the Table 1) are computed from the texture approximations. The feature sets are FS1 to FS9 and combinations, namely, FS1 to FS9 and combinations, namely, FS1 + FS6 + FS3, FS1 + FS3 + FS5, FS1 + FS4 + FS3, FS1 +

FS4 + FS5, FS3 + FS4 + FS5, FS1 + FS3 + FS5 + FS4 + FS7 + FS6, FS1 + FS3 + FS4 + FS5 + FS6, FS4 + FS5, FS2 + FS4, FS1 + FS4 + FS3 + FS5 + FS8 + FS6 + FS7 + FS9, FS1 + FS4 + FS3 + FS6 + FS5 + FS8 + FS7, FS5 + FS6, FS1 + FS5 + FS3 + FS6, FS1 + FS4 + FS3 + FS5, FS6 + FS8 + FS7 + FS9, FS3 + FS5 + FS4 + FS6, FS1 + FS6 + FS5 + FS4, FS6 + FS7, FS3 + FS4, FS8 + FS9, FS1 + FS3 are considered. Totally, thirty feature sets are used for the experimentation. The LDA is employed to obtain discriminant features. The classification is performed using k-NN classifier. The parameters of anisotropic diffusion are optimized as: conduction coefficient = 60 and lambda = 0.25. The experimentation for each feature set is executed up to 10 diffusion steps. The performance of unbiased texture classification is analyzed by performing two-fold experimentation repeated five times. The results of the proposed method based on anisotropic diffusion are compared with the methods in [2, 7]. The results obtained with the proposed method on different datasets are discussed in the following sections.

Oulu dataset

For experimentation, 16 texture classes from Oulu dataset [22] in gray scale are considered. The Table 3 shows the feature sets that yield classification accuracy more than 95% on Oulu dataset and the corresponding optimal diffusion step (DS), testing time (TsTm) and training time (TrTm) of the proposed method. It is observed from the Table 3 that the fourteen combinations of feature sets give classification accuracy above 95%. The optimum classification accuracy of 98.36% is obtained for the combination of feature set FS1 + FS3 + FS5 + FS6 with testing time of 1.15 s. and training time of 18.36 s.

Table 3. The feature sets that yield classification accuracy more than 95% on Oulu dataset and the corresponding optimal diffusion step (DS), testing time (TsTm) and training time (TrTm) of the proposed method.

Sl. No.	Feature set	DS	CA (%)	TrTm (s)	TsTm (s)	Sl. No.	Feature set	DS	CA (%)	TrTm (s)	TsTm (s)
1	FS1 + FS3 + FS4	5	95.23	7.29	0.46	8	FS1 + FS3 + FS5	6	97.81	9.89	0.62
2	FS1 + FS3 + FS4 +FS5	**7**	96.88	13.10	0.82	9	FS1 + FS3 + FS6	7	97.58	7.30	0.46
3	FS1 + FS3 + FS4 +FS5 + FS6	8	97.03	14.63	0.92	10	FS1 + FS4 + FS5	8	97.19	13.31	0.84
4	FS1 + FS3 + FS4 +FS5 + FS6 + FS7	2	95.78	11.39	0.71	11	FS3 + FS4 + FS5	2	96.41	10.95	0.69
5	**FS1 + FS3 + FS5 + FS6**	**10**	**98.36**	**18.36**	**1.15**	12	FS6 + FS7	6	96.17	6.89	0.43
6	FS3 + FS4 + FS5 + FS6	6	97.42	12.97	0.82	13	FS4 + FS5	4	97.03	11.17	0.71
7	FS1 + FS4 + FS5 + FS6	6	97.42	12.85	0.81	14	FS5 + FS6	2	97.73	8.68	0.55

Kylberg dataset

The 28 classes of Kylberg [23] dataset are considered for the experimentation. The experimental setup is same as in Oulu dataset. The Table 4 shows result of the proposed method for various feature sets that yield classification accuracy more than 97%. It is been noticed from the Table 4 that the seventeen combinations of feature sets give

Table 4. The feature sets that yield classification accuracy more than 97% on Kylberg dataset and the corresponding optimal diffusion step (DS), testing time (TsTm) and training time (TrTm) of the proposed method.

Sl. No.	Feature set	DS	CA (%)	TrTm (s)	TsTm (s)	Sl. No.	Feature set	DS	CA (%)	TrTm (s)	TsTm (s)
1	FS7	3	97.32	1498.92	0.71	10	FS1 + FS3 + FS5	2	100	2746.66	1.48
2	FS1 + FS3	2	99.02	1203.55	1.09	11	**FS1 + FS3 + FS6**	**1**	**100**	**1434.71**	**1.65**
3	FS1 + FS3 + FS4	1	99.38	1952.36	1.59	12	FS1 + FS4 + FS5	1	99.64	3061.37	1.35
4	FS1 + FS3 + FS4 + FS5	1	99.55	3719.23	1.65	13	FS3 + FS4 + FS5	1	99.55	3103.00	1.31
5	FS1 + FS3 + FS4 + FS5 + FS6	1	99.55	3940.71	2.12	14	FS6 + FS7	2	100	1563.18	0.99
6	FS1 + FS3 + FS4 + FS5 + FS6 + FS7 + FS8	1	97.95	4632.30	3.01	15	FS3 + FS4	2	98.13	2055.68	0.98
7	FS1 + FS3 + FS5 + FS6	1	100	3019.48	1.97	16	FS5 + FS6	2	99.64	3143.12	0.89
8	FS3 + FS4 + FS5 + FS6	1	99.55	3596.66	1.94	17	FS2 + FS4	2	97.32	5423.65	1.88
9	FS1 + FS4 + FS5 + FS6	1	99.46	4266.68	1.62						

classification accuracy above 97%, thereby making the method more effective in texture classification. This is due to the large datasets considered for the experimentation. The variations in the training and testing time for each of the feature set is because of the features being obtained and the optimal diffusion step that gives better classification accuracy. The better classification accuracy of 100% is attained for the combination FS1 + FS3 + FS6 with testing time of 1.65 s. and training time of 1434.71 s.

VisTex dataset

For experimentation, 19 texture classes of VisTex dataset [24] in gray scale are considered. The Table 5 shows feature sets that yield classification accuracy (CA) more than 95% on VisTex database, along with the corresponding optimal number of diffusion steps (DS), testing time (TsTm) and training time (TrTm) of the proposed method. It is observed from the Table 5 that totally eleven feature sets give classification accuracy more than 95%. The optimum classification rate of 97.43% is obtained for combination FS1 + FS4 + FS5 + FS6 with testing time of 0.87 s. and training time of 16.45 s. The dominant feature sets FS1, FS3, FS4, FS5 and FS6 perform better in classification.

Table 5. The feature sets that yield classification accuracy more than 95% on VisTex dataset and the corresponding optimal diffusion step (DS), testing time (TsTm) and training time (TrTm) of the proposed method.

Sl. No.	Feature set	DS	CA (%)	TrTm (s)	TsTm (s)	Sl. No.	Feature set	DS	CA (%)	TrTm (s)	TsTm (s)
1	FS1 + FS3 + FS4 + FS5	8	96.51	15.80	0.83	7	**FS1 + FS4 + FS5 + FS6**	**8**	**97.43**	**16.45**	**0.87**
2	FS1 + FS3 + FS4 +FS5 + FS6	10	96.58	18.61	0.98	8	FS1 + FS3 + FS5	10	95.33	10.91	0.58
3	FS1 + FS3 + FS4 +FS5 + FS6 + FS7	10	95.46	19.42	1.03	9	FS1 + FS3 + FS6	10	95.33	10.91	0.58
4	FS1 + FS3 + FS4 + FS5 +FS6 + FS7 + FS8	8	95.39	18.22	0.96	10	FS1 + FS4 + FS5	8	96.38	15.51	0.82
5	FS1 + FS3 + FS5 + FS6	8	96.51	14.38	0.76	11	FS3 + FS4 + FS5	8	95.72	15.60	0.83
6	FS3 + FS4 + FS5 + FS6	8	96.91	16.58	0.88						

Table 6. Comparison of classification accuracy, testing and training time of the proposed method for the three datasets and that of the other methods in the literature [2, 7].

Dataset	Oulu			Kylberg			VisTex		
	Proposed method (WT + PDE)	Method in [2] (NSCT + LDBP)	Method in [7] (PDE + LDBP)	Proposed method (WT + PDE)	Method in [2] (NSCT + LDBP)	Method in [7] (PDE + LDBP)	Proposed method (WT + PDE)	Method in [2] (NSCT + LDBP)	Method in [7] (PDE + LDBP)
Classification accuracy (%)	98.36	99.68	99.84	100	83.42	100	97.43	96.38	98.55
Training time (s)	18.36	291.48	135.97	1434.71	48771.46	7575.67	16.45	343.05	69.05
Testing time (s)	1.15	10.06	8.50	1.65	1930.46	2.77	0.87	11.33	3.63
Diffusion step	10	–	200	1	–	15	8	–	65

The Table 6 shows the comparison of classification accuracy, testing time and training time for the datasets for the proposed method and the methods in the literature [2, 7]. The effectiveness of the proposed method is observed from the Table 6 in terms of increased classification accuracy as compared to the other methods in the literature [2, 7]. It is also observed from the Table 6 that the proposed method shows the computational efficiency. The average reduction of 93.6% and 95.32% of testing time and training time is observed as compared to the method in [2] for VisTex, Kylberg and Oulu datasets. The average reduction of 81.25% and 67.65% of training time and testing time is observed as compared to the method in [7] for VisTex, Kylberg and Oulu datasets. This can be attributed to the features obtained from the texture component of anisotropic diffusion of input texture classes.

5   Conclusion

The paper presents an efficient method to obtain discriminant features for texture classification using WT and PDE approach. The wavelet transform is used to obtain the directional information of the input image using one level decomposition. Anisotropic diffusion is used to extract texture component from wavelet decomposed images. Statistical features are obtained from texture component. The discriminant features are obtained using LDA. The robustness of the proposed method is tested on VisTex, Kylberg and Oulu datasets. The k-NN classifier with two fold validation repeated five times is used to classify images. The proposed method performs better than the other methods in the literature. For texture image classification, the optimal values of diffusion steps are ten, one and eight for Oulu, Kylberg and VisTex datasets, respectively. The experimental evaluation reveals that the feature sets taken in combination performs better than considering the feature set individually. The enhanced classification accuracy is exhibited for the feature sets FS1, FS3, FS4, FS5 and FS6, taken in combination for all the datasets. The advantages of the proposed method are increased classification accuracy and reduced computational time. Therefore the proposed method is suitable for real time applications. The novelty of the proposed method arises in terms of capturing the discriminant texture features for the texture components obtained using anisotropic diffusion for directional subbands. The results obtained demonstrate the computational efficiency of the proposed method.

References

1. Hiremath, P.S., Shivashankar, S.: Wavelet based co-occurrence histogram features for texture classification with an application to script identification in a document image. Pattern Recogn. Lett. **29**, 1182–1189 (2008)

2. Hiremath, P.S., Bhusnurmath, R.A.: Texture image classification using nonsubsampled contourlet transform and local directional binary patterns. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **3**(7), 819–827 (2013)

3. Hiremath, P.S., Bhusnurmath, R.A.: Multiresolution LDBP descriptors for texture classification using anisotropic diffusion with an application to wood texture analysis. Pattern Recogn. Lett. **89**, 8–17 (2017). https://doi.org/10.1016/j.patrec.2017.01.015

4. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. **12**(12), 629–639 (1990)

5. Hiremath, P.S., Bhusnurmath, R.A.: RGB - based color texture image classification using anisotropic diffusion and LDBP. In: Murty, M.N., He, X., Chillarige, R.R., Weng, P. (eds.) MIWAI 2014. LNCS (LNAI), vol. 8875, pp. 101–111. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-13365-2_10

6. Bai, J., Miri, M.S., Liu, Y., Saha, P., Garvin, M., Wu, X.: Graph-based optimal multi-surface segmentation with a star-shaped prior: application to the segmentation of the optic disc and cup. In: IEEE International Symposium on Biomedical Imaging, pp. 525–528 (2014)

7. Hiremath, P.S., Bhusnurmath, R.A.: Diffusion approach for texture analysis based on LDBP. Int. J. Comput. Eng. Appl. Part I **9**(7), 108–121 (2015)

8. Weszka, J.S., Dyer, C.R., Rosenfeld, A.: A comparative study of texture measures for terrain classification. IEEE Trans. Syst. Man Cybern. **6**(4), 269–285 (1976)

9. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **3**(6), 610–621 (1973)

10. Amadasun, M., King, R.: Texural features corresponding to texural properties. IEEE Trans. Syst. Man Cybern. **19**(5), 1264–1274 (1989)

11. Wu, C.-M., Chen, Y.-C.: Statistical feature matrix for texture analysis. CVGIP: Graph. Models Image Process. **54**(5), 407–419 (1992)

12. Laws, K.I.: Rapid texture identification. SPIE **238**, 376–380 (1980)

13. Haralick, R.M., Shapiro, L.G.: Computer and Robot Vision, vol. 1. Addison-Wesley, Boston (1992)

14. Mandelbrot, B.B.: The Fractal Geometry of Nature. Freeman, San Francisco (1982)

15. Rosenfeld, A., Weszka, J.: Picture recognition. In: Fu, K. (ed.) Digital Pattern Recognition, pp. 135–166. Springer, Berlin (1980). https://doi.org/10.1007/978-3-642-67740-3_5

16. Paci, M., Nanni, L., Lathi, A., Setala, A.K., Hyttinen, J., Severi, S.: Non binary coding for texture descriptors in sub-cellular and stem cell image classification. Curr. Bioinform. **8**, 208–219 (2013). https://doi.org/10.2174/1574893611308020009

17. Hiremath, P.S., Bhusnurmath, R.A.: Colour texture classification using anisotropic diffusion and wavelet transform. In: Proceedings of the Seventh International Conference on Advances in Communication, Network, and Computing - CNC 2016, De Gruyter, pp. 44–61 (2016)

18. Hiremath, P.S., Bhusnurmath, R.A.: Performance analysis of anisotropic diffusion based colour texture descriptors in industrial applications. Int. J. Comput. Vis. Image Process. **7**(2), 50–63 (2017). https://doi.org/10.4018/IJCVIP.2017040104

19. Daubechies, I.: Ten Lectures on Wavelets. SIAM, Philadelphia (1992)
20. Mallat, S.G.: A theory of multiresolution signal decomposition: the wavelet representation. IEEE Trans. Pattern Anal. Mach. Intell. **11**, 674–693 (1989)
21. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification. Wiley Publication, New York (2001)
22. Ojala, T., Maenpaa, T., Pietikainen, M., Viertola, J., Kyllonen, J., Huovinen, S.: Outex - new framework for evaluation of texture analysis algorithms. In: Proceedings of International Conference on Pattern Recognition, pp. 701–706 (2002). http://www.outex.oulu.fi/
23. Kylberg, G.: Kylberg texture dataset v. 1.0 (2012). http://www.cb.uu.se/∼gustaf/texture/
24. http://vismod.media.mit.edu/vismod/imagery/VisionTexture
25. Pan, Z., Fan, H., Li, Z.: Texture classification using local pattern based on vector quantization. IEEE Trans. Image Process. **24**(12), 5379–5388 (2015)
26. Mehta, R., Egiazarian, K.: Texture classification using dense micro-block difference. IEEE Trans. Image Process. **25**(4), 1604–1616 (2016)
27. Dash, J.K., Mukhopadhyay, S., Gupta, R.D.: Multiple classifier system using classification confidence for texture classification. Multimeda Tools Appl. **76**(2), 2535–2556 (2017). https://doi.org/10.1007/s11042-015-3231-z
28. Liu, L., Fieguth, P., Guo, Y., Wang, X., Pietikäinen, M.: Local binary features for texture classification: taxonomy and experimental study. Pattern Recogn. **62**, 135–160 (2017). https://doi.org/10.1016/j.patcog.2016.08.032

Position Error Analysis of IRNSS Data Using Big Data Analytics

M. Geetha Priya[(✉)] and D. C. Kiran Kumar

Centre for Incubation, Innovation, Research and Consultancy,
Jyothy Institute of Technology, Bangalore, Karnataka, India
geetha.sri82@gmail.com, kiran.dc88@gmail.com

Abstract. The current investigation involves the study and analysis of position error of IRNSS (Indian Regional Navigation Satellite System) receiver system data with respect to a fixed position using big data analytics tool namely, Hadoop. IRNSS is an indigenously developed navigation system by India with constellation of seven satellites. IRNSS data pertaining to week number 918 has been used to carry out the statistical analysis in three different modes of operation. Queries/commands were created in Hadoop to process more than 6 million data (week no. 918 data). The processed data and calculated statistical parameters show that IRNSS navigation system is a promising solution for various navigation related applications. The results obtained highlights that the position error is minimum in Hybrid (GPS-L1 + IRNSS-S + IRNSS-L5) mode of operation with better accuracy when compared with individual frequency mode of operation (L5 & S) of IRNSS receiver system.

Keywords: GNSS · IRNSS · Satellites · Position error · Hadoop

1 Introduction

Global Navigation Satellite System (GNSS) is used to provide autonomous geo-spatial positioning in terms of longitude and latitude using the constellation of satellites. GPS (Global positioning System) is a one such type of GNSS with network of satellites used to provide the precise information about position that is widely being used by many users all over the world. The GNSS signals are received by the receivers to provide exact location and timing information. IRNSS is India's own regional navigation system [1] renamed as NAVIC (Navigation with Indian constellation). The NAVIC system is used for facilitating position, navigation, and tracking and also to calculate the time with high accuracy. IRNSS receiver system is designed to deliver position/location information to Indian users and the coverage area extends up to one thousand five hundred (1500) km [2] from the boundary of India, which is its primary service area. Figure 1 shows IRNSS receiver system, antenna and sky plot of available satellites. The IRNSS System is providing a position accuracy of 20 m in principal service area. IRNSS provides two types of services namely RS (Restricted Service) that is encrypted & available only for authorized users with high precision and Standard Positioning Service (SPS) to all users. Presently there are seven satellites in the constellation [3–5] placed above 36,000 km from the surface of earth. Three satellites are

placed in geostationary orbit at 32.5° East, 83° East, and 131.5° East longitude with fixed position. The other four satellites are in geosynchronous orbit, where each set of satellites will cross the equator at 55° and 111.75° East.
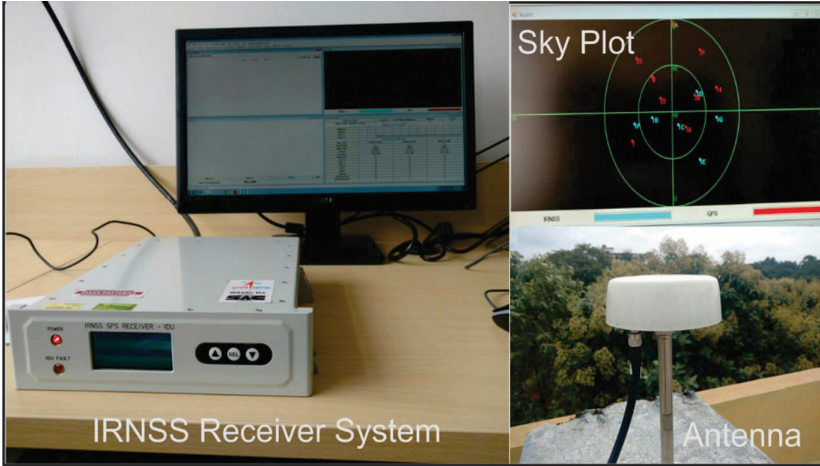
Fig. 1. IRNSS receiver system, antenna and sky plot

The IRNSS satellite signals are available at L5 band (1176.45×10^6 Hz) and S band (2492.028×10^6 Hz) microwave frequencies. The SPS (Standard positioning Service) signal is modulated by 1×10^6 Hz BPSK (Binary Phase shift keying) [6, 7] signal and RS (Restricted Service) signal is modulated with BOC (Binary offset carrier) for high precision and security. To maintain necessary coverage area and signal strength, the signal is being transmitted through phased array antenna.

A Java written open source framework (Apache) big data analytics tool-Hadoop is used for processing of large datasets in a distributed way across network of computers using robust programming models. Hadoop manages applications with MapReduce algorithm by processing the data on different CPU nodes in parallel. In simple, Hadoop framework allows the users to develop applications that are compatible to run on clusters of computers and also it can achieve complete statistical study for a large quantity of data. In this study, Hadoop frame work is used to analyze the IRNSS data statistically to assess the position error. This paper has been organized with method-ology & mathematical modeling in Sect. 2 with simulation results and conclusion in Sects. 3 and 4 respectively.

2   Methodology

2.1   IRNSS Data and HDFS

The data from the IRNSS receiver installed at the site with respect to a fixed position (Latitude = 12.77514° N and Longitude = 77.48794° E with X coordinate = 1345092.73, Y coordinate = 6073169.66, Z coordinate = 1408458.60) is utilized to carry out the

statistical analysis using Hadoop framework. The IRNSS data pertaining to week No 918 (27-3-2017 to 1-4-2017) received with the installed receiver is the data of interest for this analysis. The data received from the IRNSS SPS RX is loaded to the HDFS (Hadoop Distributed File System) Environment using Sqoop. Sqoop is a tool designed to transfer data between relational database servers and Hadoop. Hadoop is supported with tools like Sqoop and Hive in a Linux platform. Hadoop uses Hive to process data which is a data warehouse structure tool. Figure 2 gives the software framework with architecture storage and functional flow of Hadoop.
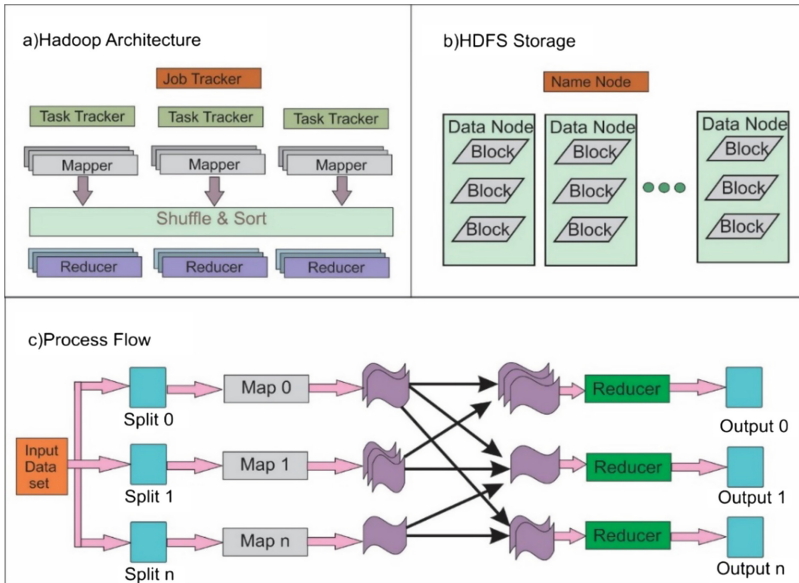

Fig. 2. Software framework (a) Hadoop architecture (b) HDFS storage (c) process flow

In general, the IRNSS data consists of files related to 3 solutions namely sol_1 (Mode 1), sol_2 (Mode 2), sol_3 (Mode 3) along with satellite information, PR (pseudo random) log & iono tropo. For week no 918, Sol_1 was assigned as hybrid mode, which is the combination of GPS (L1) + IRNSS (L5 + S), Sol_2 was assigned as IRNSS S Band and Sol_3 was assigned as IRNSS L5 Band. The three solutions (modes) consists of parameters like week no, user time, system time, solution type, iono type, X coordinates, Y coordinates, Z coordinates, latitude, longitude, altitude, geometric dilution of precision (GDOP), position dilution of precision (PDOP), vertical dilution of precision (VDOP), horizontal dilution of precision (HDOP), time dilution of precision (TDOP), position error, number of satellites, velocity information, etc. Satellite information data (file) consists of week no, user time, system time, pseudo random noise (PRN), elevation, azimuth, carrier to noise density ratio (CNO) and X, Y, Z coordinates of all the satellites. PR log data (file) consists of week no, user time, system time, Doppler Frequency (FDOPP) and Iono tropo data (file) consists of week no, user time, system time, iono, tropo information.

2.2   Mathematical Modeling for Statistical Analysis Using Hadoop

The Hive structural table created for sol_1 (Mode 1), sol_2 (Mode 2), sol_3 (Mode 3), satellite information, PR log & iono tropo using Hive queries can be used for all weeks of IRNSS data. The data is loaded to the hive structured table and week number 918 data is selected by removing all the invalid data from the other sources. The process flow for the IRNSS data analysis using Hadoop is shown in Fig. 3.
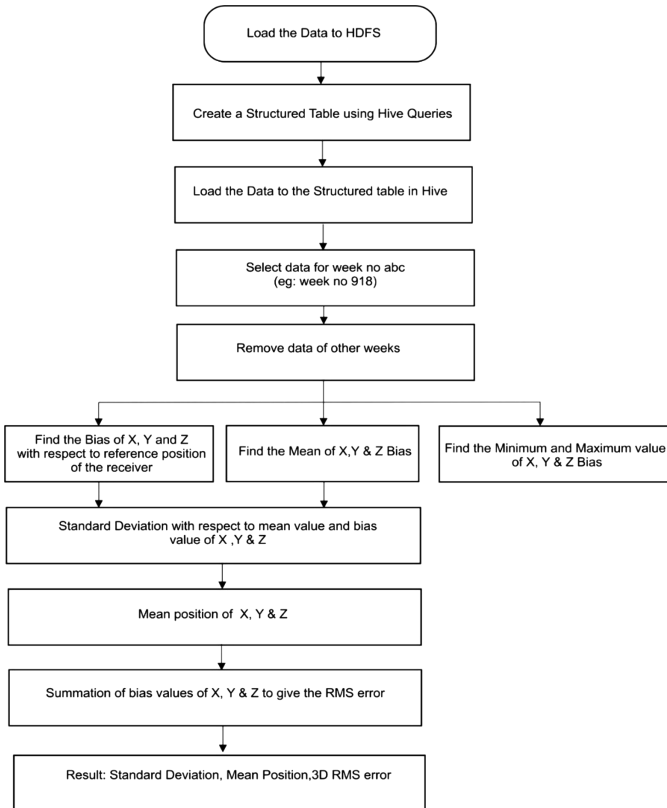
Fig. 3. Process flow

The data processed by Hadoop gives the minimum and maximum values of X, Y & Z coordinates, standard deviation, position errors, RMS error and mean position of X, Y & Z. The 3D RMS position is found by taking the mean square error with the help of standard reference values that was calculated by averaging X, Y and Z coordinates for a specific period of time. The ECEF (Earth center earth fixed) values are taken to calculate the mean X, Y, Z positions by using hive queries.

Hadoop uses Hive to process data which is a data warehouse structure tool. Hive exists on top of Hadoop to querying and analyzing Big Data and this makes querying/analyzing much easier for the user. For the present study, Hive queries are

used for measuring the minimum, maximum, standard deviation, RMS error and mean position error of the X, Y & Z coordinates (of IRNSS data week no. 918) to find the positional information of the receiver. A sample of queries used for this study is given in Table 1.

Table 1. Sample Queries

Data/file	Queries
Sol_1:	CREATE EXTERNAL TABLE Sol_1(Col_1 datatype, Col_2 datatype ….)
Sol_2:	CREATE EXTERNAL TABLE Sol_2(Col_1 datatype, Col_2 datatype ….)
Sol_3:	CREATE EXTERNAL TABLE Sol_3(Col_1 datatype, Col_2 datatype ….)
Satellite info:	CREATE EXTERNAL TABLE Satellite_info(Col_1 datatype, Col_2 datatype …)
Iono tropo:	CREATE EXTERNAL TABLE iono_tropo(Col_1 datatype, Col_2 datatype …)
PR log:	CREATE EXTERNAL TABLE pr_log(Col_1 datatype, Col_2 datatype …)
Commands:	
Select the data of week no 918:	select * from sol_1 where week no = 918
Select only the valid data	select * from sol_1 where ux > 0.0
Bias_X, Bias_Y, Bias_Z:	select ux, uy, uz, round((ux - 1345092.73), 2) as bias_x, round((uy - 6073169.66), 2) as bias_y, round((uz - 1408458.60), 2) as bias_z from sol_1_f
Mean_X, Mean_Y, Mean_Z:	select avg(bias_x) as mean_x, avg(bias_y) as mean_y, avg(bias_z) as mean_z from sol_1_f_xyz
Min & Max of X, Y & Z bias:	select min(bias_x) as min_x, max(bias_x) as max_x, min(bias_y) as min_y, max (bias_y) as max_y, min(bias_z) as min_z, max(bias_z) as max_z from sol_1_bias
Standard deviation:	select stddev_pop(bias_x) as std_x, stddev_pop(bias_y) as std_y, stddev_pop (bias_z) as std_z from sol_1_bias
ux, uy & uz pos:	select ux, uy, uz from sol_1_f
Standard deviation of pos:	select stddev_pop(ux) as std_x_pos, stddev_pop(uy) as std_y_pos, stddev_pop (uz) as std_z_pos from sol_11
Mean_X, Mean_Y, Mean_Zpos:	select avg(ux) as mean_x_pos, avg(uy) as mean_y_pos, avg(uz) as mean_z_pos from sol_11
Sum of bias_X, bias_Y, bias_Z:	select sum(bias_x) as sum_x, sum(bias_y) as sum_y, sum(bias_z) as sum_z from sol_1_bias

3   Results and Discussion

In this study, the position error of IRNSS receiver data is analyzed with respect to a fixed position (user location latitude = 12.77514° N and longitude = 77.48794° E with X coordinate = 1345092.73, Y coordinate = 6073169.66, Z coordinate = 1408458.60). This study is an attempt to utilize open source tools like Hadoop to carry out the statistical analysis of position error from the IRNSS data. Figures 4 and 5 gives the position error plot for full week no. 918 & for 8 h of a specific day in all three modes respectively. Using the queries and commands shown in Table 1, the minimum,

maximum, standard deviation, mean and other statistical parameters required for the analysis are obtained using Hadoop tool with respect to position error. The statistical parameter values obtained are listed in Table 2 for all modes.
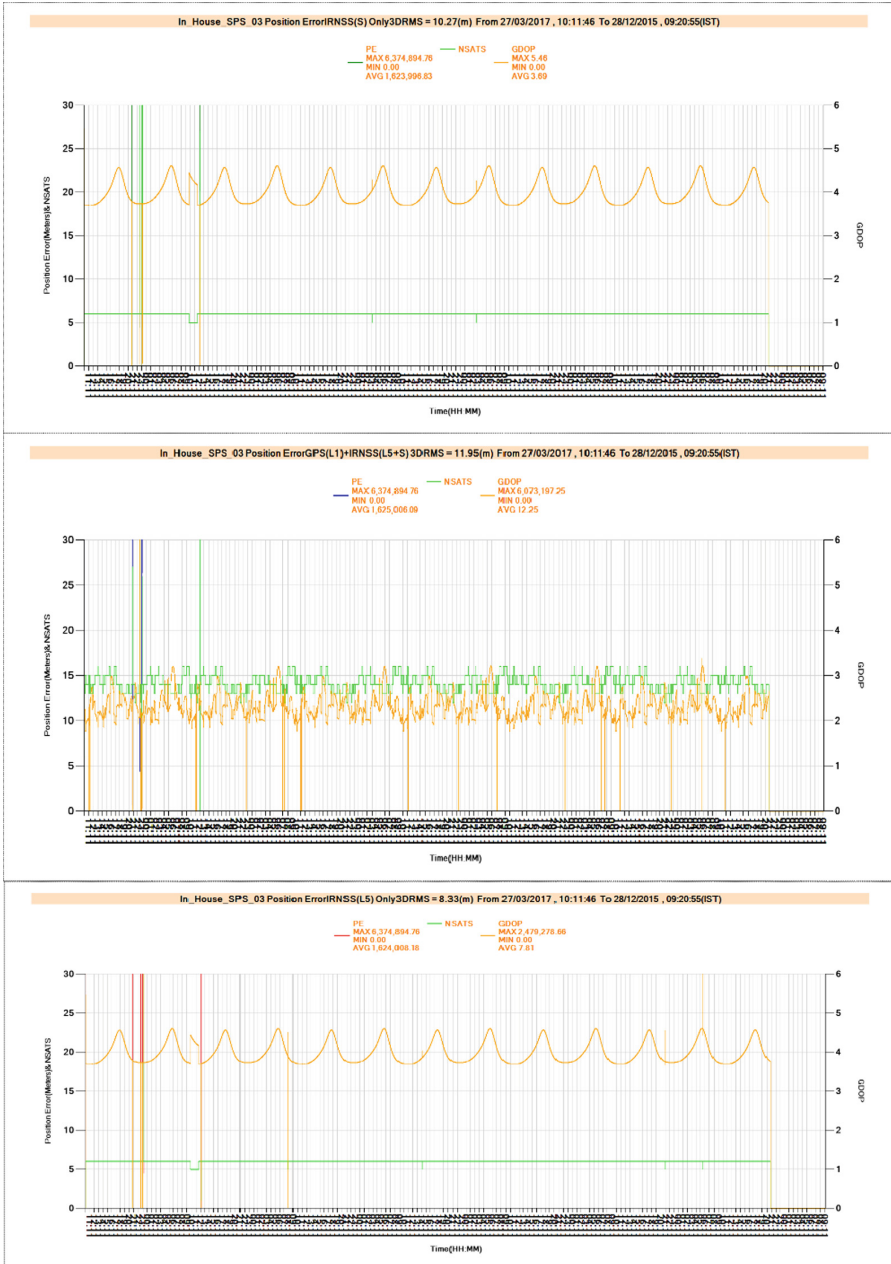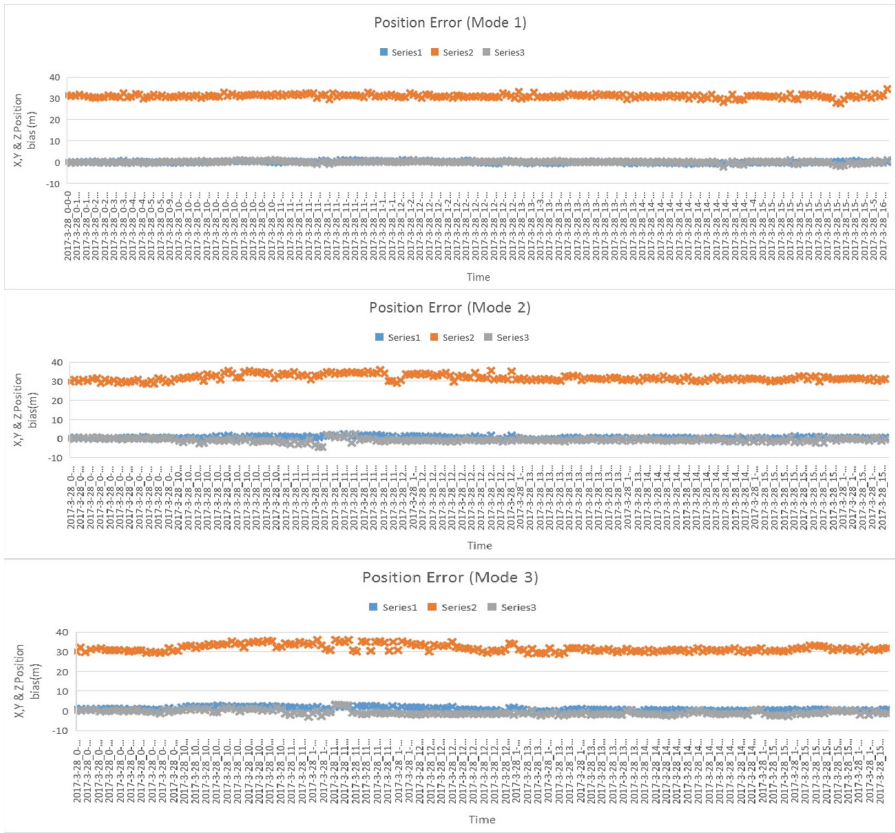

Fig. 4. Position error plot for week no. 918 for all three modes.

Fig. 5. Position error plot for 8 h (12 pm to 8 pm, 28-3-2017 of week no. 918) of data.

From Table 2, it is observed that,

- Rows 3, 5 & 7 shows that IRNSS-S band is giving better results in terms of maximum position error (X, Y & Z) compared to Hybrid and IRNSS-L5 band.
- From rows 4 & 6, it is observed that hybrid mode is better in giving minimum position error for X & Y coordinates compared to IRNSS-S & L5 band. Row 8 shows that IRNSS-S band gives better in terms of minimum position error for Z coordinates in comparison with Hybrid and IRNSS-L5 band.
- From row 12, 13 & 14, it is seen that the standard deviation (X, Y & Z error) with respect to reference value is minimum in hybrid mode in comparison with IRNSS-S & L5 band. The second best solution is IRNSS-S band.
- From rows 9, 10 & 11, it is observed that IRNSS-L5 band is giving minimum average position error when compared to Hybrid and IRNSS-S band.
- From row 15 & 18, it is seen than the difference between mean X position and reference X position is minimum in hybrid solution compared to IRNSS-S & L5 band.

- From row 16 & 19, it is seen than the difference between mean Y position and reference Y position is minimum in hybrid solution compared to IRNSS-S & L5 band.
- From row 17 & 20, it is seen than the difference between mean Z position and reference Z position is minimum in IRNSS-S band solution compared to hybrid & IRNSS-L5 band.

The proposed study shows that hybrid mode of operation is giving the best position information with minimum position error and maximum accuracy than the individual L5 or S band of IRNSS receiver system. The second best solution is given by IRNSS-L5 band.

Table 2. Statistical analysis of GPS & IRNSS data using Hadoop framework for week no. 918

Parameters	Hybrid: GPS-L1 + IRNSS-S + IRNSS-L5 (m)	IRNSS-S (m)	IRNSS-L5 (m)
3D RMS position error	3.24727	3.23492	3.235524
Max X error	46.23	2.63	4.88
Min X error	−16.91	−8.63	−9
Max Y error	87.08	39.73	47.83
Min Y error	10.95	20.72	19.81
Max Z error	6.76	6.05	7.72
Min Z error	−20.01	−103.35	−102.7
Mean X error	4.97	−5.06	−7.28
Mean Y error	0.0029	7.54	−4.02
Mean Z error	−8.12	−5.55	−8.59
SD X error	0.35	0.57	0.84
SD Y error	0.96	1.41	1.55
SD Z error	0.57	1.42	2.02
Mean X position	1345097.70	1345087.66	1345085.44
Mean Y position	6073169.66	6073177.20	6073165.63
Mean Z position	1408450.47	1408453.04	1408450.00
Reference X position	1345092.73		
Reference Y position	6073169.66		
Reference Z position	1408458.60		

4   Conclusion

This study is an attempt to utilize open source tools like Hadoop to carry out the statistical analysis of position error from the IRNSS data. The Data is collected in three different modes like Hybrid (GPS + IRNSS (L5 + S)), IRNSS(S) and IRNSS (L5) and

processed individually to measure the effectiveness of IRNSS. This study cum analysis shows that, there is an improvement in accuracy and visibility with respect to hybrid mode of operation with IRNSS L5 band giving the second best solution. The position error analysis computed using Hadoop for IRNSS receiver, is essential for understanding how IRNSS works, and for knowing what degree of errors could be anticipated. This helps in making corrections for receiver clock errors and other effects. This kind of analysis could be carried out for any GNSS receivers. This paper provides an overview for researchers in GNSS field to choose the appropriate solution depending on the applications.

References

1. Montenbruck, O., Steigenberger, P.: IRNSS orbit determination and broadcast ephemeris assessment. In: ION International Technical Meeting (ITM), Dana Point, CA, pp. 26–28 (2015)
2. Raghu, N., Manjunatha, K.N., Kiran, B.: Determination and preliminary analysis of position accuracy on IRNSS satellites. In: International Conference on Communication and Signal Processing, pp. 6–8. IEEE Press, India (2016). https://doi.org/10.1109/iccsp.2016.7754248
3. Khatri, R.R., Chauhan, S.: Indian regional navigation satellite system. Int. J. Innov. Res. Technol. **2**, 380–384 (2016)
4. Indian Space Research Organization. https://www.isro.gov.in/irnss-programme
5. Raghu, N., Kiran, B., Manjunatha, K.N.: Tracking of IRNSS, GPS and hybrid satellites by using IRNSS receiver in STK simulation. In: International Conference on Communication and Signal Processing, India, pp. 6–8 (2016). https://doi.org/10.1109/iccsp.2016.7754276
6. Guangya, X.W., Zhiqiang, S., Xiaofeng, L., Yang, H.J.: Research on the 3D visualization of information operations based on STK. In: International Conference on Audio Language and Image Processing, pp. 921–925. IEEE Press, China (2010). https://doi.org/10.1109/icalip.2010.5685180
7. Fan, S., Zhao, L., Xiao, W., Li, Z.: Performance analysis and simulation of iridium navigation satellite based on STK. In: 2nd International Workshop on Earth Observation and Remote Sensing Applications, China, pp. 291–295 (2012). https://doi.org/10.1109/eorsa.2012.6261185

Cluster Representation and Discrimination Based on Regression Line

M. S. Bhargavi[1(✉)] and Sahana D. Gowda[2]

[1] BIT, Bangalore, India
ms.bhargavi@gmail.com
[2] BNMIT, Bangalore, India
sahanagowdad@gmail.com

Abstract. Clustering aims to group data into coherent groups based on the nearness of samples in multiple feature space where the coherency enriches the uniqueness of the clusters with respect to others. A cluster representation based on intrinsic cohesiveness would aid in efficient discrimination between clusters. In this paper, a novel method for representation and discrimination amongst the clusters has been proposed. Distances computed between every pair of samples in a cluster reveal the cohesiveness of samples in multi-dimensional space. As distances computed between every pair of samples enormously increase with the number of samples, distances are assimilated by histograms. The range of the bins in a histogram specifies the distance amongst the samples in a cluster. For effective discrimination, histograms are further transformed into a regression line by constructing cumulative histograms. Each cluster is represented by slope, intercept and error characterizing the regression line. The extent and angle of the slope is determined by the diameter of the cluster ranged by the bins and distribution of distances in the bins. To discriminate clusters represented by regression line, a statistical test called probability-value hypothesis testing is performed. Based on the probability obtained, the clusters are discriminated to be similar or dissimilar. Experimentation on real and synthetic clusters demonstrates the efficiency of the proposed approach in extracting unique cluster representation for discrimination.

Keywords: Cluster cohesion · Representation · Discrimination
Histogram · Regression

1 Introduction

Clustering, an unsupervised learning approach clusters the data based on the coherency of samples in feature space [1]. Each cluster formulated is unique with respect to the other. Discrimination of clusters is important in various applications such as anomaly detection [2, 3], knowledge updation in incremental learning [4, 5], comparing the results of clustering algorithms [6, 7] and others. For faster comparison and efficient storage, it is essential to maintain cluster representations rather than raw clustered samples. Cluster representation must be so unique that the representation must be able to depict the difference between the clusters quantitatively.

Samples in any cluster have unique cohesiveness based on which the samples are grouped homogenously. The cohesion is due to the relative closeness of samples in multiple feature space which is elucidated by distances computed between every pair of samples. Large number of distances computed between every pair of samples in a cluster can be assimilated by histograms [8]. Histogram assimilates the distances into the bins where each bin elucidates the frequency of distance values in the designated bin encapsulating certain range of distance value. The bins ranges between the smallest and largest distance computed between any pair of samples with the largest distance being the diameter of the cluster [9].

To provide effective discrimination between clusters, histograms are further transformed into regression line. A histogram is transmuted to a regression line through an intermediate step involving the computation of cumulative histogram. Each regression line of the form $y = mx + c + e$, abridges one entire histogram into slope, intercept and error. The extent and angle of the slope is determined by the diameter of the cluster ranged by the bins and the distribution of distances in the bins. A cluster with different distribution of distances or diameter will result in a different slope. Hence, this proposal provides a means for effective discrimination between clusters.

Regression line representations of respective clusters are statistically compared through Probability-value (p-value) hypothesis testing. The p-value is the probability of finding observed or more extreme results of a statistical model when the null hypothesis is true [10]. On comparison, p-value predicts the probability between 0-1 for slopes and intercepts. Based on the ranging between 0 and 1, the clusters can be discriminated to be similar or non-similar.

In this paper, a novel cluster representation is proposed using regression line based on assimilation of distances computed amongst the samples of a cluster in histogram. Each cluster is represented by slope, intercept and error characterizing the regression line. A transformation of clusters to regression line representation aids in faster and efficient discrimination of clusters.

2   State-of-Art

Researchers have proposed two different approaches to provide discrimination amongst the clusters. In the first approach, clusters are directly discriminated without representation [11–14]. In the second approach, mostly used in streaming scenarios, clusters are represented and discriminated [15–19].

Permutational multivariate analysis of variance [11], Multi-response permutation procedure [12], Mantel's test [13] and Analysis of group similarities [14] are non-parametric tests to discriminate between the clusters without representing the clusters.

Permutational multivariate analysis of variance [11] computes the ratio 'F' of average distance among the clusters to average distance within the cluster. Larger the ratio more pronounced is the separation. Several permutations of the samples are performed for gaining significance. It is generally accepted that any separation between clusters is not significant if more than 5% of the permuted F-statistics have values greater than that of the observed statistic.

Multi-response permutation procedure [12] computes the average distance within each cluster. Delta, the weighted mean within-cluster distance is computed for k clusters. The weight assigned is the ratio of number of samples within a cluster to all the other clusters. Samples and distances computed among them are permuted several times to recalculate the delta. Fraction of deltas obtained on permutation that are less than the observed ones is the significance test.

Mantel's test and its extensions [13] is a statistic test for finding correlation between proximity matrices that can be used to test differences between clusters. In any one of the proximity matrix, the rows and columns are permuted and correlation recalculated. A deviation from zero correlation indicates significant difference over several permutations.

Analysis of group similarities [14] is an extension to the Mantel Test. It is based on the standardized rank correlation of proximity matrices. The ratio 'R' is calculated which is the ratio of difference in among-and-within cluster ranks to the number of sample pairs used to compute distance. The R statistic is compared to a set of R' values achieved by means of randomly shuffling cluster labels between the samples and the percent of times the actual R surpassed the permutations derived R' values is used to compute difference.

These methods require original samples of the clusters to be made available. They also undergo many permutations of samples to depict the amount of discrimination thus increasing the complexities of analysis. This kind of analysis may not be suitable in scenario where there is real-time requirement to reveal the difference between the clusters.

The approach where the clusters are represented before discrimination is seen in scenarios where clusters are obtained from streaming data. Clusters are represented and compared temporally.

Aggarwal et al. [15] proposed CluStream algorithm which maintains a representation of clusters called micro-cluster of tuple list – ['N', 'LS', 'SS', 'LST', 'SST'] where N is the number of samples, LS is the linear sum of samples, SS is the squared sum of the samples, LST is the sum of time stamps, SST is the sum of the squares of the time stamp. Clusters are compared by measuring the closeness in terms of the representative parameters of the cluster.

Yang and Zhou [16] proposed Hclustream in which clusters are represented using the histogram and centroid of discrete and continuous features respectively termed as micro-Hclusters. To compare the clusters, a heterogeneous distance measure is formulated to compute distance between sample-sample, sample-microcluster and microcluster-microcluster.

Komkrit et al. [17] proposed E-Stream algorithm where the cluster representation is named as Fading cluster structure as it reduces the weight of older data. Each cluster maintains a vector of weighted summation of the values of features, weighted sum of squares of feature values, sum of the weights in the samples and α-bin histogram computed over each feature projecting the distribution. Fading model reduce weight of older data in order to explore change in cluster structure.

Meesuksabai et al. [18] proposed HUE-Stream which is an extension to E-Stream algorithm. It provides support for both categorical and numerical features along with uncertainty in data. Clusters are represented similar to E-stream but extended to support

uncertainty through a distance function with probability distribution. Change in the clustering structure is detected through the distance function.

Nagabhushan et al. [19] proposed a representation for clusters by maintaining triple information for each feature of a cluster – regression line transformed from histogram, standard deviation and mean. This representation is used for maintaining knowledge of clusters and merging nearest clusters based on some distance threshold during incremental learning.

The representations provided are feature based representations, where the representations must be scaled to varying number of features of different stream. Comparison of clusters become computationally complex with increasing number of features and may not be feasible in real-time environment.

Discrimination between the clusters with original samples is tedious and time consuming. Hence, cluster representation before discrimination is a better alternative. However, the representation must not be based on individual features in order to maintain uniform complexity of analysis between low dimensional and high dimensional data. This work proposes a novel cluster representation and discrimination based on regression line. This approach of cluster representation aids in faster and efficient discrimination between clusters.

3   Proposed Methodology

3.1   Cluster Representation

A cluster is a set of samples with high coherency within clusters and low coherency between clusters. Cohesion determines how closely related the samples of a cluster are. Depending on the number of samples and the cohesiveness of samples within a cluster, clusters may have different geometry/shape, size, diameter and density. Distances computed between every pair of samples in a cluster aids in elucidating the pattern of cohesion of samples in a cluster. Clusters are represented based on distances computed amongst the samples in a cluster.

Let 'k' be the number of clusters formulated. Let $n_1, n_2.....n_i$, be the number of samples in cluster $k_1, k_2....k_i$ respectively. For each cluster, pairwise Euclidean distance between the samples is computed. Say, the data samples in cluster k_i consist of m_j features. Given an m_j x n_i matrix of data samples in cluster k_i, the distance amongst the observations X_1 and X_2 in Euclidean space is given by [20]

$$D_{12} = \sqrt{\sum_{j=1}^{m_j} \left(X_{1j} - X_{2j} \right)^2} \tag{1}$$

Euclidean distance computed between every pair of samples of a cluster is assimilated by histograms. Before explaining the process of assimilating the distances by histogram, a brief introduction to histogram is provided in the next section.

Histogram. The histogram [8] represents the distribution of specified values over a range of bins. The height of each bin is the count of values in the designated range (Fig. 1).
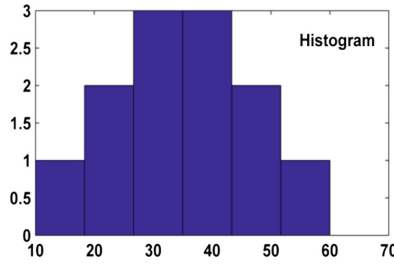
Fig. 1. A sample histogram

Pairwise distances are assimilated by histogram into bins. Bins ranges between the smallest and the largest distance computed between any pair, the latter being the diameter of the cluster. Each bin encompasses the frequency of the distances in the designated range.

Histogram is further transformed into a regression line representation through an intermediate step involving the computation of cumulative histograms [21, 22]. A transformation of histogram to regression line further assimilates the distribution and range of distances in a histogram into slope, intercept and error values. The process of transformation of histogram to regression line is explained in the next section.

Transformation of Histogram to Regression Line. Let a histogram H consists of 10 bins,

$H \rightarrow \{a_1\ a_2\ a_3\ a_4\ a_5\ a_6\ a_7\ a_8\ a_9\ a_{10}\}$
where, a_i is the frequency counts of the bins.

A cumulative histogram is formulated over these 10 peak indices of the histogram as shown Fig. 2a.

$CH \rightarrow \{c_1\ c_2\ c_3\ c_4\ c_5\ c_6\ c_7\ c_8\ c_9\ c_{10}\}$
where $c_i = \sum a_k$ for k = 0 to i;

The number of bins of the histogram is retained to be same for any type of cluster. The number of bins of the histogram is chosen to be '10' for any range of distances computed across any clusters. This is essential for the effective characterization of distances into histograms and to maintain uniformity among cluster representations that aid in accurate comparisons.

Normalization of cumulative histogram is achieved by dividing all the peak indices by the largest peak index of the cumulative histogram which is c_{10} as shown in Fig. 2b. The height of the bins of the histogram represents frequency counts of the distances that differ with the number of samples in a cluster. The height of a cumulative histogram increases with the number of samples and the corresponding distances. This results in a steeper slope when a regression line is fit over the cumulative histogram. The discriminative measure proposed in Sect. 3.2 cannot yield valid results due to the large variation. Hence, the cumulative histogram is normalized. A 1st order polynomial is fitted across 10 peak indices of the bins in the normalized cumulative histogram. It is illustrated in Fig. 2c.

A typical simple regression line is shown in Fig. 2d.
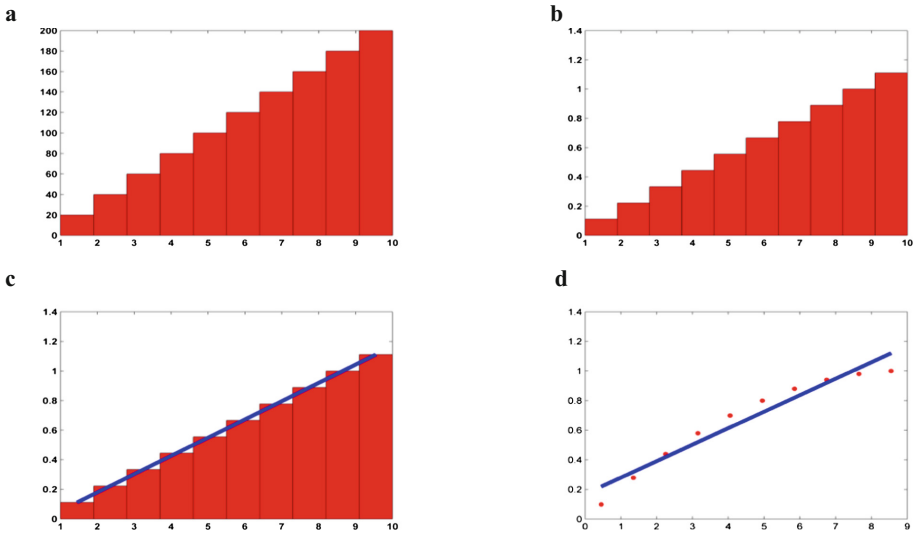
Fig. 2. (a) Cumulative histogram (b) Normalized cumulative histogram (c) Fitting a regression line (d) A sample regression line

The mathematical expression for the straight line is given by, $y = mx + c + e$ where m represents slope, c represents intercept and e is the error between the model and observation.

Slopes differ when clusters have different range and distribution of distances. Even a small change in the distribution of distances or range is easily reflected on the regression line where the slopes, intercepts and the error values change. Two clusters can have same range of distance but different distribution. Two clusters can have different range of distance but same distribution. In either case, the regression lines vary indicating difference in clusters.

3.2   Discrimination of Clusters

To determine the similarity or dissimilarity of two clusters, a comparative measure is needed. Since the clusters are represented by regression line, a comparison of clusters yields comparison of two regression lines. Each regression line is represented by slope intercept and error. Along with these, additional parameters - sum of squared errors, mean and standard deviation of range of distance values are retained. Sum of squared errors is the sum of the squared deviations from actual observed values.

Determining statistically significant difference between the two slopes and intercepts indicates that the regression lines are different. To determine the statistical difference, a p-value hypothesis testing is performed. The test is initiated by considering two hypotheses.

H_0 – Null hypothesis indicating slopes/Intercepts are similar which is tentatively held true.
H_1 – Alternate hypothesis indicating slopes/Intercepts are different

P-value, a computed probability is used which forms a boundary between rejecting and not rejecting the hypothesis. P-values range between 0 and 1 providing evidence against null hypothesis. Probability towards 0 indicates null hypothesis has strong evidence to be untrue. A critical value – alpha α, represents an error rate when the null hypothesis is rejected when true. P-value can be compared to critical alpha value to cull the hypothesis.

Cull H_0, when p-value is less than or equal to α
Do not cull H_0, when p-value is greater than α

Thus, p-value determines whether the difference between two regression lines is statistically significant or not based on the critical value. Initially, p-value is computed to find the difference between the slopes. If the slopes are not significantly different then the p-value comparing intercepts is computed. This is required because the lines could be near parallel with distinct intercepts. If p-value of slope and intercept is greater than α, then it is concluded that the lines are not significantly different. If p-value of slope comparison is less than α, then there is no need to compute p-value for intercept comparison. It is concluded that the lines are significantly different. In literature, detailed procedure for calculating p-value statistics for slope and intercept comparisons are provided [23].

Procedure for Calculating P-value for Comparison of Slopes. Let S_1 and S_2 be the slopes of two regression lines. Let SE_{S1} and SE_{S2} be the standard error of slope 1 and slope 2 respectively. The test statistic is computed as

$$T_S = (S_1 - S_2)/\sqrt{SE_{s1}^2 + SE_{s2}^2} \tag{2}$$

P-value for slope comparison can be obtained using Student's T cumulative distribution function on the test statistic [24].

Procedure for Calculating P-value for Comparison of Intercepts. Let I_1 and I_2 be the Intercepts of two regression lines. Let SSE_1 and SSE_2 be the sum of squared errors of regression lines. The test statistic is computed as,

$$T_I = (I_1 - I_2) * /(S_{I_1 - I_2}) \tag{3}$$

where,

$$S_{I_1 - I_2} = \sqrt{S_{yx}\left[\frac{1}{n_1} + \frac{1}{n_2} + \frac{M_1^2}{SD_{x_1}^2} + \frac{M_2^2}{SD_{x_2}^2}\right]} \tag{4}$$

where, M_1 and M_2 are means of predictor variables of regression lines, n_1 and n_2 are number of samples of regression lines, $SD_{x_1}^2$ and $SD_{x_2}^2$ are the standard deviations of predictor variables multiplied by number of samples on each regression line, S_{yx} is given by

$$S_{yx} = \frac{(SSE_1 + SSE_2)}{n_1 + n_2 + 4} \tag{5}$$

Threshold Determination for Concluding Similarity or Dissimilarity of Clusters.
Significance level α is the threshold to cull the null hypothesis i.e. the null hypothesis is rejected when the p-value is less than α. This indicates that slopes portray statistically significant difference. As per literature in statistical research, widely acceptable threshold for declaring the difference between slopes to be significant is 0.01, 0.05 and 0.1, also termed as statistical significance levels [25, 26]. An acceptable significance level can be dependent on the application under use. A larger $\alpha = 0.1$ can be chosen for critical applications to be more certain of not missing detecting a difference that might exist. Hence significance level of $\alpha = 0.1$ is chosen. Two clusters are different if the p-value computed between them is less than 0.1. Two clusters are similar if the p-value computed between them is greater than 0.1. An advantage of using p-value hypothesis testing and significance level is that the same significance levels can be imposed irrespective of the range and type of clusters and irrespective of the applications under study. P-value would always range between 0-1. Hence, it offers a generalized threshold that is applicable to varied types of clusters across different applications.

4   Experimental Analysis and Results

The proposed cluster representation provides a means for efficient discrimination among clusters. To prove the effectiveness of the proposed cluster representations for discrimination, experimentations on synthetically generated clusters and clusters of real-datasets is provided. Experiments are separated into two sections. Section 4.1 demonstrates first set of experiments where similar clusters are compared. It is shown that when clusters are similar, p-value hypothesis test yields very high probability value. Section 4.2 demonstrates the second set of experiments where dissimilar clusters are compared. It is shown that when clusters are dissimilar, p-value hypothesis test yields very low probability value. In cases where, slope-value is different, Intercept p-value is irrelevant as angle and extent of slope would already be different resulting in different intercept values.

4.1   Comparison of Similar Clusters

Two pairs of synthetically generated clusters are used to demonstrate similarity. Comparisons are named as Synthetic_Similar_1 and Synthetic_Similar_2. In the ensuing section, an overview of the clusters used for comparison, scatter plots of the clusters and corresponding experimental analysis and results are provided for each experiment.

Synthetic_Similar_1. This experimentation contains 2 synthetically generated spiral clusters with 105 data points in one cluster and 106 data points in another cluster. Each cluster is defined by two dimensions. Clusters are limited to two dimensions for visualization. Clusters are similar and this is evident in the scatter plots shown in Fig. 3a and b. The plot of the distances computed amongst the samples in the clusters is shown in Fig. 3c and d respectively. Y-axis portrays the distances whereas x-axis
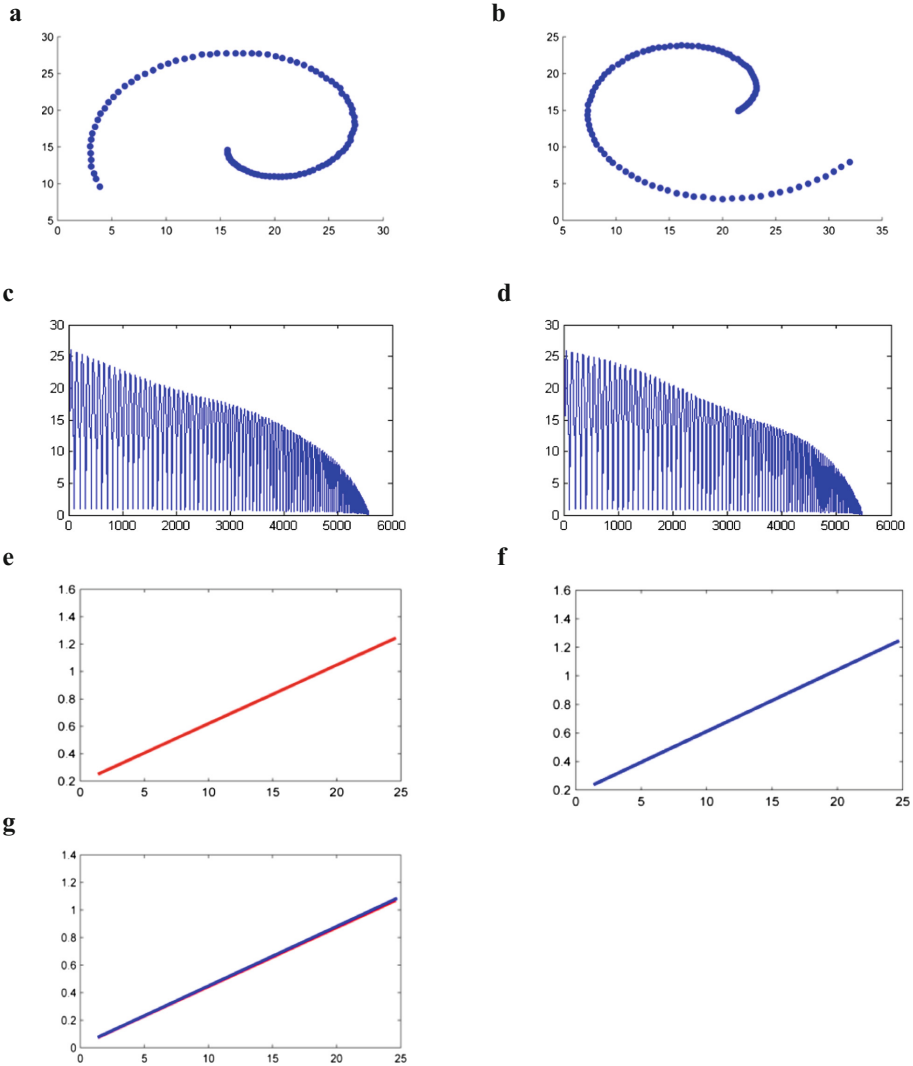
Fig. 3. (a and b) Scatter plots of cluster-1 & cluster-2 of Synthetic_Similar_1 Comparison, (c and d) Plot of the distances computed amongst the samples in cluster-1 & cluster-2 (e and f) Regression line transformation of cluster-1 and cluster-2 respectively (g) Regression lines of both clusters on a single plot

portrays the indices of distance vector. Figure 3e and f shows the regression line transformation of clusters respectively. Figure 3g shows both the regression lines on a single plot. It can be seen that the regression lines overlap indicating high similarity between the clusters. Slopes, Intercepts and corresponding p-value computed between the regression lines for slopes and intercepts are provided in Table 1.

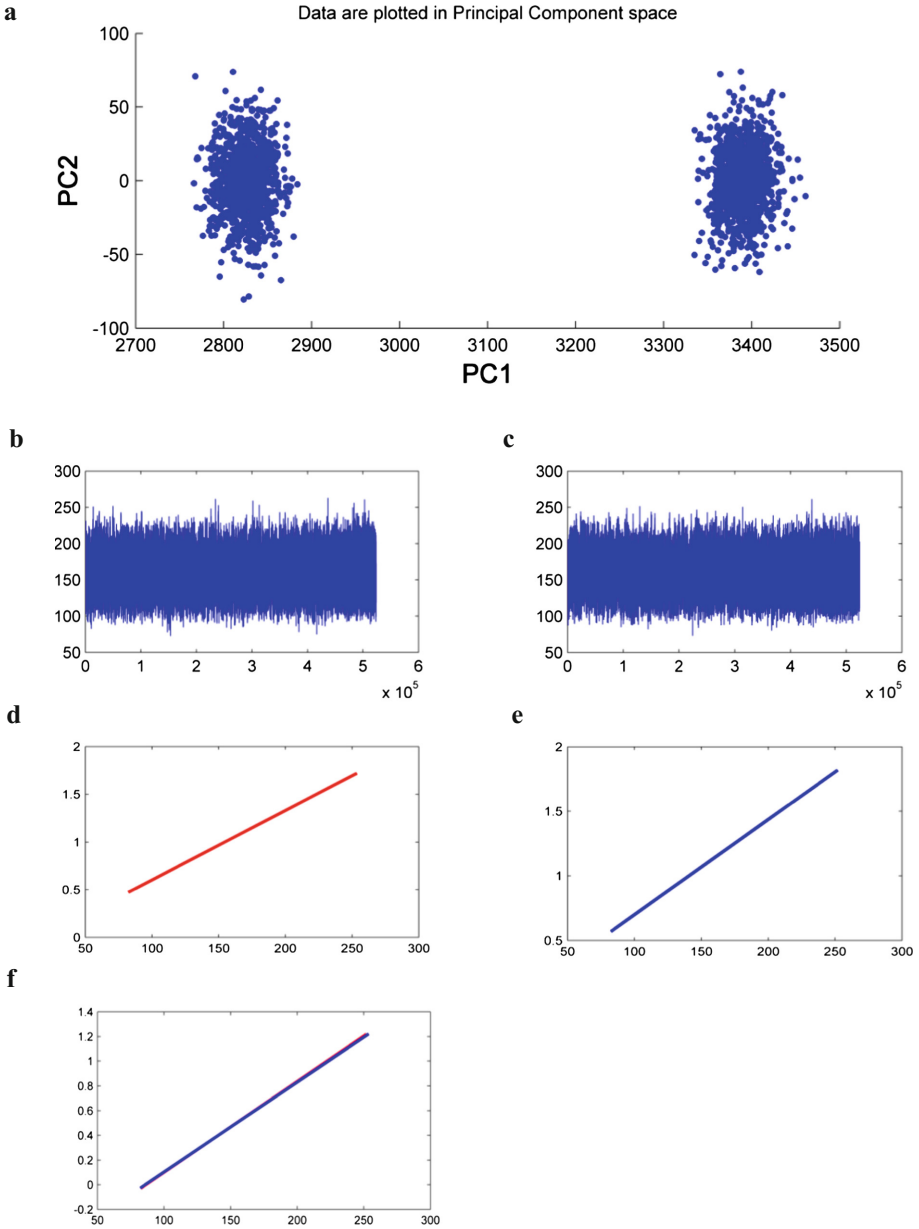**Table 1.** Slope, intercept and p-value hypothesis test results of clusters in Synthetic_Similar_1

Slope of cluster-1	0.0428
Slope of cluster-2	0.0432
Intercept of cluster-1	0.0149
Intercept of cluster-2	0.0172
P-value of slope	0.8976
P-value of intercept	0.9632

Synthetic_Similar_2. This experimentation contains two synthetically generated Gaussian clusters with normally distributed data having mean 500 and 600 respectively and standard deviation equal to 20. In each cluster, 32 attributes are used to represent 1024 samples. Since the clusters are high dimensional, the plot of the cluster is shown with principal components for visualization. Scatter plots of each cluster with respect to principal components is shown in Fig. 4a. Clusters are similar and this is evident in the scatter plot. The plot of the distances computed amongst the samples in the clusters are shown in Fig. 4b and c. Figure 4d and e displays the regression line transformation of clusters. Figure 4f displays both the regression lines on a single plot. It can be seen that the regression lines overlap indicating high similarity in clusters. Slopes, Intercepts and corresponding p-value computed between the regression lines for slopes and intercepts are provided in Table 2.

4.2   Comparison of Dissimilar Clusters

Two pairs of clusters are used to demonstrate dissimilarity. Each comparison is named as Synthetic_Dissimilar_1 and Real_Dissimilar_2. The experiment is conducted on one synthetically generated cluster and clusters from a real-dataset. In the ensuing section, an overview of the clusters used for comparison, scatter plots of the clusters and corresponding experimental analysis and results are provided for each experiment.

Synthetic_Dissimilar_1. This experimentation contains two clusters with different geometry and spread having 85 and 153 data samples respectively. Each cluster is defined by two dimensions. Figure 5a and b shows the scatter plots of each cluster. Figure 5c and d shows the plot of the distances computed amongst the samples in the clusters respectively. Figure 5e and f displays the regression line transformation of clusters respectively. Figure 5g shows both the regression lines on a single plot. It can be seen that the angle and extent of the two regression lines are different. Slopes, Intercepts and corresponding p-value computed between the regression lines for slopes are provided in Table 3. P-value of slope is less than 0.1 indicating that the clusters are very different. As the slopes are different, there is no need to compare intercepts.
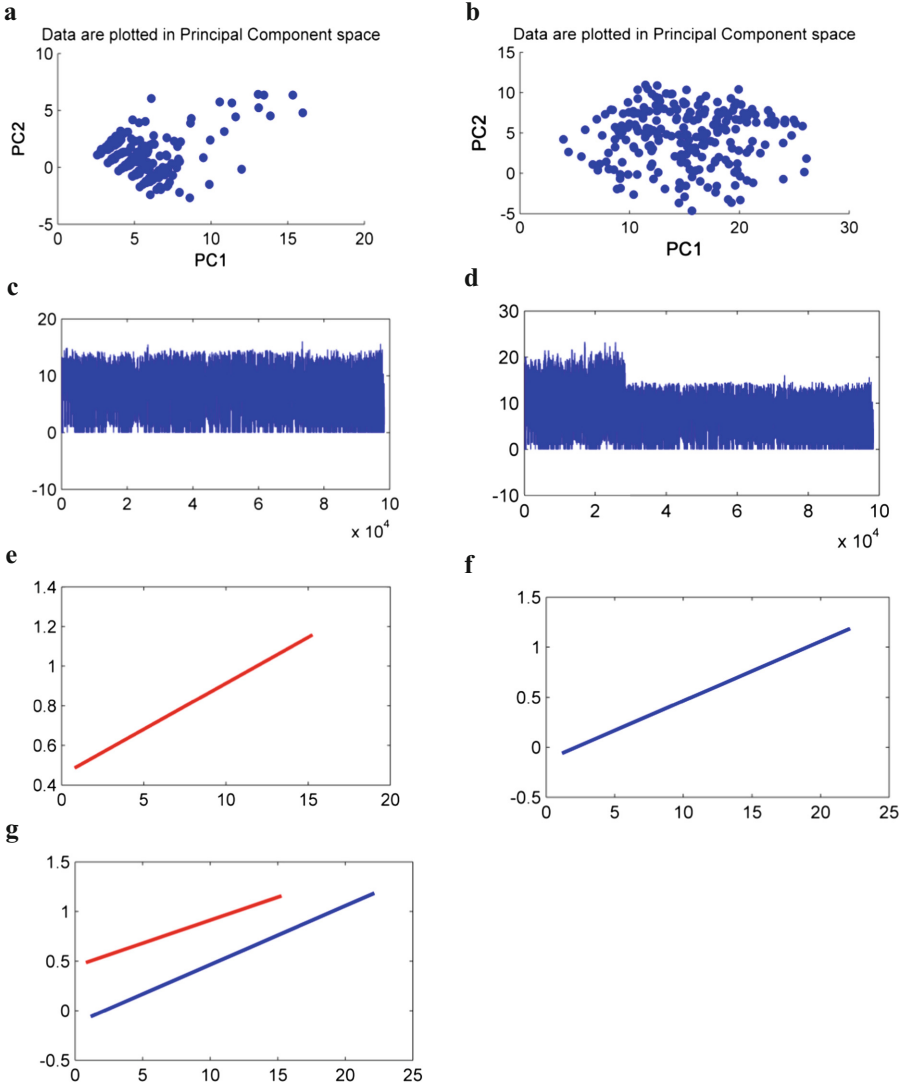
Fig. 4. (a) Scatter plots of cluster-1 & cluster-2 of Synthetic_Similar_2 Comparison, (b and c) Plot of the distances computed amongst the samples in cluster-1 & cluster-2 (d and e) Regression line transformation of cluster-1 and cluster-2 respectively (f) Regression lines of both clusters on a single plot

Table 2. Slope, intercept and p-value hypothesis test results of clusters in Synthetic_Similar_2

Slope of cluster-1	0.0073
Slope of cluster-2	0.0074
Intercept of cluster-1	−0.6248
Intercept of cluster-2	−0.6383
P-value of slope	0.95
P-value of intercept	0.93
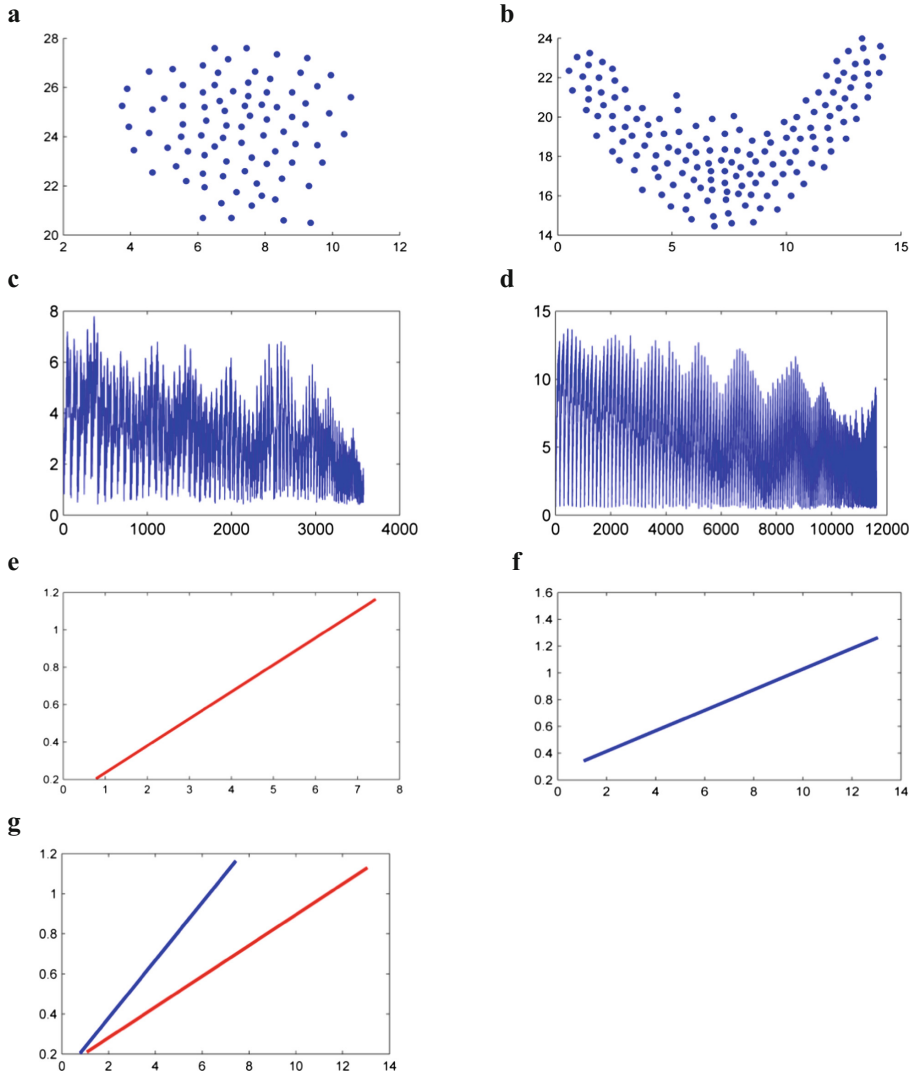
Fig. 5. (a and b) Scatter plots of cluster-1 & cluster-2 of Synthetic_Dissimilar_1 Comparison, (c and d) Plot of the distances computed amongst the samples in cluster-1 & cluster-2 (e and f) Regression line transformation of cluster-1 and cluster-2 respectively (g) Regression lines of both clusters on a single plot

Real_Dissimilar_2. This experimentation contains two clusters from real datasets downloaded from UCI machine learning repository [27]. Clusters are of Wisconsin Breast Cancer dataset. The two clusters represent Malignant and Benign cases. The Malignant cluster consists of 458 samples and Benign consists of 241 samples. Both the clusters are defined by 9 features. The clusters are dissimilar. Since the clusters are

Fig. 6. (a and b) Scatter plots of cluster-1 & cluster-2 of Real_Dissimilar_2 Comparison, (c and d) Plot of the distances computed amongst the samples in cluster-1 & cluster-2 (e and f) Regression line transformation of cluster-1 and cluster-2 respectively (g) Regression lines of both clusters on a single plot

high dimensional, cluster data is plotted with principal components for visualization. Figure 6a and b shows the scatter plots of each cluster along principal components. Figure 6c and d displays the plot of the distances computed amongst the samples in the clusters respectively. Figure 6e and f displays the regression line transformation of clusters respectively. Figure 6g shows both the regression lines on a single plot. It can be seen that the angle, extent and intercept of the two regression lines are different. Slopes, Intercepts and corresponding p-value computed between the regression lines for slopes are provided in Table 4. P-value of slope is not less than 0.1 but neither is it very high towards 1. However, P-value of Intercept is less than 0.1 indicating dissimilarity.

Table 3. Slope, intercept and p-value hypothesis test results of clusters in Synthetic_Dissimilar_1

Slope of cluster-1	0.1442
Slope of cluster-2	0.0767
Intercept of cluster-1	0.0915
Intercept of cluster-2	0.1281
P-value of slope	0.00048

5   Comparative Analysis

A comparative analysis has been tabulated with two non-parametric tests: Analysis of group similarities (ANOSIM) [14] and Multi-response permutation procedure (MRPP) [12] to discriminate between the clusters. The working of ANOSIM and MRPP is explained in the state-of-art. In both the methods clusters are directly discriminated without representation. A comparative analysis has not been tabulated with methods that represent the clusters before discrimination. These methods discriminate clusters based on pre-defined thresholds and thus the results of discrimination between the clusters cannot be validated with the proposed method. Similar to the proposed method, in ANOSIM and MRPP, the results of discrimination ranges between 0 and 1 indicating the extent of similarity between the clusters. For ANOSIM, result ranging towards 0 indicates similarity whereas towards 1 indicates no similarity. For MRPP, result ranging towards 1 indicates similarity whereas towards 0 indicates no similarity. The cluster datasets that are used in the experimental analysis are also used in comparative analysis and detailed descriptions of the clusters are provided in Sect. 3. Clusters in Synthetic_Similar_1 and Synthetic_Similar_2 are similar clusters whereas clusters in Synthetic_Dissimilar_1 and Real_Dissimilar_2 are dissimilar clusters. Table 5 shows the comparative analysis results on the cluster datasets using ANOSIM, MRPP and the proposed method. ANOSIM fails in discriminating Synthetic_similar_2 clusters. MRPP fails in discriminating Synthetic_similar_1 clusters and in all the other cases, it does not depict the results of discrimination clearly by providing extreme values towards 0 or 1. However, the results of MRPP can be considerably accepted based on the tendency of the results ranging towards 0 or 1. The proposed approach accurately discriminates the clusters in all the four cases based on slope and intercept p-values.

Table 4. Slope, intercept and p-value hypothesis test results of clusters in Real_ Dissimilar_2

Slope of cluster-1	0.0464
Slope of cluster-2	0.0593
Intercept of cluster-1	0.4492
Intercept of cluster-2	−0.1277
P-value of slope	0.31
P-value of intercept	0.0008

Table 5. Comparative results

Clusters for comparison	ANOSIM	MRPP	Proposed method	
			Slope	Intercept
Synthetic_Similar_1	0.0271	0.02	0.89	0.96
Synthetic_Similar_2	0.89	0.57	0.95	0.93
Synthetic_Dissimilar_1	1	0.22	0.00048	–
Real_Dissimilar_2	0.99	0.36	0.31	0.0008

6   Conclusion

Clusters are defined based on proximity of samples in feature space. Hence, each cluster has a unique cohesive strength. In this paper, the cohesiveness of the clusters is utilized to represent the clusters that aid in qualitative discrimination between the clusters. Cohesiveness of samples realized through distances between every pair of samples in a cluster is assimilated by histograms and then by regression line. Cluster is finally represented by slope, intercept and error characterizing the regression line. A transformation of clusters to regression line representation aids in faster and efficient discrimination of clusters. Experimentation on clusters from real datasets and synthetically generated clusters proves the efficacy of the proposed method in developing an efficient cluster representation that provides qualitative discrimination.

References

1. Aggarwal, C.C., Reddy, C.K.: Data Clustering: Algorithms and Applications, p. 652. Data Mining and Knowledge Discovery Series. Chapman and Hall/CRC (2013)
2. Wurzenberger, M., Skopik, F., Landauer, M., Greitbauer, P., Fiedler, R., Kastner, W.: Incremental clustering for semi-supervised anomaly detection applied on log data. In: Proceedings of the 12th International Conference on Availability, Reliability and Security, Article No. 31. ACM (2017). https://doi.org/10.1145/3098954.3098973
3. Burbeck, K., Nadjm-Tehrani, S.: Adaptive real-time anomaly detection with incremental clustering. Inf. Secur. Tech. Rep. **12**(1), 56–67 (2007). https://doi.org/10.1016/j.istr.2007.02.004

4. Langone, R., Agudelo, O.M., De Moor, B., Suykens, J.A.K.: Incremental kernel spectral clustering for online learning of non-stationary data. Neurocomputing **139**, 246–260 (2014). https://doi.org/10.1016/j.neucom.2014.02.036
5. Sun, Z., Mao, K.Z., Tang, W., Mak, L.-O., Xian, K., Liu, Y.: Knowledge-based evolving clustering algorithm for data stream. In: IEEE International Conference on Service Systems and Service Management, pp. 1–6 (2014). https://doi.org/10.1109/icsssm.2014.6874031
6. Rendón, E., Abundez, I., Arizmendi, A., Quiroz, E.M.: Internal versus external cluster validation indexes. Int. J. Comput. Commun. **5**(1), 27–34 (2011)
7. Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J.M., Perona, I.: An extensive comparative study of cluster validity indices. Pattern Recogn. **46**(1), 243–256 (2013). https://doi.org/10.1016/j.patcog.2012.07.021
8. Dean, S., Illowsky, B.: Descriptive statistics: histogram. Retrieved from the Connexions Web site (2009). http://cnx.org/content/m16298/1.11
9. http://infolab.stanford.edu/~ullman/mmds/ch7.pdf
10. http://www.statsdirect.com/help/basics/p_values.htm
11. Anderson, M.J.: Permutational multivariate analysis of variance. Dept. Stat. Univ. Auckland **26**, 32–46 (2005)
12. Cai, L.: Multi-response permutation procedure as an alternative to the analysis of variance: an SPSS implementation. Behav. Res. Methods **38**(1), 51–59 (2006)
13. Guillot, G., Rousset, F.: Dismantling the mantel tests. Methods Ecol. Evol. **4**(4), 336–344 (2013). https://doi.org/10.1111/2041-210x.12018
14. Clarke, K.R.: Non-parametric multivariate analyses of changes in community structure. Austral Ecol. **18**(1), 117–143 (1993). https://doi.org/10.1111/j.1442-9993.1993.tb00438.x
15. Aggarwal, C.C., Han, J., Wang, J., Yu, P.S.: A framework for clustering evolving data streams. In: Proceedings of the 29th International Conference on Very large Data bases, vol. 29, pp. 81–92. VLDB Endowment (2003)
16. Yang, C., Zhou, J.: HClustream: a novel approach for clustering evolving heterogeneous data stream. In: Sixth IEEE International Conference on Data Mining Workshops (2006). https://doi.org/10.1109/icdmw.2006.89
17. Komkrit, U., Rakthanmanon, T., Waiyamai, K.: E-stream: evolution-based technique for stream clustering. In: Alhajj, R., Gao, H., Li, J., Li, X., Zaïane, O.R. (eds.) ADMA 2007. LNCS (LNAI), vol. 4632, pp. 605–615. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73871-8_58
18. Meesuksabai, W., Kangkachit, T., Waiyamai, K.: Hue-stream: evolution-based clustering technique for heterogeneous data streams with uncertainty. In: Tang, J., King, I., Chen, L., Wang, J. (eds.) Advanced Data Mining and Applications, vol. 7121, pp. 27–40. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-25856-5_3
19. Nagabhushan, P., Ali, S.Z., Pradeep Kumar, R.: A new cluster-histo-regression analysis for incremental learning from temporal data chunks. Int. J. Mach. Intell. **2**, 53–73 (2010). https://doi.org/10.9735/0975-2927.2.1.53-57
20. Deza, M.M., Deza, E.: Encyclopedia of distances. In: Deza, M.M., Deza, E. (eds.) Encyclopedia of distances, pp. 1–583. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-00234-2
21. Nagabhushan, P., Pradeep Kumar, R.: Histogram PCA. In: Liu, D., Fei, S., Hou, Z., Zhang, H., Sun, C. (eds.) ISNN 2007. LNCS, vol. 4492, pp. 1012–1021. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-72393-6_120
22. Kumar, R.P., Nagabhushan, P.: An approach based on regression line features for low complexity content based image retrieval. In: IEEE International Conference on Computing: Theory and Applications (2007). https://doi.org/10.1109/iccta.2007.25

23. Wuensch, K.L.: Comparing correlation coefficients, slopes, and intercepts (2007). http://core.ecu.edu/psyc/wuenschk/docs30/CompareCorrCoeff.pdf

24. Grigelionis, B.: Student's t-distribution. In: Lovric, M. (ed.) International Encyclopedia of Statistical Science, pp. 1558–1559. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-04898-2

25. Krzywinski, M., Altman, N.: Points of significance: significance, P values and t-tests. Nat. Methods **10**(11), 1041–1042 (2013). https://doi.org/10.1038/nmeth.2698

26. https://onlinecourses.science.psu.edu/statprogram/node/138

27. UCI Machine Learning Repository (2017). http://archive.ics.uci.edu/ml/

Recognition of Traffic Sign Based on Support Vector Machine and Creation of the Indian Traffic Sign Recognition Benchmark

Vidyagouri B. Hemadri$^{(\boxtimes)}$ and Umakant P. Kulkarni

SDMCET, Dharwad 580 002, India
vidya_gouri@yahoo.com, upkulkarni@yahoo.com

Abstract. Traffic sign recognition, a driver assistance system informs and warns the driver about the status of the road is a challenging issue. Though, a lot of work on this topic has been carried out, but complete benchmark datasets are not freely available for comparison of different approaches. A few databases are available for benchmarking automatic detection of traffic signs. However, there is no database built considering the Indian traffic signs. The road scenarios in India are quite different from other countries, especially in rural areas. Hence, an effort to build an Indian traffic sign database considering both rural and urban situations is presented in the work. The database consists of 13000 traffic sign images of 50 different classes of traffic signs taken at different times under different environmental conditions and includes the detailed annotation of the traffic signs in terms of size, type, orientation, illumination and occlusion. The work also discusses an efficient method for identification of road signs based on two modules: (1) feature extraction based on dense scale invariant feature transform (DSIFT) and (2) a classifier trained by support vector machine (SVM). The SIFT approach transforms an image it into a large collection of local feature vectors invariant to scaling, translation or rotation of the image, and reduction in the dimensionality is achieved by applying principal component analysis (PCA). After extracting the features, the image is classified using support vector machine, a supervised learning model.

Keywords: Dense scale invariant feature transform · Pattern recognition
Principal component analysis · Support vector machine

1 Introduction

Recognition of traffic sign is a categorization problem having several classes of traffic signs. Even though, much work is done on the identification of road signs, but a complete benchmark datasets for identification of traffic signs are less considered. Hence, the comparison of the work on the recognition of traffic sign is a bit difficult. A small number of datasets are available as a benchmark for automatic recognition of traffic signs. Stallkamp *et al.* have contributed a German traffic sign dataset which has more than thirty thousand images of German traffic signs in forty three classes for public accessibility [1]. Larsson and Felsberg have contributed a Swedish traffic sign dataset and, have nearly twenty thousand frames with 20% being labeled [2].

Katholieke Universiteit of Leuven (KUL) Belgium Traffic Signs Dataset (BelgiumTS) is a dataset with 10000 + traffic signs. The material is captured in Belgium in urban environments [3]. Ruta *et al*. have focused on the use of discriminative feature selection algorithms for identification and tracking of traffic signs [4]. Mogelmose *et al*. have described various approaches used in the traffic sign recognition, especially in various stages like segmentation, feature extraction, and sign detection and, also explained the deficiency in the use of publicly available image datasets [5]. Recognition of traffic sign based on Eigen values was proposed by Fleyeh and Davami and was tested on database of traffic sign's borders and speed limit pictograms [6]. Zhu *et al*. have created Tsinghua-Tencent 100 K benchmark dataset of road sign and database has 100000 Tencent Street View panoramas containing 30000 traffic-sign instances [7]. Yang *et al*. focused on the identification of traffic signs based on color probability model with convolutional neural network and also built a Chinese Traffic Sign Dataset (CTSD) with 1100 images [8].

Soilan *et al*. developed a publicly available Spanish traffic sign dataset. The dataset has approximately one thousand five hundred images with fifty to hundred classes of Spanish traffic signs [9]. The ensemble of bag-of visual- phrases and hierarchical deep models to recognize the traffic signs was discussed in the work [10]. Detection of traffic signs based on cascaded convolutional neural networks (CNN) was presented in the paper [11]. Detection of location of traffic signs based on hybrid region model and classification based on Fast R-CNN was presented in the work [12]. The use of off-line detector, online detector, and motion model predictor for simultaneously detecting and tracking road signs and a scale-based intra-frame fusion method for classification of the traffic signs was discussed in the work [13]. Recognition of traffic signs based on the use of kernel-based extreme learning machine with deep perceptual features was described in the paper [14]. The use of fuzzy rule based color segmentation method for the detection of traffic signs followed by the recognition of the traffic signs based on histogram oriented gradient features and speeded up robust features with artificial neural network was presented in the work [15]. Recognition of traffic signs using support vector machine with band of features including local binary patterns, histogram of oriented gradients, and Gabor features was described in the work [16]. Use of random forest and support vector machine classifiers with the feature descriptor obtained by the ensemble of histogram of oriented gradients features and local self-similarity feature to detect and recognize the traffic signs was focused in the work [17]. Use of convolutional neural network to detect and recognize the traffic sign was presented in the work [18]. Recognition of traffic sign based on extreme learning machine with histogram of oriented gradient variant feature was presented in the work [19]. Use of linear support vector machine to detect a traffic sign in a real time is described in the work [20]. Detection and recognition of traffic signs based on three dimensional point cloud models was described in the work [21].

But, not much work is done on the Indian traffic sign recognition benchmark. Hence, an Indian traffic sign recognition benchmark considering both rural and urban situations is presented in the work. The dataset consists of 13000 traffic sign images of 50 different classes of traffic signs taken at different times under different environmental conditions and at different places. Classification accuracy of k-nearest neighbor and support vector machine classifiers by extracting scale invariant feature transform

followed by principal component analysis features is computed on the database. The proposed work also discusses the recognition of the unknown traffic sign using SVM with DSIFT.

2   Indian Traffic Sign Database

The Indian traffic sign database (INDTRDB) [22] is primarily designed as a benchmark for traffic sign recognition research keeping in view the systems for assisting the driver. The INDTRDB consists of 13000 traffic signs images of mandatory, cautionary and informatory signs taken under various conditions, such as during daytime, afternoon, night, and ideal images, partially occluded and distorted images. Out of 50 different classes of traffic sign images in a database, 22 images are of mandatory, 21 of cautionary and 07 of informatory sign images. The detailed annotation of the INDTRDB in terms of size, type, orientation, illumination and occlusion is given with a database. The design of the database consists of the following steps:

1. Data collection
2. Selection and cropping of traffic signs
3. Data organization and restructuring
4. Manual annotation.

3   Data Collection

The images of traffic signs are captured using Prosilica GX 1920c camera with automatic exposure control with frame rate of 112fps having resolution of 1936×1456. The images are taken at different roads of India considering both urban and rural areas. Figure 1 shows the picture having the traffic sign taken from some roadside.

Fig. 1. Picture containing traffic signs

4   Selection and Cropping of Traffic Signs

Traffic signs are designed such that they are easily identified by human drivers. In this work the images containing the traffic signs are captured under various conditions like varying lighting conditions (daytime and night time), different seasons, at different places. After the frames are selected, the traffic signs are manually cropped. The signs are cropped in such a way that the images contain at least five pixels around the sign in order to maintain consistency across the images as the signs may vary with respect to orientations. The images are not necessarily in square and the size varies from 30×30 to 160×160. The dataset contains nearly 13000 images of 50 different classes. Figure 2 shows few samples of different classes of Indian traffic signs in a dataset. Every sign in the dataset varies with other sign either with respect to size, orientation, illumination, occlusion, background, contrast and quality. Figure 3 shows the sample images of mandatory traffic sign 'NO PARKING'.

5   Data Organization and Restructuring

The duplicate or similar images were removed from each class. Each image of a class is varied in size from 30×30 to 160×160. The images are resized in such a way that few images are square images and few are rectangular images. Different instances are taken for each class of traffic signs such as at different places, with different orientations, perfect signs, distorted signs, partially occluded signs. The contrast of each image is varied manually and considered as different instance of that sign. The dataset was split into two subsets with 50% images taken for training and remaining 50% for testing.

Fig. 2. Different classes of traffic signs in INDTRDB

Fig. 3. Sample images of "NOPARKING" traffic signs in INDTRDB database

6   Manual Annotation

A manual annotation of every image for INDTRDB dataset is provided with the following attributes. The annotation details of the traffic sign "STOP" is shown in Fig. 4. XML description is used to store annotated information.

・Size	In the range of 30×30 to 160×160
・Illumination	Good, Bad
・Orientation	Frontal, Left, Right
・Occlusion	Partial, No
・Contrast	Ideal, Dark, Bright
・Coordinate	(x, y) coordinate of the top left corner and bottom right corner.
・Category	Category of the traffic sign images.
・Folder name	The name of the folder which contains the required file
・File name	Name of the sign image in the folder

Fig. 4. XML annotation of STOP traffic sign

7   Database Complexity Analysis

The Indian traffic sign dataset INDTRDB has a wide range of variations in the traffic signs. Classification accuracy of KNN and SVM classifiers with PCA and DSIFT feature extractor is computed on the database. The experiment is implemented using MATLAB-2012b. The system divides the dataset created into two sets namely test set and training set. Various combinations of feature extractor and classifier algorithms are applied to get the best classification accuracy for the dataset created. The dataset shows good result for DSIFT with SVM. The classification accuracy of KNN and SVM classifiers on the INDTRDB, GTSRB, BelgiumTS database is shown Table 1 for various feature dimensions ranging from 20 to 80 with value of K taken as one. There is no much variation in the recognition rate by further increase of feature dimension. The comparison of INDTRDB dataset with GTSRB and BelgiumTS is shown in the Fig. 5. The result shows that both KNN and SVM methods has shown good classification accuracy on the INDTRDB database as compared to GTSRB and BelgiumTS database. The complexity of the dataset is better understood by the distribution of images in the dataset. For further comparison, a total of 600 images were selected from

GTSRB, BelgiumTS, and INDTRDB with 100 images from six different classes and the Eigenvalue spread of the covariance matrix of the preferred signs is analyzed for all three datasets. The magnitude of Eigenvalue indicates the prominent directions along which the data have highest variations and it will be less, if the dataset has a smaller number of variations. Figure 6 shows the Eigenvalue range of GTSRB, BelgiumTS and INDTRDB. It is clear that, number of components needed to cover the space for INDTRDB and GTSRB is large compared to BelgiumTS dataset. This is mainly because of large variations of the signs in the dataset. Figures 7 and 8 shows the related mean signs and Eigenvectors for the same subset of datasets. The mean sign of INDTRDB dataset is more spread compared to BelgiumTS, but not as GTSRB database. The mean sign of BelgiumTS resembles the structure of traffic sign indicating less variation.
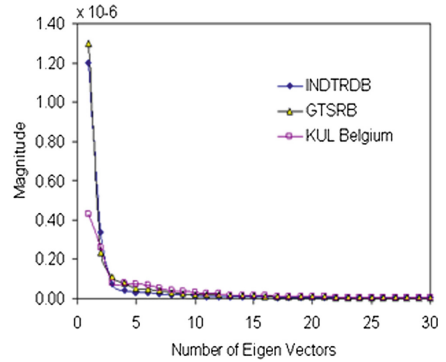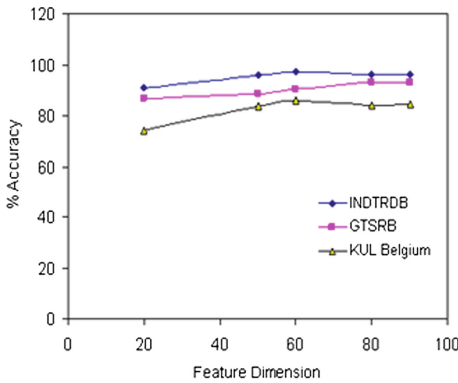
Fig. 5. Comparison of INSTRDB, GTSRB and BelgiumTS database

Fig. 6. Eigen value spectrum of subset of INDTRDB, GTSRB, and Belgium TS
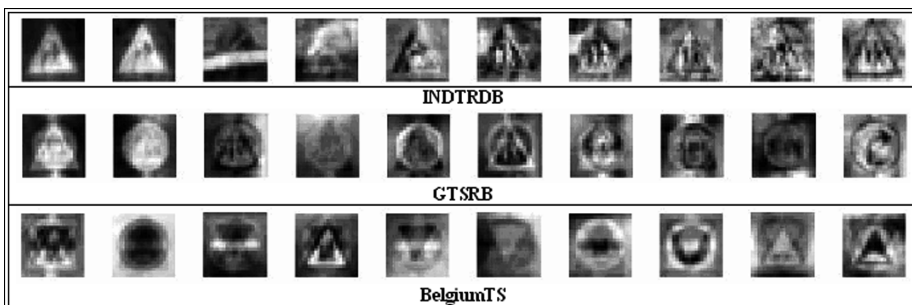
Fig. 7. Eigen vectors (top 10) corresponding to subset of INDTRDB (top), GTSRB (middle), and BelgiumTS (bottom) traffic signs
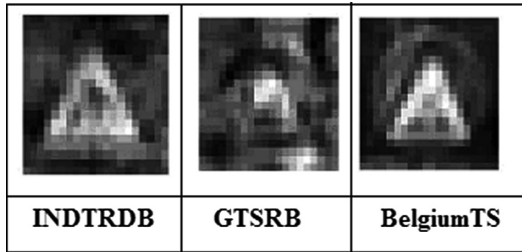
Fig. 8. Mean traffic signs of subset of databases with 6 different traffic signs from (left) INDTRDB, (middle) GTSRB and (right) BelgiumTS

Table 1. Classification accuracy of three datasets.

Feature dimension	INDTRDB				GTSRB				BelgiumTS			
	PCA		DSIFT		PCA		DSIFT		PCA		DSIFT	
	KNN	SVM	KNN	SVM	KNN	SVM	KNN	SVM	KNN	SVM	KNN	SVM
20	91.13	91.01	91.37	90.89	87.37	83.45	72.03	86.67	70.37	76.71	69.49	74.28
50	90.78	94.56	92.08	95.86	89.21	86.72	78.38	88.63	76.49	86.57	75.34	83.44
60	89.81	91.26	93.25	97.08	89.32	86.83	78.89	90.15	76.66	89.17	75.39	85.69
80	89.41	92.72	92.59	96.03	89.65	88.6	79.73	92.89	77.15	88.2	76.18	84.06
90	90.87	92.72	91.93	96.03	89.45	90.68	79.73	92.89	77.59	86.13	76.22	84.5

8   Recognition of Traffic Sign

In the recognition phase, the system is provided with an unknown test image to identify to which class the traffic sign image belongs. Road signs have a predetermined shape and color with some information and are distinguished by a fixed number of relatively constant colors such as white, red, and blue. To have a clear visibility, signs have enough color contrast with the background. In this work, detection of the road sign from the scene is instigated using color information. The work uses HSV color space to locate the signs from the scene. Hue and Saturation components are used to detect the red and blue colors. The images in the scene having the required colors are further classified based on shapes. Canny algorithm is used to find the required shapes in the image. The classification of the detected image is done using support vector machine with dense scale invariant feature transform feature extractor. After extracting the features, the image is classified using support vector machine, a supervised learning model. The method to recognize unknown traffic sign is given in Algorithm 1. The function takes SVM model, test image, the candidate image that needs to be predicted and will produce output, which is the prediction of the class with the highest probability. The summary of the traffic sign recognition system is shown in Fig. 9. Out of

100 unknown traffic signs 97 signs are correctly classified and 3 are misclassified due to higher occlusion, high distortion, etc. Illustration of four correctly classified and one misclassified traffic sign image is shown in Table 2. The sign is misclassified as "NO ENTRY" instead of "NO PARKING" as both signs have same shape with red color with a cross bar inside the sign. The only information that differentiates these signs is the presence of letter "P" in case of "NO PARKING" whereas an "Arrow Mark" in case of "NO ENTRY".

Fig. 9. Recognition of traffic sign

```
Algorithm 1 ( Recognition of Traffic Sign )
      Input : Indian Traffic Sign Database ( INDTRDB), Sign Image
      Output :  Type of Traffic Sign
Method
  1. Read the sift features of INDTRDB
  2. For i ← 1 : sizeof(INTTRDB)
          // Checked for various feature dimension between 20 to 90
        [fea] = dsift(single(INDTRDB(i)), 'step', 20)
        [fea]=PCA(fea, feature_dim);
        INDTRDB_sift_feaures(:,i) ← double(fea( : ))  // Returns matrix
                                                         with descriptors
    Endfor
  3. [fea] = dsift(single(SIGN),'step',20);
  4. [fea]= PCA(fea, feature_dim);
  5. Signfea = double(fea( : ) )
        // Returns SVM model
  6. [SVMmodel, mean_train, std_train] = SVMTrain(INDTRDB_sift_features,
    labels, 'RBF')
    // Returns class of SIGN image
  7. Class = SVMPredict(Signfea,SVMmodel,mean_train,std_train)
  8. If ( class = class label in INDTRDB database)
    Print 'Type of the traffic sign'
   Else
     Print 'NOT Traffic sign'
   EndIf
End of algorithm
```

Table 2. Recognition of unknown traffic sign

Unknown Image	Recognition Result	Class Recognized
	Successfully Classified	STOP
	Successfully Classified	LEFT HAND CURVE
	Successfully Classified	OVERTAKING PROBIHITED
	Successfully Classified	NO PARKING
	Unsuccessful Classification	Recognized as: NO PARKING Actual Class: NO ENTRY

9   Conclusion

A new dataset called INDTRDB, an Indian traffic sign dataset is developed to provide a general reference for traffic sign recognition and for the related research area. The INDTRDB dataset has large numbers of images nearly 13000 sign images taken from 50 diverse classes which are taken at diverse illumination and environmental conditions with a detailed manual XML annotation. The classification accuracy of support vector machine on the INDTRDB, GTSRB, and BelgiumTS databases is found to be 98.67%, 93% and 86% respectively. The number of major principal components required to span the space for INDTRDB and GTSRB is large compared to BelgiumTS dataset, which shows the INDTRDB and GTSRB dataset has large variations of the signs in the dataset. The work also describes the recognition of the unknown traffic sign using DSIFT and SVM with a recognition accuracy of 98%. Out of 120 unknown traffic sign images 117 signs are correctly classified and, three are misclassified due to higher occlusion, high distortion, etc.

References

1. Stallkamp, J., Schlipsing, M., Salmen, J., Igel, C.: Man vs. computer: benchmarking machine learning algorithms for traffic sign recognition. Neural Netw. **32**, 323–332 (2012). https://doi.org/10.1016/j.neunet.2012.02.016
2. Larsson, F., Felsberg, M.: Using fourier descriptors and spatial models for traffic sign recognition. In: SCIA, vol. 11, pp. 238–249 (2011). https://doi.org/10.1007/978-3-642-21227-7_23
3. Timofte, R., Zimmermann, K., VanGool, L.: Multi-view traffic sign detection, recognition, and 3D localisation. Mach. Vis. Appl. **25**(3), 633–647 (2011). https://doi.org/10.1007/s00138-011-0391-3
4. Ruta, A., Li, Y., Liu, X.: Real-time traffic sign recognition from video by class-specific discriminative features. Pattern Recogn. **43**(1), 416–430 (2010). https://doi.org/10.1016/j.patcog.2009.05.018
5. Mogelmose, A., Trivedi, M.M., Moeslund, T.B.: Learning to detect traffic signs: comparative evaluation of synthetic and real-world datasets. In: 21st International Conference on Pattern Recognition (ICPR), pp. 3452–3455. IEEE (2012)
6. Fleyeh, H., Davami, E.: Eigen-based traffic sign recognition. IET Intell. Transp. Syst. **5**(3), 190–196 (2011). https://doi.org/10.1049/iet-its.2010.0159
7. Zhu, Z., Liang, D., Zhang, S., Huang, X., Li, B., Hu, S.: Traffic-sign detection and classification in the wild. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2110–2118 (2016)
8. Yang, Y., Luo, H., Xu, H., Wu, F.: Towards real-time traffic sign detection and classification. IEEE Trans. Intell. Transp. Syst. **17**, 2022–2031 (2016). https://doi.org/10.1109/TITS.2015.2482461
9. Soilan, M., Riveiro, B., Martinez-Sánchez, J., Arias, P.: Traffic sign detection in MLS acquired point clouds for geometric and image-based semantic inventory. ISPRS J. Photogrammetry Remote Sens. **114**, 92–101 (2016). https://doi.org/10.1016/j.isprsjprs.2016.01.019
10. Yu, Y., Li, J., Wen, C., Guan, H., Luo, H., Wang, C.: Bag-of-visual-phrases and hierarchical deep models for traffic sign detection and recognition in mobile laser scanning data. ISPRS J. Photogrammetry Remote Sens. **113**, 106–123 (2016). https://doi.org/10.1016/j.isprsjprs.2016.01.005
11. Zang, D., Zhang, J., Zhang, D., Bao, M., Cheng, J., Tang, K.: Traffic sign detection based on cascaded convolutional neural networks. In: 2016 17th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), pp. 201–206. IEEE (2016). https://doi.org/10.1109/snpd.2016.7515901
12. Qian, R., Liu, Q., Yue, Y., Coenen, F., Zhang, B.: Road surface traffic sign detection with hybrid region proposal and fast R-CNN. In: 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), pp. 555–559. IEEE (2016). https://doi.org/10.1109/fskd.2016.7603233
13. Yuan, Y., Xiong, Z., Wang, Q.: An incremental framework for video-based traffic sign detection, tracking, and recognition. IEEE Trans. Intell. Transp. Syst. **18**(7), 1918–1929 (2017). https://doi.org/10.1109/TITS.2016.2614548
14. Zeng, Y., Xu, X., Shen, D., Fang, Y., Xiao, Z.: Traffic sign recognition using kernel extreme learning machines with deep perceptual features. IEEE Trans. Intell. Transp. Syst. **18**(6), 1647–1653 (2017). https://doi.org/10.1109/TITS.2016.2614916

15. Abedin, Z., Dhar, P., Hossenand, M.K., Deb, K.: Traffic sign detection and recognition using fuzzy segmentation approach and artificial neural network classifier respectively. In: International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 518–523. IEEE (2017). https://doi.org/10.1109/ecace.2017.7912960

16. Berkaya, S.K., Gunduz, H., Ozsen, O., Akinlar, C., Gunal, S.: On circular traffic sign detection and recognition. Expert Syst. Appl. **48**, 67–75 (2016). https://doi.org/10.1016/j.eswa.2015.11.018

17. Ellahyani, A., El Ansari, M., El Jaafari, I.: Traffic sign detection and recognition based on random forests. Appl. Soft Comput. **46**, 805–815 (2016). https://doi.org/10.1016/j.asoc.2015.12.041

18. Zhu, Y., Zhang, C., Zhou, D., Wang, X., Bai, X., Liu, W.: Traffic sign detection and recognition using fully convolutional network guided proposals. Neurocomputing **214**, 758–766 (2016). https://doi.org/10.1016/j.neucom.2016.07.009

19. Huang, Z., Yu, Y., Gu, J., Liu, H.: An efficient method for traffic sign recognition based on extreme learning machine. IEEE Trans. Cybern. **47**(4), 920–933 (2017). https://doi.org/10.1109/TCYB.2016.2533424

20. Zaklouta, F., Stanciulescu, B.: Real-time traffic sign recognition in three stages. Rob. Auton. Syst. **62**(1), 16–24 (2014). https://doi.org/10.1016/j.robot.2012.07.019

21. Vahid, B., Arash, J., Saha, G.M.: Multi-class US traffic signs 3D recognition and localization via image-based point cloud model using color candidate extraction and texture-based recognition. Adv. Eng. Inform. **32**, 263–274 (2017). https://doi.org/10.1016/j.aei.2017.03.006

22. INDTRDB. https://tinyurl.com/indtrdb

Neural Network Based Characterization and Reliable Routing of Data in Wireless Body Sensor Networks

Biradar Shilpa[1](✉) ⓘ, S. G. Hiremath[2] ⓘ, and G. Thippeswamy[3] ⓘ

[1] Dr. AIT, Bangalore, India
biradarshilpa@gmail.com
[2] EWIT, Bangalore, India
sghiremath_2001@yahoo.com
[3] BMSIT, Bangalore, India
gt_swamy@yahoo.com

Abstract. Health is very important to each one and for the early discovery of different diseases to get timely treatment Wireless Body Sensor Networks (WBSNs) a subgroup of wireless sensor networks can be used. WBSNs works for collecting the health related data such as ECG, EEG, Glucose values, temperature etc. from human body and route the data towards the destination so that the patients information can be sent to the concerned person. In WBSN the sensor node can either be appropriately located on the body or inserted inside the body. In this work firstly the EEG signals characterization is done with the help of neural networks. EEG signal processing is done with the help of DWT. Then in the secondly routing mechanism used to send the data to the destination. In the health monitoring system routing has an important role. The facts sensed by the sensors need to be routed to the destination effectively. In this work a routing algorithm proposed which can be used to transfer the data to destination efficiently. To route the data to the sink multi hop process is used and the cost function is used to find the node next closest node in the route from source to destination. The experiments are done in MATLAB environment and simulation result shows that efficient classification of EEG signal is done and the proposed scheme provides the less end-to-end as well as good throughput.

Keywords: Wireless Body Sensor Network (WBSN)
Wireless Sensor Networks (WSN) · Daubechies Wavelet Transforms (DWT)
Electroencephalogram (EEG)

1 Introduction

Wireless sensor networks can be acclimated for the assorted purposes like traffic monitoring, taking care of plants, observing the infrastructure as well as health monitoring. Use of WSN in health is the emerging subfield entitled as Wireless body sensor network (WBSN). It is an interdisciplinary area used for nonstop monitoring of health conditions with real-time informs of medical records. In this for timely detection of patient's health

conditions is done with the help of computer assistance and integration of number of sensors is done into wearable wireless body area network [1]. In this either the tiny sensors are inserted inside the human body such a way that the person is comfortable and insertion of sensor shouldn't damage normal activities or they can be mounted on the body surface.

The sensors inserted in/mounted on the body surface will gather several functional changes of the body which are useful to observe healthiness status of the patient in this the physical location of patient doesn't effect. The health condition information gathered by the sensors will be transmitted to a processing device that is an external device wirelessly. This externally situated device will immediately communicate all up to date information to the related doctors all over the world with the help on internet. In case of emergency situations, the physicians will immediately advise the concerned person by communicating proper messages or gives some alarms [18].

The identification and analysis of epileptic seizure activity in humans can be done using Electroencephalogram (EEG). Usually Professionally skilled persons do the analysis of the EEG signal manually but number of such peoples are very small. Diagnosing the anomalous activities in the brain working is very important concern. Basically EEG signals contains the brain functionality information. In time domain the visual study of EEG signals may be inadequate as there is no certain criterion assessed by the experts. As regular clinical diagnosis is required for analyzing the EEG signals, some computerization techniques can be used. Temperature, respiration rate etc. can be sensed with the help of sensors and after sensing the health data it need to be send to the destination.

The different parts that can be focused in Body Sensor networks are shown in Fig. 1.
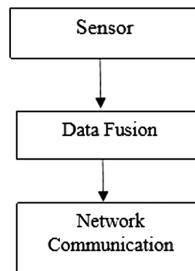
**Fig. 1.** Research areas in WBSN.

The work on design of sensors and its generally focuses on the sensor nodes wear ability of [5], The fault analysis capacity and the way of handling fault [6], energy consumption [7] and deployment of sensor [8], the data fusion basically deals with denosing concept in which the different method and their respective implementation for removal of noise is concentrated [9], extraction of features is discussed in [10], classification of data [11] and other key technologies. The next part is network communication [3], research the main focus is on topological planning of the network [12], characterization of channels [13], controlling channel access [14, 15] design of routing algorithm [16] etc.

2   Data Fusion

The data fusion process basically includes preprocessing, feature extraction, classification.

In human physiological signals usually there will be some kind of noise and interference, because of which the collected signal cannot be used as it is. For filtering the noise from the collected signals and getting the valuable information preprocessing of the signal can be done. The different preprocessing methods are Wavelet Transform (WT), Fourier Transform (FT), low pass as well as high pass filters are some of the examples that can be used for noise removal.

Feature extraction is concept about reducing the quantity of input data required to describe a large set of data. Selecting the features is nothing but the determination of a subgroup of the original features. The features are selected such that they have the information connected to the input data and the anticipated task can be completed by using this reduced representation. Once preprocessing is done feature extraction is carried out, where the representative data is assembled into feature vectors to do the segregation among different activities or phenomena [17]. In this stage, information analysis is done with the help of Fast Fourier Transform, wavelet analysis, etc., by doing extraction the features of gathered signals. Artificial Neural Network can be used for this purpose. The next phase is about classification. All these steps are shown in Fig. 2.
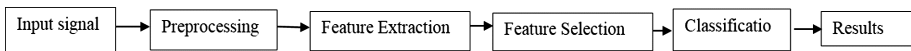
Fig. 2. Steps in classification of signals.

2.1   Classification of EEG Signal

In the proposed work DWT and ANN is used for the classification of EEG signal. In human it's surface amplitude value lie in the extend about 10 to 100 µV. The EEG signal's s recurrence value has some lower and upper bound but from the physiological viewpoint the most significant frequencies lie among the values 0.1 to 30 Hz.

Clinically the guidelines portrayed for eeg signal groups are provided likewise: those delta band will be Hosting recurrence in the reach from 0.1 Hz to 3.5 Hz, theta band is hosting recurrence in the go (4 with 7.5 Hz), alpha band is having frequency in the extent (8 to 13 Hz), Furthermore beta band will be having the frequency in the range (14 on 30 Hz) where as gamma waves constitute signs that have frequencies more over 30 Hz. whereas gamma waves constitute signals that have frequencies greater than 30 Hz. For both the normal and epileptic patients he data of EEG signal consists of five groups [2].

The wavelet analysis is basically used to break down the signals into a number of frequency bands. While doing the analysis of signals with the help of DWT, selecting suitable wavelet and the number of decomposition levels are very important. The number of levels for decomposition is selected based on the main frequency components of the signal.

Selecting the number of levels for decomposition is very important in DWT. The levels has to be done in such a way that the selected signal parts associate properly

with the frequencies required for signal classification and are taken in wavelet coefficients. As the smoothing feature of DB4 is appropriate for identifying changes of the EEG signals so Daubechies 4 (db4) is used for computing the wavelet coefficients. Daubechies wavelets are the one of the standard form of wavelets demonstrating basics of wavelet signal processing. The signals were decomposed into details D1–D5 and one final approximation A5. Pattern recognition techniques used for the EEG signal classification. Pattern recognition is a method of observing a pattern of a given object based on the information already possessed [4]. Different computerized reasoning systems like artificial neural networks (ANN), fuzzy logic (FL) and adaptive fuzzy logic (AFL) are the automatic pattern recognition used to do the categorization of aggravation signs. Newly, probabilistic model based methods like Hidden Markov models, Dynamic time wrapping, Dempster-Shafer theory of evidence are also recommended. The method used here uses two steps: first stage is feature extraction (FES) where as other is classification stage (CS). To do the Classification process the input used is preprocessed signal. In this case, prior to giving the EEG signal as an input to the (CS) the signal transformation from time domain into the wavelet domain is done. In the classification stage, the suggested wavelet energy dissemination features are used as input to Neural Network (NN). NN is one of the prevailing tool used for pattern recognition. It can considered as the intermediater that is existing among input and output data, inclosing nonlinear relationships. Use of Feed-Forward Neural Network (FFNN) is done for EEG signal classification. Neuron is the fundamental unit of a NN, which can be denoted as a function of weighted summation. A FFNN structure can be deliberated as a mathematical operator, such as weighted summation and multiplication. So, it is conceivable to recreate a varied class of algorithms by utilizing multiplier module.

3 Routing

Once the EEG signals are classified as either normal or abnormal it need to be sent to the respective person in case of any abnormilty. The next phase of work is about sending the information which include network initialization, identification of source, establishment of reliable path Data transmission. The routing algorithm used is given below.

3.1 Parameter Based Reliable Routing Algorithm (PBRR)

It is a multilevel routing algorithm, as the main purposes of this algorithm is to reduce the loss of packets and Drop of packet, which will select the best value route of the network for reliable data transfer. The parameters considered for choosing the best path selection from the source to destination are hop count, free Buffer Size and distance between them.

Algorithm used to determine best path

Step 1: Determine Source, Destination node
Step 2: Describe the communication Range, Opening (initial) Energy
Step 3: Apply greedy Method and separation distance among nodes can be decided using Euclidian distance formula

Dist = In the Euclidean plane, if p = (p1, p2) and q = (q1, q2) then the distance is given by

$$d(p, q) = \sqrt{(q1 - p1)^2 + (q2 - p2)^2} \tag{1}$$

Step 4: Compute and show opening energy of the nodes
Step 5: Determine node states with highest parameter
Step 6: see whether the destination lies within the communication rang Range? 1:0 Update source and destination location
Step 7: if destination distance ≤ transmission range Update the source and routers
Step 8: Find the nodes distance, which all are within the communication limit of source node and update nodes distance in range
Step 9: Select the node within transmission range which is near to destination based on the threshold value If dist < threshold Threshold = dist Src = NextSrc Goto Step 5

End

4   Results

A. For EEG Signals
See Figs. 3, 4 and 5.
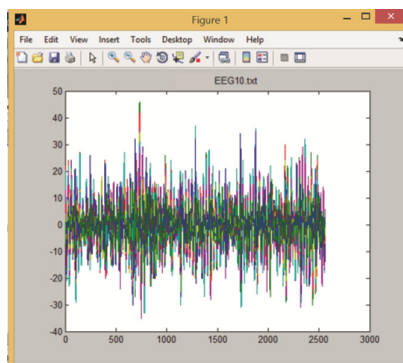
Fig. 3. Input EEG signal
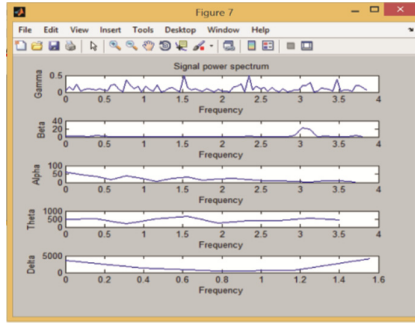
Fig. 4. Signal power spectrum
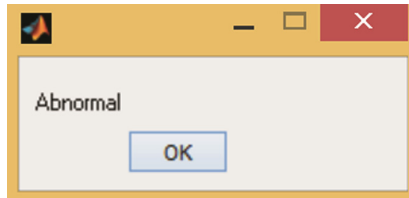
Fig. 5. Signal classification result

B. Routing Results:
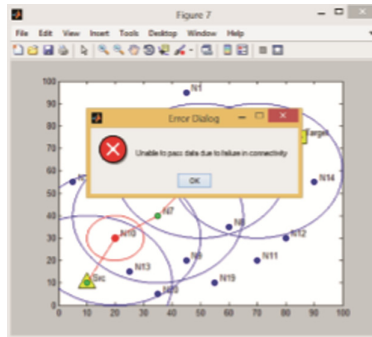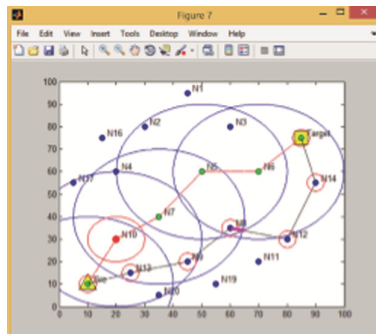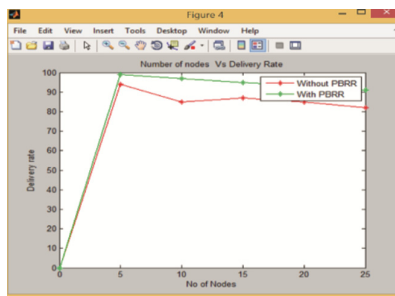See Figs. 6, 7, 8 and 9.
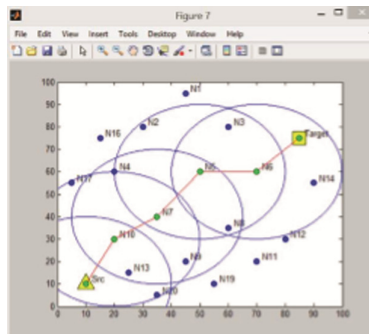
Fig. 6. Route establishment

Fig. 7. Handling of path failure

Fig. 8. Alternate path selection

Fig. 9. Performance of packet delivery

5 Conclusion

This work focuses on the categorization of EEG signal and sending those majority data to particular end utilizing proper directing plan. As EEG signals are non-stationary, the conventional system of frequency analysis is not exceedingly effective in diagnostic classification. In this paper, a procedure is suggested for the EEG signal classification,

it is based on WT. DWT is used to decompose EEG signal at resolution levels of the components of the EEG signal and to extract the percentage dissemination of energy features of the EEG signal at different determination levels. The FFNN classifies these extracted features to recognize the EEGs type according to the percentage distribution of energy features. The results shows that the proposed routing algorithm is able to handle the path failure efficiently and route the data to destination through alternate path. The packet deliver rate is good and the end-to-end delay is less.

References

1. Gonzalez, E., Peña, R., Vargas-Rosales, C., Avila, A., deCerio, D.P.-D.: Survey of WBSNs for pre-hospital assistance: trends to maximize the network lifetime and video transmission techniques. Sensors **15**, 11993–12021 (2015). https://doi.org/10.3390/s150511993

2. Vedavathi, B.S., Biradar, S., Hiremath, S.G., Thippeswamy, G.: Wavelet transform based neural network model to detect and characterize ECG and EEG signals simultaneously. In: IEEE International Advance Computing Conference Number #35547 Xplore Compliant, BMS College of Engineering Bangalore (2015). www.iacc2015.com. ISBN 978-1-4799-8047-5

3. Sharma, S., Vyas, A.L., Thakke, B., Mulvaney, D., Datta, S.: Wireless body area network for health monitoring. In: IEEE 4th International Conference on Biomedical Engineering and Informatics (BMEI) 978-1-4244-9352-4/11/$26.00 ©2011. IEEE (2011)

4. Murugappan, M., Rizon, M., Nagarajan, R., Yaacob, S., Zunaidi, I., Hazry, D.: EEG feature extraction for classifying emotions using FCM and FKM. Int. J. Comput. Commun. **1**, 21–25 (2007)

5. Jafari, R., Lotfian, R.: A low power wake-up circuitry based on dynamic time warping for body sensor networks. In: Proceedings of the 2011 International Conference on Body Sensor Network (BSN), Dallas, TX, USA, pp. 83–88 (2011)

6. Luo, X., Dong, M., Huang, Y.: On distributed fault-tolerant detection in wireless sensor networks. IEEE Trans. Comput. **55**, 58–70 (2006)

7. Kwong, J., Ramadass, Y.K., Verma, N., Chandrakasan, A.P.: A 65 nm sub-Vt microcontroller with integrated SRAM and switched capacitor DC-DC converter. IEEE J. Solid State Circ. **44**, 115–126 (2009)

8. Atallah, L., Lo, B., King, R., Yang, G.Z.: Sensor placement for activity detection using wearable accelerometers. In: Proceedings of the 2010 International Conference on Body Sensor Networks (BSN), Biopolis, Singapore, pp. 24–29 (2010)

9. Song, K.T., Wang, Y.Q.: Remote activity monitoring of the elderly using a two-axis accelerometer. In: Proceedings of the CACS Automatic Control Conference, Tainan, Taiwan (2005)

10. Preece, S.J., Goulermas, J.Y., Kenney, L.P.J., Howard, D., Meijer, K., Crompton, R.: Activity identification using body-mounted sensors—a review of classification techniques. Phys. Meas. **30**, 353–363 (2009)

11. Akin, A., Stephan, B., Mihai, M.P., Paluca, M.P., Paul, H.: Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: a survey. In: Proceedings of the 23rd International Conference on Architecture of Computing Systems (ARCS), Hannover, Germany, pp. 1–10 (2010)

12. Natarajan, A., Motani, M., de Silva, B., Yap, K., Chua, K.: Investigating network architectures for body sensor networks. In: Proceedings of the 1st ACM SIGMOBILE. In: international Workshop on Systems and Networking Support For Healthcare and Assisted Living Environments (HealthNet 2007), San Juan, Puerto Rico, pp. 19–24 (2007)

13. Tachtatzis, C., Graham, B., Tracey, D., Timmons, N.F., Morrison, J.: On-body to on-body channel characterization. In: Proceedings of 2011 IEEE Sensors Conference, Limerick, Ireland, pp. 908–911 (2011)

14. Ullah, S., Shen, B., Riazul, I.S., Khan, P., Saleem, S., Kwak, K.: A study of MAC protocols for WBANs. Sensors **10**, 128–145 (2010)

15. Seo, S.H., Gopalan, S.A., Chun, S.M., Seok, K.J., Nah, J.W., Park, J.T.: An energy-efficient configuration management for multi-hop wireless body area networks. In: Proceedings of 3rd IEEE International Conference on Broadband Network and Multimedia Technology, Beijing, China, pp. 1235–1239 (2010)

16. Latré, B., De Poorter, E., Moerman, I., Demeester, P.: MOFBAN: a lightweight modular framework for body area networks. In: Kuo, T.W., Sha, E., Guo, M., Yang, L.T., Shao, Z. (eds.) EUC 2007. LNCS, vol. 4808, pp. 610–622. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-77092-3_53

17. Yang, J.Y., Wang, J.S., Chen, Y.P.: Using acceleration measurements for activity recognition: an effective learning algorithm for constructing neural classifier. Pattern Recognit. Lett. **29**, 2213–2220 (2008)

18. Movassaghi, S., Abolhasan, M., Lipman, J., Smith, D., Jamalipour, A.: Wireless body area networks: a survey. IEEE Trans. **6**(3) (2014). third quarter

Automata Approach to Reduce Power Consumption in Smart Grid Cloud Data Center

J. Usha , S. R. Jayasimha$^{(\boxtimes)}$, and S. G. Srivani

R V College of Engineering, Bengaluru, India
{ushaj, jayasimhasr, srivanisg}@rvce.edu.in

Abstract. Today, in the cloud data center reduction in power consumption is one of the most difficult tasks in the Information and Communication Technology. In the cloud, data center includes reservoirs which process the power to meet user needs while computing. Presently. Data centers have become increasingly popular by user acceptance. The challenging task is to reduce high energy consumption in data center servers in the cloud computing environment which results in increased cost of maintenance and CO_2 emission. In this paper, automata approach has been proposed and then compared with various approaches which optimize the hardware, software components reducing the power utilization in the data center servers using virtualization technology. Dynamic voltage and frequency scaling approach based on the learning automata are used to reach a compromise and decrease the energy consumption in the data center. Power can be saved by shutting down the servers during idle period and consumption is maximum during peak period.

Keywords: Learning automata · Virtual machine consolidation
Threshold · MMT · DVFS

1 Introduction

The data centers provide storage and services [1] in cloud computing environment. Amazon, IBM, Google, Microsoft, workday, OMNITURE, CITRIX and many other companies have developed data centers to manage, share and improve the performance in computing. Energy efficiency is the main issue in the huge data centers due to huge cost and ecological concern with respect to large amount of Carbon Dioxide emissions [2]. According to the recent survey of CISCO, traffic of cloud is 14-folds in 2016 and will increase 17–18 Folds by 2020 [3].

Consumption of energy in the data center accounted for 1.5% of the total energy consumed in US [4]. In future, 2018 to 2020, the expected consumption of energy in data center through virtual machines may increase to 2.7% of total energy consumption in the cloud [5]. Face book, Google+, YouTube consumes 10.52%, 7.74% and 3.27% of the total energy utilization of the IT industry. The maximum energy usage of data centers is equivalence to 25,000 households [6, 7]. Due to increase in the energy and temperature in the data center, the reliability and durability of various hardware

resources will be reduced and will cause the carbon dioxide (CO_2) emission resulting in environmental issue [8]. The reduction of power consumption in the data center CPU power, storage and networking components need to be focused [9, 10]. The various energy consumption equipment devices [11] like server, switches and other facilities are compared in the Fig. 1.
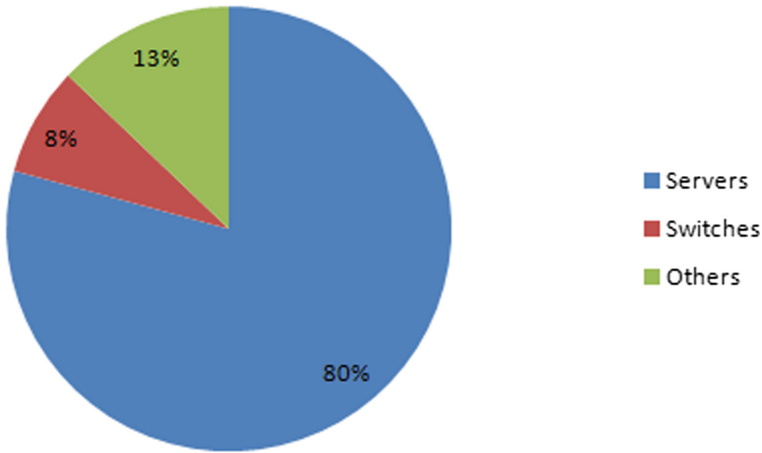

Fig. 1. Energy used by various devices

Other hardware equipments in the data center like CPU, memory and Disk devices with storage capacity consumes 58%, 28% and 14% respectively is [12] shown in the Fig. 2.
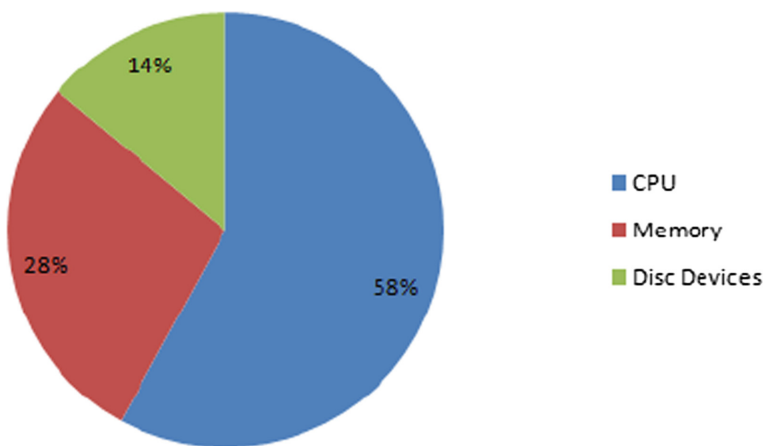

Fig. 2. Power consumption by various resources

In the cloud data center idle servers consumes more wastage of energy through various hardware resources because data centers are not fit to any power [13]. During 2 years survey Google scholars identified 20000 physical machines ran at 40 to 80% efficiency and the energy usage close to 100% [14].

Virtualization technology is the solution for the reduction of energy consumption in the data center. Virtualization enables multiple machines to locate a single Virtual Machine (VM) through VM and it can reduce the number of hardware resources requirement and can improve the productivity of the machine [15, 16]. The VM consolidation in the data center can be tested through forward, backward and none methods compared with different components and workload distribution with different time and achieved 76.9% to 97.8% energy efficiency [17].

With less number of Virtual Machines and Physical Machines (PM) human operator can manage the VM migration or movement is easy [18] where as if the VMs and PMs are increased it is very difficult to migrate physically. At this instance of time automation is necessary [19, 20]. For the automation requires the evaluation of VMs and PMs through Eq. 1 can be calculated using total number of physical machines considered for the experiment to the square i.e., double of Virtual Machine need to be selected to calculate the energy efficiency in the data center:

$$(Total\,Number\,of\,PMs)^{(Total\,Number\,of\,VMs)} \tag{1}$$

The objectives identified as M Physical Machines given with N Virtual Machines need to be placed and while mapping PMs on to VM the capacity of PMs should not exceeded and the parameters considered as Memory, CPU power, network, bandwidth dimensions.

2 Proposed Approach

Learning automata (LA) isthe best wayto requires multiple correlation coefficients in the data center. Random environment variables are good response in learning automata depicts in Fig. 3.
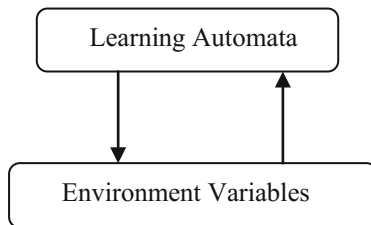
Fig. 3. Learning automata

Automata structures in various pieces are LA = {α, ß, P, Q, R}

$\alpha = \{\alpha_1, \alpha_2, \alpha_3, \cdots\cdots \alpha_n\}$ defined with set of actions
$\ss = \{\ss_1, \ss_2, \ss_3, \cdots\cdots \ss_m\}$ set of input
$P = \{p_1, p_2, p_3, \cdots\cdots p_r\}$ Action performance
$Q = P(r + 1) = Q[\alpha(n), \ss(n), p(n)]$
$R = \{R_1, R_2, R_3, \cdots\cdots R_n\}$ possibility of action

In the learning automata has chosen α_i then the environment will give favorable response. If $P_i(r)$ is increasing their performance then decreases other probabilities. With the better response of $P_i(r)$ decreases then the other probabilities are increase. All the above changes are the sum of $P_i(r)$ others will remain constant to equal.

(i) Favorable response defined as

$$P_i(r+1) = P_i(r) + \alpha[1 - P_i(r)]$$
$$P_j(r+1) = (1 - \lambda) P_j(r) \in jj \neq I \qquad (2)$$

(ii) Favorable response with action made as

$$P_i(r+1) = (1 - \lambda) P_i(r)$$
$$P_j(r+1) = \lambda/n\text{-}1 + (1 - \lambda) P_j(r) \in jj \neq I \qquad (3)$$

The Eq. 2 depicts that increase the number of VMs considered and decrease the probability status based on the workload and Eq. 3 depicts decrease the probability and increase the total number of VMs considered to identify the workload condition in the data center both Eqs. 2 and 3 used to calculate λ penalty is discussed in further equations.

3   Multiple Correlation Coefficients

The multiple correlation coefficients are used for the dependent variables in the analysis of multiple regression model. The variables value predicts dependent variable on real time entity. The dependent variables depicts **x** results for the variables: **y_1, y_2, y_3, y_4**. The regression relation observes **i** from **x** variable to **Q** as shown in Eq. 4.

$$X_i = \ss_0 + \ss_1 y_{1i} + \cdots + \ss_q y_{rp} + e_i \qquad (4)$$

In Eq. 4 X_i are the i values dependent variable. Q is the number of predictions in sum j are the coefficients where as j = 0, 1,···, q, e_i are the observed values from i different cases

Suppose y is the matrix of N * (Q + 1), whereas always first column considered with 1, and the various data considered from the independent variables as x vector, which is 1 * N form the real dependent variable.

The matrix considered with

$$X = x_1, x_2, x_3 \cdots x_n \tag{5}$$

and

$$y = 1 \cdots y_{ip}, 1 \cdots y_{np} \tag{6}$$

Equations 5 and 6 defined as the variables and covariant in column matrix.

The dependent variable x is shown as $\hat{x}$ and which is predicted the vector values shown in Eq. (7)

$$\hat{x} = y.a$$
$$a = (yy^m)^{-1}y_x^m \tag{7}$$

whereas y^T shows the transport of y matrix the accessed quality calculated prediction by multiple correlation coefficient which can be calculated through linear relationship with its values as Zero and One shown in Eq. 8.

$$R^2_{x.1...Q} = \left[\frac{Covarience(x_i, \hat{x})}{\sqrt{var(x_i)var(\hat{x})}} \right] \tag{8}$$

The covariance can be calculated in Eq. 8 and correlation between the variables shown in the Eq. 9.

$$R^2_{X.1...Q} = \left[\frac{\sum_1^n (x_i - m_x)^2 (\hat{x}_1 - n_{\hat{x}})^2}{\sum_1^n (x_i - m_x)^2 \sum_1^n (\hat{x}_i - m_{\hat{x}})^2} \right] \tag{9}$$

In the Eq. (9) the correlation between the covariance can be calculated between VM and the collection VMs selected different physical machines. The matrix y depicts the productivity of VMs in the cloud all the column matrix includes M to N productivity with history of VMselected host in the server.

4   Proposed Method

The experiment is carried out in the CloudSim toolkit. The CloudSim tool is most environmental support toolkit for the cloud platform in Infrastructure as a Service (IaaS). CloudSim facilitate and reproducing energy efficient support in virtualized data center in Cloud Computing. Time varying with memory, size and bandwidth is considered with number of physical host and virtual machine considered in the data center. The Simulation is energy aware power computing resources. The modified best fit decreasing method based on dynamic migration scheduling technique is considered and comparing with dynamic voltage frequency scaling by selecting with different number of VMs with Thresholds frequency in VM selection policy in CloudSim. The detailed approach is discussed.

The automation of energy efficiency is considered by the number of physical host machine equals to the number of action taken is same. Only one action should taken one at a time for one host. The automation vector probability function array can be called as A[i]. And its single values equals to one divides number of host (1/N). The amount of probability selecting by the host ID is equal to i.

Each and every host selects the automated Virtual Machines in the data center and the selection process will not be overloaded then the VMs host is obtained. The power estimation can be calculated between the CPU and energy consumption in the memory devices.

The selection of VM power is calculated through estimation linear relationship. The energy identified in the current VM is estimated and the VM allocation through correlation among the VMs in the other hosts to be correlated through multiple correlation coefficient and reward to the dedicated automation.

The process repeated till the number of host and high probability will be considered at the destination host the experiment repeated for different RAM size of various VMs can be considered with constant memory size is evaluated.

5   Performance Evaluation

Virtual Machine consolidation and evaluated by considering 6 VM for simulation experiment and which are shown in the Table 1, the different types of Personal Machine are shown in Table 2, and the VMs runs on the various days results shown in Table 3.

Table 1. Types of VMs

Type	Size	BW	RAM	MIPS
1	3.0 GB	200 Mbit/s	980	3300
2	3.0 GB	200 Mbit/s	220	2500
3	3.0 GB	200 Mbit/s	225	1200
4	3.0 GB	200 Mbit/s	843	650
5	3.0 GB	200 Mbit/s	1200	700
6	3.0 GB	200 Mbit/s	1450	850

Table 1 by considering 6 types of different RAM is verified with the constant bandwidth of 200 Mbits/s and Size of 3.0 GB memory.

Table 2. Types of PMs

Type	Storage	BW	RAM	MIPS
1	2 GB	2 Gbit/s	5020	2430
2	2 GB	2 Gbit/s	5020	4220
3	2 GB	2 Gbit/s	5020	3300

Table 2 shows three different types of workload condition analyzed with the same bandwidth, memory and RAM.

Table 3. VMs run on various days

Date	Number of VMs
02052017	1800
04052017	956
08052017	1076
20052017	1715
24052017	1600
02062017	1860
04062017	1944
08062017	1560
20062017	1423
24062017	1640

Table 3 shows different days verified with the multiple number of VMs. VM selection policy used a Thresholds (THR) to identify overloaded hosts and identified as 12.0 error confidence level. And selection of virtual machine used by Minimum Migration Time (MMT) policy. The results of CloudSim Simulation tool were are compared with dynamic voltage and frequency scaling (DVFS) and the method used to model based dynamic migration scheduling. The Figs. 4, 5 and 6 show the simulation results.
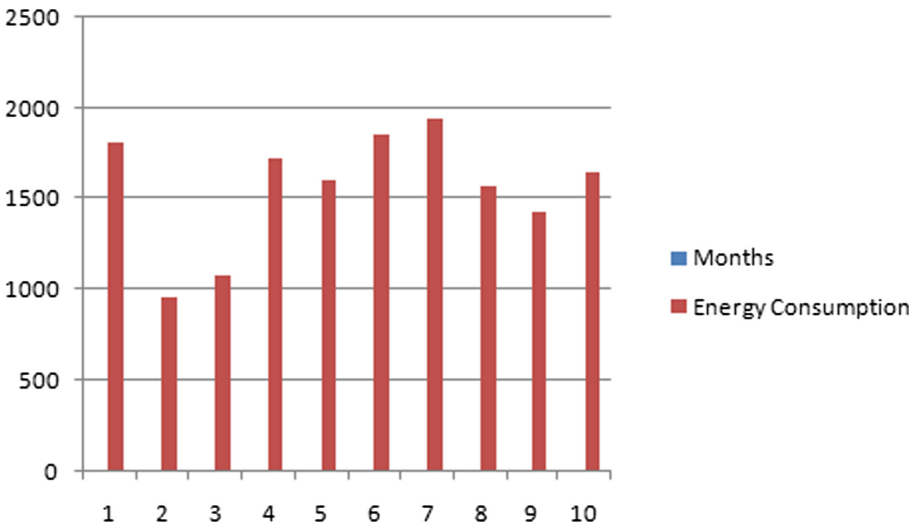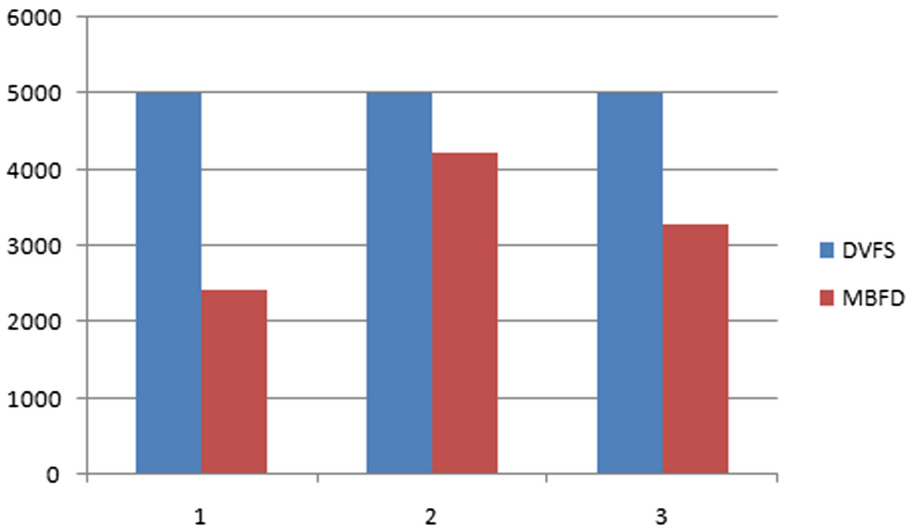

Fig. 4. Energy consumption analysis

Figure 4 shows total number of months considered in x axis, total number of VMs considered in y axis and it shows the Thresholds frequency in VM selection policy.
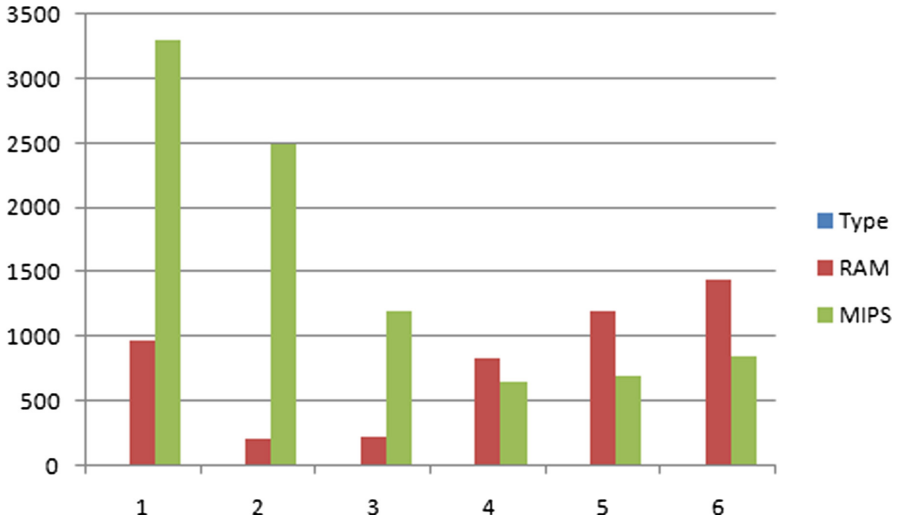
Fig. 5. Average RAM v/s MIPS

Figure 5 shows amount of RAM v/s MIPS with the size of 3.0 GB memory with the bandwidth of 200 Mbit/s is considered.

Fig. 6. Comparison of DVFS and MBFD method

Figure 6 depicts the comparison of dynamic voltage frequency scaling and modified best fit decreasing method based on dynamic migration scheduling technique.

6   Conclusion

The High energy consumption in the data center causes more CO_2 emission in the environment. Now a day's cloud computing energy efficiency and reduce the power consumption is the hot technology in the research. In this paper VM selection policy used by CloudSim tool to reduce power utilization in the data center. The power-aware energy efficiency is automated by considering the number of physical host machines and double the number of virtual machines considered with different workload conditions which includes memory usage and the bandwidth. The automataapproach, minimum energy and minimum correlation coefficient are the two parameters considered to find the efficiency in cloud. Multi correlation coefficient method is used to find the power aware energy efficiency in the cloud environment the learning automata is one of the best approach is used to find the efficiency in the cloud. our experimental result shows the comparison of DVFS and MBFD method and The Modified best-fit decreasing method is accurate to reduce the power consumption 20% to 25% less than the existing best-fit decreasing method by varying the parameters of memory, bandwidth and MIPS.

References

1. Dabbagh, M., Hamdaoui, B., Guizani, M., & Rayes, A.: Energy-efficient resource allocation and provisioning framework for cloud data centers (2015)
2. Cisco global cloud index: frocast and methodology. Cicso Inc., White Paper (2016)
3. Gandhi, A., Balter, H., Das, R., Lefurgy, C.: Optimal power allocation in server farms. In: Proceedings of the Eleventh International Joint Conference on Measurement and Modeling of Computer Systems. ACM, New York (2009)
4. Peer1 hosting site puts a survey on "Visualized: ring around the world of data center power usage". www.engadget.com
5. Kaplan, J., Forrest, W., Kindler, N.: Revolutionizing data center energy efficiency, Technical report McKinsey & Company (2010)
6. Brooks, D., Martonosi, M.: Dynamic thermal management for high-performance microprocessors. In: The Seventh International Symposium on High-Performance Computer Architecture, Article (CrossRef Link) (2001)
7. Ye, K., Huang, D., Jiang, X., Chen, H., Wu, S.: Virtual machine based energy-efficient data center architecture for cloud computing: a performance perspective. In: IEEE/ACM International Conference on Green Computing and Communications, on Cyber, Physical and Social Computing, Article (CrossRef Link) (2010)
8. Jayasimha, S.R., Usha, J., Srivani, S.G.: Minimizing power consumption and improve the quality of service in the data center, 9(November) (2016). http://doi.org/10.17485/ijst/2016/v9i43/104388
9. Kansal, A., Zhao, F., Liu, J., Kothari, N., Bhattacharya, A.: Virtual machine power metering and provisioning. In: Proceedings of the 1st ACM Symposium on Cloud Computing (2010)
10. Jayasimha, S.R., Usha, J., Srivani, S.G.: Analysis of power consumption under different workload conditions in the data center, (ICECS). IEEE (2016)
11. Jayasimha, S.R., Narasimha Prasad, D., Hamsa, K., Sumithra Devi, K.A.: Prevention of data from the data leakage in cloud computing. In: ISRASE, Firse International Conference on Recent Advances in Science and Engineering (ISRASE 2014) (2014)

12. Jayasimha, S.R., Hamsa, K.: A survey on efficient broadcast, geocasting and scalability recital in MANET'S, Icon RFW-2014 (2014)
13. Computing, S.: An Energy Efficient VM Allocation Approach for Data Centers. http://doi.org/10.1109/BigDataSecurity-HPSC-IDS.2016.62
14. Giacobbe, M., Celesti, A., Fazio, M., Villari, M., Puliafito, A.: An approach to reduce energy costs through virtual machine migrations in cloud federation (2015)
15. Goyal, A.: Bio inspired approach for load balancing to reduce energy consumption in cloud data center (2015)
16. Tripathi, A.: Storage cost efficient approach for cloud storage and intermediate data sets in clouds (2015)
17. Yang, T., Lee, Y.C., Zomaya, A.Y.: Collective energy-efficiency approach to data center networks planning, **7161**(c), 1–12 (2015). http://doi.org/10.1109/TCC.2015.2511732
18. Wang, S., Zhou, A.O., Hsu, C., Member, S., Xiao, X., Yang, F., Member, S.: Provision of data-intensive services through energy- and QoS-aware virtual machine placement in national cloud data centers, **4**(2) (2016)
19. Chisca, D.S., Casti, I., Barry, M.: On energy-and cooling-aware data centre workload management (2015). http://doi.org/10.1109/CCGrid.2015.141
20. Deguchi, T., Taniguchi, Y., Hasegawa, G., Nakamura, Y.: Impact of workload assignment on power consumption in software-defined data center infrastructure (2014)

Hybrid Optimization in Big Data:
Error Detection and Data Repairing
by Big Data Cleaning Using CSO-GSA

K. V. Rama Satish[(✉)] and N. P. Kavya

RNS Institute of Technology, Bengaluru, India
{ramasatish.k.v,n.p.kavya}@rnsit.ac.in

Abstract. Data cleaning is an important process in the history of data acquisition, data storage, data management and data analytics, and is still go through rapid development. In fact, cleaning of data is considered as a very important challenging task in the Big data era, due to the exponential growth of data in terms volume and variety of data in most of the applications. This paper focus to prove an accurate data extraction system in different ways of Data cleaning, i.e., error detection methods and data repairing algorithms. To achieve the accuracy of data extraction and improve the quality of data, this paper proposes a hybrid Cuckoo Search Optimization along with Gravitational Search algorithm (CSO-GSA) which is used to effectively detect the error from the data received by the source file and repairs the data before delivering it. Through the experiment on the MATLAB platform, it is exhibits the proposed approach to bringing down the time for error detection and correction in huge data sets with acceptable error detecting accuracy.

Keywords: Big data · Data cleaning · Data repairing · Error detection
Data quality · Hybrid Cuckoo Search Optimization
Gravitational Search Algorithm

1 Introduction

With the exponential utilization of electronic gadgets, day by day life is facing challenges of Big data. Big Data refers to an information set which gathers vast and complex information that is difficult to process utilizing customary applications [1]. Enormous information are described by three perspectives: (an) information are various, (b) information can't be arranged into normal databases, and (c) information are generated, captured, and handled quickly [2]. Because of its high operational and key potential, remarkably in creating business worth, "huge information" has as of late turned into the center of scholarly and corporate examination [3]. The surviving writing recognizes 'enormous information' as the 'following huge thing in advancement'; "the fourth worldview of science"; 'the following boondocks for development, rivalry, and profitability'; 'the following "administration unrest" and that 'huge information' is "acquiring an upheaval science and innovation" [4].

In the developing field of enormous information mining, administration and investigation, data extraction has been the greatest roadblock for activity data administrations and its applications [5]. Be that as it may, the investigation and information extraction process from enormous information turn out to be extremely troublesome errands for the vast majority of the traditional and propelled information mining and machine learning devices [6]. The happening to the Big data Age presents huge open doors and enormous difficulties for current society [7]. To meet the tightly coupled ongoing prerequisites for breaking down mass measures of information and over-hauling demands within fraction of milliseconds, an in-memory framework/database that maintains the information in the Random Access Memory every time is fundamental [8]. Database frameworks have been developing throughout the most recent couple of decades, primarily determined by advances in equipment, accessibility of a lot of information, gathering of information at an uncommon rate, rising applications thus on [9]. The scene of information administration frameworks is progressively divided taking into account application areas [10].

In spite of the fact that information cleaning has been a long standing issue, it has turned out to be basically vital again in view of the expanded enthusiasm for huge information and web information [11]. The greater part of the center of the work on huge information has been on the volume, speed, or assortment of the information; in any case, an imperative piece of making huge information valuable is to guarantee the veracity of the information [12]. This issue has gotten noteworthy consideration throughout the years in the conventional database writing, the cutting edge approaches miss the mark regarding a compelling answer for huge information and web information [13]. Traditional methods include outlier detection [14], noise removal [15], entity resolution and imputation [16].

In particular, exploiting so as to cut edge approaches attempt to clean information designs in the data, which they express as restrictive useful conditions [17]. In any case, these methodologies rely on upon the accessibility of a perfect information corpus or an outside reference table to learn information quality standards or examples before settling the blunders in the messy information. Frameworks, for example, Conquer rely on an arrangement of clean limitations gave by the client [18]. Such spotless corpora or limitations might be anything but difficult to set up in a firmly controlled undertaking environment yet are infeasible for web information and huge information [19]. One might attempt to gain information quality principles specifically from the uproarious information. Unfortunately, in any case, our exploratory assessment demonstrates that even little measures of commotion extremely debilitates the capacity to take in helpful imperatives from the information [20].

To improve the accurate process of data extraction in Big Data cleaning we propose a new hybrid CSO-GSA algorithm i.e. Cuckoo Search Optimization algorithm and Gravitational Search algorithm. The main contribution of this paper is: (1) to repair and detect the errors in the Big data (2) to optimize the Big data by a novel hybrid technique and to improve the quality of the data. The paper is organized as: 1. Some of the recent research works related to Big data optimization are highlighted as Sect. 2. The detailed explanation of our proposed Big data optimization method is presented in Sect. 3. Performance analysis of the experiment, their outcomes, ands comparison of the proposed method is specified as Sect. 4 and in the last section, the conclusion is written as Sect. 5.

2   Related Works

A few research works done in the past decades related to optimized Big Data cleaning is presented in this section.

Daneshmand *et al*. [21] have proposed an effective parallel, hybrid random/deterministic decomposition algorithm in that, in every loop, the variables subset of was updated, simultaneously minimizing a convex surrogate of the original nonconvex function.

Wu *et al*. [22] have proposed a HACE theorem, explaining the Big data features, and processing model with respect to Big data was proposed from the techniques of data mining. This model which is more data intensive, describes demand-driven aggregation derived out of various data lakes, data mining and data analysis, user interests and security considerations.

Wang *et al*. [23] have proposed an algorithm to improving dictionary knowledge using the Particle-Swarm-Optimization (PSO). In every iteration, special atoms were chosen within the dictionary and PSO is applied to enhance state of dictionary-learning model.

Zhang *et al*. [24] have studied various methods of minimizing Big data and geo-dispersed data, time and cost of uploading on the cloud and process them using a modified variant type of Map Reduce. Propose an "online-lazy-migration algorithm (OLM) and a randomized-fixed-horizon-control algorithm (RFHC)" for optimizing at any specified time interval, the choice of the data center for data aggregation and processing, as well as the data movement.

Zheng *et al.* [25] have studies different ways of integrating Big data analytics using network optimization from QoS. Different techniques in of data acquisition and analysis were deliberated keeping view of network optimization.

3   Optimization of Big Data Using Proposed Hybrid Techniques

Big data has hidden knowledge and patterns which, if extracted, can help to reduce production and maintenance costs, identify trends in businesses, predict the expectations of the consumers and help us live better. A Big data problem is not only about its size, but the generation frequency and rapid variety of changing trends of the data, heterogeneous data sources, different levels of errors, and other factors defines complexity of a Big data problem.

If a source file directly delivers the data without cleaning the data the receiver will receive an unexpected data. Many procedures have been proposed on optimization for both single and multi-objective problems. Cuckoo search is one of the optimization algorithm which can be used for selection of data from the Big data globally. However due to the complexity of Big data and to extract the exact data with better quality some hybridization techniques are required. By this we have analyzed that there is a vital requirement to improvise evolutionary optimization techniques like hybrid optimization techniques, that can solve such problems with millions of variables in the Big data.

To overcome the above issues in Big data in this appear we propose a new hybrid Cuckoo Search Optimization with Gravitational Search algorithm (CSO-GSA) which helps to effectively detect the error from the data received by the source file and repairs the data before delivering it. In this paper hybridization is done in the phase which combine the ability to find the global area of CSO using the local search capability of GSA for Data cleaning and optimal data extraction. The architecture of our proposed method is shown in the below Fig. 1.
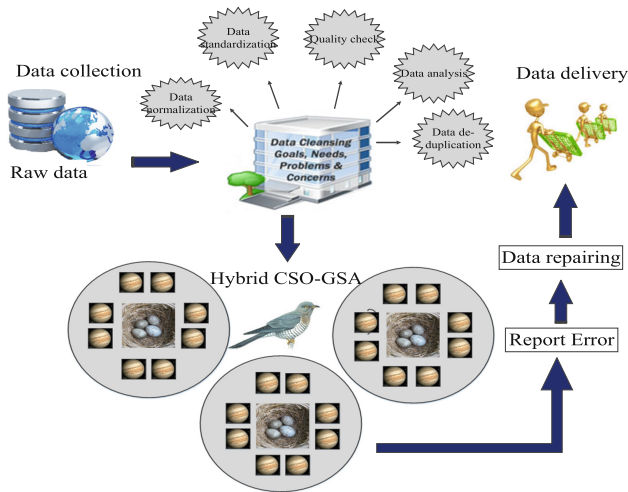
Fig. 1. Architecture of our proposed Big data cleaning method

From Fig. 1 it is evident that when the source file receives a request from the user to extract some data, searching of the exact data from the collection of raw data in the source will be done. To reduce the complexity of this process Data cleaning by hybrid CSO-GSA which eliminates the unwanted data's and repair the data is proposed in this paper. After Data cleaning the out coming data is imposed to a quality check and when the data reaches the required quality standard will be delivered to the user. To evaluate the performance of our proposed method it is necessary to design a system model with the unwavering problem which is described in the upcoming section.

3.1 Problem Identification

Data cleaning is a costly task, avoiding incomplete, irrelevant and invalid data to be fed to the system is certainly an important phase to minimize the data cleaning task. It holds a logical restructuring of the database schema and integrity constraints for the data entry applications. The problems of every sources are integrated as multiple sources also aggregated. The important talk of data cleaning is to identify multiple sources for overlapping data records implies to the same entity. Frequently, the information is only partly redundant and the sources may complement each other by providing additional information about an entity.

Consider 'O', set of objects with unique identity, representing a particular feature of a real world entity. The object has a different values at different instant of time 't' because of the current updating of database. We consider the lifespan of O as a sequence of transitions and we handle the dataset by adapting techniques to detect and repair errors. The O is associated with the two results such as exist and non-exist. Consider the given structured database as D = {d_1, d_2...d_n}. Where d represents the number of data's that represent the real world entities. We record observation of data from different sources at various intervals of time, comparing current observation with its previous, Mathematically it is denoted by P = {t0, t1,...tn} and the set of observations points I(D, ti), i→[0, n], the updates we get at time t_i, as a special case, I(D, t0) contains values D provided at t_0^2. Due to the updation of data at any time there is a chance to miss an overwritten data while observation. Thus frequent collection can often store huge information and advantage. In order to eliminate the above issues in this paper proposes a novel hybridization technique for Data cleaning. To evaluate the proposed methods' performance, it is necessary to define a system model with the problem constraints which is explained in the subsequent section.

3.2   System Model

The implementation of the hybridization technique with the basic strategy is explained in the next section. In our model, the database D contains fixed set of Table 1 {d_1, d_2... d_n} with O number of objects. Each table d is associated with the number of objects O. we use D(d, O) to denote the value of a given attribute and for the subset we define D (d, X) to represent the projection of d on attributes in X.

Table 1. Dataset of employee before data cleaning

Name	Gender	Age	Status	BOP
Maria	Female	24	Married	USA
George	NaN	32	Married	Alaska
Maria	Female	24	Married	USA
Kiran	Male	30	NaN	Russia
Ram	Male	25	Married	Texas
Smith	Male	27	NaN	UK
Shreya	Female	21	Married	Paris
Sam	NaN	23	Married	Washington
Rahul	Male	28	um	Goa
Catherine	Female	31	Married	Calcutta
Stancy	Fe	12	um	Haryana
Zeeria	Female	55	Married	Darjiling
Jack	Male	10	Married	France

The following are the possibility of constraints in our system model.

(a) Functional dependencies (FD): each FD has the form A[X]→A[Y] where X and Y are the subset attributes. The dataset is said to be satisfy the FDA[X]→A[Y] if foe every pair of tables d1, d2 belongs to A such that D(d1, X) = D(d2, Y) and also D (d1, Y) = D(d2, Y)

(b) Inclusion dependencies (IND): ID has the form A1[X] $\subseteq$ A2[Y] where X and Y are the subset attributes.. A dataset D is to satisfy the IND A1[X] $\subseteq$ A2[Y] iff for every tuple d1 belongs to A1 there exists d2 belongs to A2 such that D(d2, Y) = D(d1, X).

To clean the data it is necessary to check the quality of data. The factors of Data Quality is explained and derived in the below sections.

3.3   Core Factors of Big Data Quality

The cleaned Big data should be in high quality which requires to satisfy the following objectives such as: accuracy, completeness, deduplication, consistency and currency as whole we refer it as AC^{3}D measures. A source is said to be high-quality iff the true value of the object advances to the value; we capture these five conditions by the five measures respectively.

3.4   Data Repairing

Let us assume for a database D, the repair of is a database D' such that (i) tuples present in D are carried over to D' having id, with altered values of attribute, (ii) zero or more newly added tuples seen in D' but not in D and (iii) D' satisfies the constraint set C. An erratic database may be neither complete nor sound, and therefore the model enhances every value modifications and insertion of tuples. We change the tuples value in D rather merely removing them in so as to minimize the data loss.

3.4.1   Repairing FD Violation

We consider that the tuple d is repaired FD F = A[X]$\rightarrow$A[Y], if for all other tuples d'$\rightarrow$A either D'(d, A) not equals to D'(d', A) for some A$\rightarrow$X or for every B$\rightarrow$Y eq(d, B) = eq(d', B). if d is resolved; as no violation in D' but the converse is not some B$\rightarrow$Y, (d, B) might have the similar final value as (d', B) without (d.B) and (d',B) being in the similar class. Using as a pair of tuples we can explain a violation, we define the act of resolving a tuple d w.r.t. Fein terms of a set S of tuples from R. Here D includes d and all other tuples that agree with d on (target values of) attributes in X. The procedure to repair FD constraints is given below.

Steps in Repairing FD constraints

Input: *Dataset D with number of tuples with attribute set X.*

   1. *FD F=A[X]$\rightarrow$A[Y]*
   2. *Begin*
   3. *For each attribute A in Y do*
   4. *{*
   5. *eq(A)=eq(d,A)*
   6. *$\mathcal{E}$ =$\mathcal{E}$-{eq(d,A):dϵD}U{eq(A)}*
   7. *}*
   8. *end*

3.4.2 Repairing IND Constraints

For an IND I = A1[X] ⊆ A2[Y] a tuple d is said to be repaired with respect to I if there exist tuple d'→A2 such that (d, A) and (d;, B) are in the same class for every corresponding pair attributes A→X and B→Y. it is quiet manageable to notice the in contrast to FDs a tuple repaired w.r.t an IND will become resolved a fact we refer to as the stable property of IND resolution. The steps in Repairing IND constraints is given below.

Steps in Repairing IND constraints

Input: Dataset D with number of tuples with attribute set X and Y.

1. IND I=A1[X] ⊆ A2[Y]
2. Begin
3. If(target=new) then
4. {
5.   d'=new null tuple in A2 with zero weight
6. $\mathcal{E} = \mathcal{E}$ U{{d',A}:A{→attr(A2)}
7. }
8. For each attribute A in X and corresponding b in Y do
9. $\mathcal{E} = \mathcal{E}$ -eq(d,A), eq(d',B)})U(eq(d,A)Ueq(d',B)};
10. End

3.5 Optimization of Big Data

In this section before explaining the hybridization of our proposed novel CSO-GSA with the implementation and consequence of our proposed hybridization technique we explain the primary phases of CSO and GSA algorithm to make better to understand our proposed strategy.

3.5.1 Cuckoo Search Optimization Algorithm (CSOA)

Cuckoo Search Optimization is an "evolutionary algorithm", designed on the egg lying habitat of cuckoos. The basic strategy behind this optimization algorithm is the selection of best nests of the host birds to lay down the eggs of cuckoos. Initially the cuckoos lay some eggs along with the host bird's egg. The eggs which are very identical to the eggs of host bird have the only opportunity to survive for growth and the other eggs that are identified are destroyed. The developing eggs indicate the fittingness of the area in the nests. The more number of eggs survive in the particular region, the profit is gained in that region. So the position of eggs pattern in which more eggs persist will be the pattern order that CSOA is about to optimize. The cuckoo starts to lay eggs in rapid order at nests with in its eggs laying region. The egg laying continues until the optimum value (best position and maximum gain) is achieved.

The following rules explain the cuckoo search algorithm:

- A cuckoo randomly chooses a nest and lays eggs one by one at a time.
- The nest which carries the eggs to next generation with high quality is selected as best nest.

- The number of nests available for laying eggs are assumed as static and the host can investigate to find an alien egg with the probability $A_e \rightarrow [0.1]$.

Based on the above rules the Pseudocode for the cuckoo search algorithm is summarized below.

Pseudocode for CSOA

Initialize the population of N host nests, Get a cuckoo (say i) randomly, and choose a host among N (say j) randomly

1. *Evaluate the fitness function.*
2. *If(Fi > Fj):*
3. *Record the new solution by replacing the value j.*
4. *End_if*
5. *Replace Ae from N nests, Save the best solutions, Rank them and find the optimum(Current) solution.*
6. *End.*

To achieve maximization of the problem, the quality or fitness of a solution depends on the objective function. When producing new solutions, $s_i^{(t+1)}$ for, say cuckoo i, a L´evy flight is performed

$$s_i^{(t+1)} = s_i^{(t)} + \beta \oplus le'vy(\lambda) \tag{6}$$

Where the step size $\beta > 0$, relevant to the measures of the problem. In majority, we consider $\beta = O(1)$. The $\oplus$ symbol indicates the entry-wise multiplications with L´evy (λ) provide a random walk for random steps values picked from a L´evy distribution.

$$Le'vy \sim u = t^{-\lambda}, \ (1 < \lambda \leq 3) \tag{7}$$

Equation (7) tends the mean and variance value to infinity. The immediate iteration of a cuckoo generates a random walk which follows a power-law step length distribution with a heavy tail.

3.5.2 Gravitational Search Algorithm (GSA)

The Gravitational Search algorithm (GSA) is another latest stochastic algorithm for searching which also situated in the light of the mass interactions and law of gravity. The algorithm runs with number of searching objects that interact with each other through the gravity force. Each object indicates a complete solution or a part of it to the problem. All the objects work under the principal of gravitational force, attracting each other and this force is the main causes for driving all objects towards heavier mass objects. The optimal solution to the problem can be gained using higher fitness function of these heavier masses. As these objects move gradually, the lighter objects contributes the worst case. In GSA, every object mass has attributes such as inertial mass (Mii), active gravitational mass (Mai), passive gravitational mass (Mpi) and its position. The initial position of the object mass is same as the solution of the problem and the inertial masses and gravitational mass are directly proportional to the fitness function.

The algorithm stats with the initial position of a system, described as M masses i.e. dimension of the search space.

$$P_i = (P_i^1 \ldots P_i^d \ldots P_i^m) \text{ for } i = 1, 2, \ldots, M \tag{8}$$

Where M indicates the space-dimension and P_i^d indicates the position in the dth dimension of the ith agent.

Initially, the random definition of the solutions agents are assumed and the same respectively to the Newton's law of Gravity, a gravitational force among masses M_j acts M_i at an interval t is indicated as follows:

$$F_{ij}^d(t) = G(t) \cdot M_i(t) \times M_j(t) \cdot [R_{ij}(t) + \varepsilon]^{-1}(p_j^d(t) - p_i^d(t)) \tag{9}$$

where M_i, M_j indicates the mass of the objects i and j, G(t) depicts the gravitational constant at a specific time t, ε is a small constant and $D_{ij}(t)$ is the Euclidian distance between objects i and j defined as follows:

$$D_{ij}(t) = ||P_i(t), P_j(t)|| \tag{10}$$

Calculation of the gravitational constant G(t) is based on:

$$G(t) = G_0 \times \exp(-\alpha \times iter/\mathrm{max}iter) \tag{11}$$

Where G_0 *and* α are initial value and descending coefficient respectively, *iter* is the present loop, and *maxiter* is maximum limit of the loop. The overall force acting on the ith agent is calculated as follows:

$$F_i^d(t) = \sum_{j \in kbest, j \neq i}^{M} rand(j) \cdot F_{ij}^d(t) \tag{12}$$

where rand (j) is a random number between [0, 1] and kbest is the set of agents with best fitness value and biggest mass (first K). In order to calculate the acceleration $a_i^d(t)$ of the ith agent, in the dth dimension, at t time, law of motion is used. In accordance with this law, acceleration $a_i^d(t)$ is proportional to the force acting on i^{th} agent, and inversely proportional to the mass of the agent.

$$a_i^d(t) = F_i^d(t) \cdot [M_{ii}(t)]^{-1} \tag{13}$$

The strategy of search on Eq. (13) can be defined to find the upcoming next position of an agent and its velocity. The velocity of an agent $v_i^d(t + 1)$ is indicated as sum of its velocity and its current acceleration. So, the next position and velocity can be computed as follows:

$$v_i^d(t + 1) = rand(i)v_i^d(t) + a_i^d(t) \tag{14}$$

$$P_i^d(t + 1) = P_i^d(t) + v_i^d(t + 1) \tag{15}$$

where $v_i^d(t)$ and $P_i^d(t)$ represents the velocity and position of the agent i, at t time, in d dimension and rand (i) is a random number between 0 and 1 used to randomized feature to the search. Initially, the gravitational constant, G will be decreased respectively to control the search accuracy with accordance with time. The functioning of GSA is summarized by the Pseudocode shown below.

Pseudocode for GSA

---

1. *Produce initial population.*
2.   *Calculate each objects' fitness.*
3.   *Replace/update worst of population and Gbest.*
4.   *Compute space dimension M and acceleration a of every object*
5.   *Update position of each object and velocity of each object*
6. *If stopping criteria get satisfied*
7.   *Return the best solution*
8. *else*
9.   *Return to step 2*
10. *end*

3.5.3 Hybridization of CSO

From the above sections we acquire a basic knowledge about the working of GSA and CSO algorithms. The novel hybridization emerged from both CSO-GSA is now explained in this section from the basic knowledge obtained from the sections. From the above knowledge we gained that the CSO finds the global area to lays the eggs and to find the best solution (best area) which makes the egg to grow and mature. The basic idea behind the hybridization of CSO-GSA is the determination of the global area of CSO is done by the local search capability of GSA. We hybridize the functions of both CSO with GSA using a low level co-evolutionary heterogeneous hybrid. Since both are different type of algorithms run in parallel and one depends another while running to produce the final results. In CSO-GSA, there are number of data's present to complete users request these *Dn* are stored as 'n' initially. The objective function corresponds to *Dn*, the population of data is derived as *Dn*. The population generation of data is influenced from Lëvy Flight. The Lëvy Flight process runs the cuckoo until the data extraction becomes maximum related data is generated. After populating the data then the cuckoo checks for Quality and Fitness D_{fit}. The selected data's closely related to the expected data and the remaining data is discarded. As a result of hybridizing the above two algorithms Eq. (6) can be updated as follows:

$$S_i(t + 1) = S_i(t) + rand \times v_i^d(t) + a_i^d(t) \tag{16}$$

The fitness function of our proposed hybrid CSO-GSA is given by the following equation expressed as

$$D_{fit} = D_{accuracy} + D_{deduplication} + D_{consistency} + D_{currency} + D_{completeness} \qquad (17)$$

Where the values to calculate the above expression can be used from the Eqs. (1), (2), (3), (4) and (5). The fitness function is calculated for number of iteration after Repairing the data based on the constraints defined in the paper and when the D_{fit} value satisfies the Data Quality standard the iteration get stopped and the resulting will produce the complete cleaned data set.

In CSO-GSA initially all the objects are randomly assumed. Each object represents a solution. Then the gravitational force, gravitational_constant and resultant forces are calculated using Eqs. (9), (11) and (12). The accelerations of objects are calculated using Eq. (13). Then update the best solution values after each iteration. Determine the velocity of all the objects after updating the best solutions and accelerations using Eq. (14). Finally, the positions of objects are calculated using Eq. (16). This updation process repeats until the stopping criteria is met. The process flow of proposed Hybrid CSO-GSA is shown in the Fig. 2.
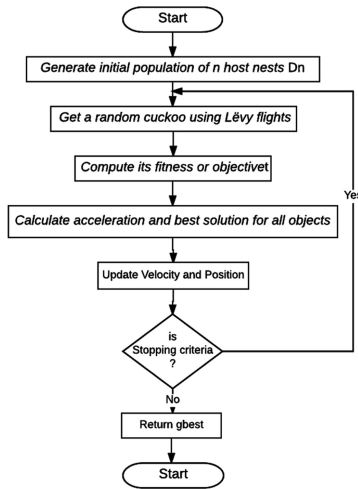
Fig. 2. Steps in hybrid CSO-GSA

4   Results and Discussion

This section shows the proposed Hybrid Optimization in Big data analysis mainly the Error Detection, Data Repairing and the results obtained by them.

4.1   Simulation Results

For validation of our method, datasets of employee details are used, as shown in Table 1. These dataset contains some Big data complexities such as Consistency,

Currency, Deduplication and Completeness. The "NaN" term in the dataset represents the data consistency, the errors such as um, fe are said to be data completeness. In the dataset there is two employees namely Maria this type of error is known as deduplication. And sometimes the employee's details are not updated in the dataset such as Jack with age 10. These type of errors are called as data currency. Our proposed algorithm is implemented to avoid these complexities in big data.

The above dataset contains some complexities which have to be eliminated. The data duplication error is completely removed from the big data by using our hybrid optimization. Our proposed algorithm performs well on these complexities and generates better accuracy and the resulting dataset of our proposed data cleaning method is shown in the Table 2 below.

Table 2. Dataset of employee After Big Data cleaning

Name	Gender	Age	Status	BOP
Maria	Female	24	Married	USA
George	Male	32	Married	Alaska
Kiran	Male	30	Married	Russia
Ram	Male	25	Married	Texas
Smith	Male	27	Married	UK
Shreya	Female	21	Married	Paris
Sam	Male	23	Married	Washington
Rahul	Male	28	Unmarried	Goa
Catherine	Female	31	Married	Calcutta
Stancy	Female	12	Unmarried	Haryana
Zeeria	Female	55	Married	Darjiling
Jack	Male	25	Married	France

The above table shows the cleaned dataset by using CSO-GA hybrid optimization that eliminates deduplication, consistency, completeness, currency and finally contributes the best accuracy. The performance of our proposed hybrid optimization is evaluated using the following parameters.

4.2 Performance Evaluation

Some criteria were employed in accordance to the recommendations to assess performance of error detection such as accuracy, precision, Recall and F-measure.

4.2.1 Precision
The Precision can be defined as follows

$$Precision(p) = \frac{Sum\ of\ relevant\ tuples\ \det ected\ from\ the\ database}{Total\ sum\ of\ tuples\ \det ected} * 100 \quad (18)$$

4.2.2 Recall

The recall can be defined as follows

$$\mathrm{Re}call(r) = \frac{Sum\ of\ relevant\ tuples\ \det ected\ from\ the\ database}{Total\ sum\ of\ relevant\ tuples\ in\ the\ database} * 100 \qquad (19)$$

4.2.3 Accuracy

The data accuracy can be defined by the following equation

$$Accuracy = \frac{\mathrm{Pr}\,ecison + \mathrm{Re}call}{2} * 100 \qquad (20)$$

4.2.4 F-Measure

The F-measure can be defined as follows

$$F - measure = \frac{2 * p * r}{(p + r)} \qquad (21)$$

In the above dataset we make 5 queries to check the efficiency of our proposed data cleaning method. The performance evaluation of our proposed method based on the queries is shown in below Table 3.

Table 3. Performance evaluation of CSO-GSA method (proposed method)

Query No.	Precision	Recall	Accuracy (%)	F-measure
Q1	0.9375	1	96.8750	1.8750
Q2	0.94	0.6	97	1.2
Q3	0.9667	0.5	98.333	1
Q4	0.9833	0.25	99.1667	0.50
Q5	0.9967	0.1667	99.8333	0.3333

From the table it is evident that our proposed framework generates tremendous accuracy of 98.24% in average. The graphical representation of the performance of our proposed method is shown in Fig. 3.

4.3 Comparison and Discussion

The Big data may be analyzed and the errors are detected using several techniques. Our CSO-GSA (proposed method) uses CSO-GA hybrid optimization to detect the errors from the data. The comparison between our proposed methods with existing technique i.e. Data cleaning without optimization is shown in Figs. 4, 5, 6 and 7.

The above figures compare several of the proposed method with existing methods i.e. Data cleaning without optimization. The deduplication values should be less for an
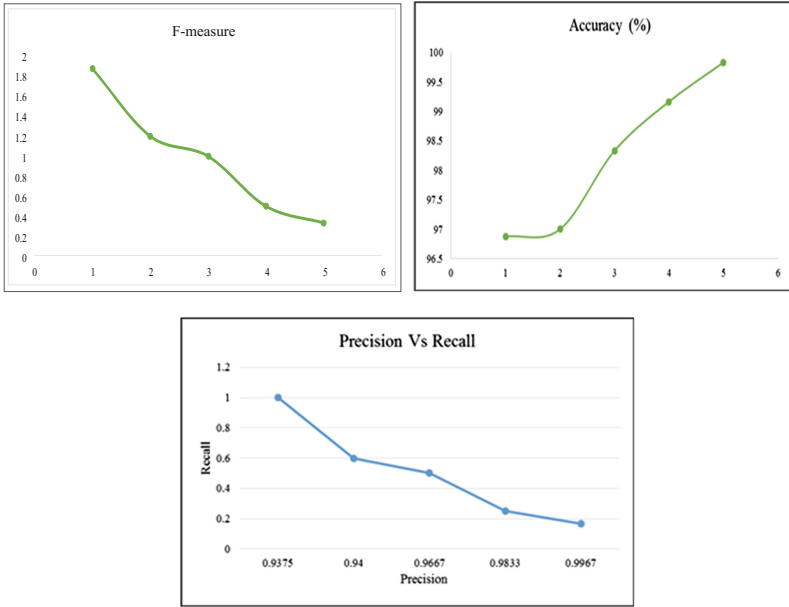
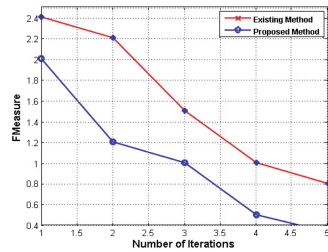**Fig. 3.** Performance of our proposed method
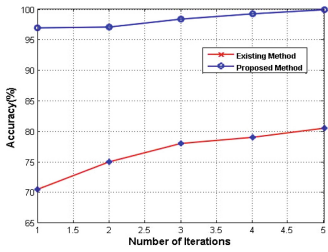
Fig. 4. Accuracy comparison
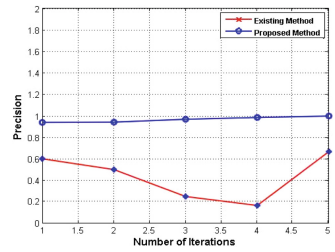
Fig. 5. Precision comparison
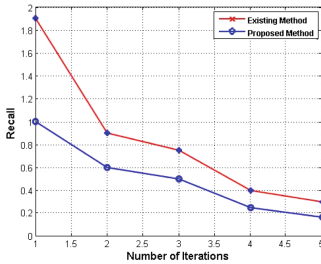
Fig. 6. Recall comparison

Fig. 7. FM measure comparison

efficient system and from the Figs. 4 to 7 indicates that the proposed model is an efficient model than the existing ones. From Fig. 4 we can conclude that our proposed hybrid optimization gives us valuable accuracy as 98.24% which is greatly higher than the existing method.

5   Conclusion

In this paper we have proposed a new method of Big data analysis of cleaning with high accuracy. For extracting some data the source file receives a request from the user, searching of the exact data from the collection of raw data in the source will be done. Our hybrid CSO-GSA is utilized to reduce the complexity of this process Data cleaning which eliminates the unwanted data's and repair the data. After Data cleaning the out coming data is imposed to a quality check and when the data reaches the required quality standard will be delivered to the user. The proposed system is evaluated in terms of various parameters and the experimental results and comparison shows that our proposed method is efficient and accurate in detecting the errors than the existing methods.

References

1. Hashem, I.A.T., Yaqoob, I., Anuar, N.B., Mokhtar, S., Gani, A., Khan, S.U.: The rise of "Big data" on cloud computing: review and open research issues. Inf. Syst. **47**, 98–115 (2015)
2. Wang, G., Gunasekaran, A., Ngai, E.W.T., Papadopoulos, T.: Big data analytics in logistics and supply chain management: certain Investigations for research and applications. Int. J. Prod. Econ. **176**, 98–110 (2016)
3. Assunção, M.D., Calheiros, R.N., Bianchi, S., Netto, M.A.S., Buyya, R.: Big data computing and clouds: trends and future directions. J. Parallel Distrib. Comput. **79**, 3–15 (2015)
4. Zhang, Y., Zhang, G., Chen, H., Porter, A.L., Zhu, D., Lu, J.: Topic analysis and forecasting for science, technology and innovation: methodology with a case study focusing on Big data research. Technol. Forecast. Soc. Change **105**, 179–191 (2016)
5. Zhang, H., Chen, G., Ooi, B.C., Tan, K.-L., Zhang, M.: In-memory Big data management and processing: a survey. IEEE Trans. Knowl. Data Eng. **27**(7), 1920–1948 (2015)
6. Jin, X., Wah, B.W., Cheng, X., Wang, Y.: Significance and challenges of Big data research. Big Data Res. **2**(2), 59–64 (2015)
7. Fong, S., Wong, R., Vasilakos, A.: Accelerated PSO swarm search feature selection for data stream mining Big data (2015)
8. Wu, D., Zhu, L., Xiwei, X., Sakr, S., Sun, D., Qinghua, L.: building pipelines for heterogeneous execution environments for Big data processing. IEEE Softw. **33**(2), 60–67 (2016)
9. Wamba, S.F., Akter, S., Edwards, A., Chopin, G., Gnanzou, D.: How 'Big data'can make Big impact: findings from a systematic review and a longitudinal case study. Int. J. Prod. Econ. **165**, 234–246 (2015)

10. Tan, K.H., Zhan, Y., Ji, G., Ye, F., Chang, C.: Harvesting Big data to enhance supply chain innovation capabilities: an analytic infrastructure based on deduction graph. Int. J. Prod. Econ. **165**, 223–233 (2015)
11. Fan, C., Xiao, F., Madsen, H., Wang, D.: Temporal knowledge discovery in Big BAS data for building energy management. Energy Build. **109**, 75–89 (2015)
12. Dong, H., Wu, M., Ding, X., Chu, L., Jia, L., Qin, Y., Zhou, X.: Traffic zone division based on Big data from mobile phone base stations. Trans. Res. Part C: Emerg. Technol. **58**, 278–291 (2015)
13. Zhou, K., Chao, F., Yang, S.: Big data driven smart energy management: from Big data to Big insights. Renew. Sustain. Energy Rev. **56**, 215–225 (2015)
14. Triguero, I., Peralta, D., Bacardit, J., García, S., Herrera, F.: MRPR: a MapReduce solution for prototype reduction in Big data classification. Neurocomputing **150**, 331–345 (2015)
15. Suresh, S.: Big data and predictive analytics: applications in the care of children. Pediatr. Clin. N. Am. **63**(2), 357–366 (2016)
16. Pääkkönen, P., Pakkala, D.: Reference architecture and classification of technologies, products and services for Big data systems. Big data Res. **2**(4), 166–186 (2016)
17. Wang, Y., Kung, L., Byrd, T.A.: Big data analytics: understanding its capabilities and potential benefits for healthcare organizations. Technol. Forecast. Soc. Change (2016)
18. Zhang, Y., Qiu, M., Tsai, C.W., Hassan, M.M., Alamri, A.: Health-CPS: healthcare cyber-physical system assisted by cloud and Big data (2015)
19. Zhong, R.Y., Huang, G.Q., Lan, S., Dai, Q.Y., Chen, X., Zhang, T.: A Big data approach for logistics trajectory discovery from RFID-enabled production data. Int. J. Prod. Econ. **165**, 260–272 (2015)
20. D'Oca, S., Hong, T.: Occupancy schedules learning process through a data mining framework. Energy Build. **88**, 395–408 (2015)
21. Daneshmand, A., et al.: Hybrid random/deterministic parallel algorithms for convex and nonconvex Big data optimization. IEEE Trans. Sig. Process. **63**(15), 3914–3929 (2015)
22. Wu, X., Zhu, X., Gong-Qing, W., Ding, W.: Data mining with Big data. IEEE Trans. Knowl. Data Eng. **26**(1), 97–107 (2014)
23. Wang, L., Geng, H., Liu, P., Ke, L., Kolodziej, J., Ranjan, R., Zomaya, A.Y.: Particle swarm optimization based dictionary learning for remote sensing Big data. Knowl.-Based Syst. **79**, 43–50 (2015)
24. Zhang, L., Chuan, W., Li, Z., Guo, C., Chen, M., Lau, F.: Moving Big data to the cloud: an online cost-minimizing approach. IEEE J. Sel. Areas Commun. **31**(12), 2710–2721 (2013)
25. Zheng, K., Yang, Z., Zhang, K., Chatzimisios, P., Yang, K., Xiang, W.: Big data-driven optimization for mobile networks toward 5G. Network **30**(1), 44–51 (2016)

Application of Optimization Technique
for Performance and Emission Characteristics
of a CNG-Diesel Dual Fuel Engine:
A Comparison Study

A. Adarsh Rai[✉] 🄳, B. R. Shrinivasa Rao 🄳,
Narasimha K. Bailkeri 🄳, and P. Srinivasa Pai 🄳

NMAM Institute of Technology, Nitte, Udupi District, Karnataka, India
{adarsh.rai, shrinivasaraobr, nkbailkeri,
srinivasapai}@nitte.edu.in

Abstract. This work is concerned with development of a hybrid optimization technique established on particle swarm optimization (PSO) and Firefly (FA) algorithm for simultaneous optimization of performance and emission characteristics (output characteristics) of CNG-Diesel dual fuel engine (DFE). The study focusses on a suitable objective function (OF) development based on ANN paradigm using RBF-ELM network and application of PSO, FA and hybrid PSO-FA algorithm for optimization. Though all the algorithms give similar predicted values for the inputs to obtain the optimized outputs, the focus on getting the optimal results in the least possible time with necessary stability has been exhibited by the hybrid model, which has not been attempted before for engine modeling applications, as per the authors' literature survey.

Keywords: CNG-Diesel · Firefly · PSO · Hybrid optimization
Dual fuel engine

1 Introduction

Increased fossil fuel utilization and stringent emission regulations imposed by the government and other organizations has led to the use of alternative fuels like hydrogen, compressed natural gas (CNG), liquid petroleum gas (LPG) etc., which decreases the dependency over fossil fuels. These alternatives have performance characteristics comparable with that of diesel engine [1, 2]. Liu et al. [2] considered the impacts of primary fuel amount with optimized injection timings on CNG diesel DFEs and found that emissions of Nox and particle matter was decreased. Papagiannakis et al. [3] showed that diesel can be substituted with natural gas by 80% for a constant inlet temperature, under different speeds and load. Mahla et al. [4] conducted experiments on CNG-Diesel dual fuel engine using cooled exhaust gas recirculation (EGR) and found that NOx decreased, whereas smoke and carbon monoxide increased at higher EGR ratio. Pirouzpanah and Sarai [5] performed experiment on CNG-Diesel DFE under different loads and EGR ratio. It is found that, application of EGR decreased emissions like NOx, unburnt hydrocarbon (HC), soot emissions with lowest

performance sacrifice for part and full load conditions. Reducing engine emissions and fuel utilization are the major difficulties encountered by researchers to find optimal engine design and input parameters to improve engine performance and decrease emissions.

Therefore, stable emission and fuel utilization modeling is necessary for diesel engine design. In the present day, rapid development in computer technology has made modelling of the complex data using tools like artificial neural network (ANN) which can learn from examples and can handle noisy data, thus it is better than thermodynamic modelling, which is costly and time-consuming [6–8]. Wong et al. [9] used Extreme Learning Machine (ELM) to model an diesel engine for small data set, along with logarithmic transformation to reduce the effect of data scarcity and data exponentiality simultaneously. The results showed that predicted value from the model is in close acceptance with the experimental data.

Optimization is most importantly used in many engineering application, to utilize the valuable resources, time and money under given constraints. From last two decades nature inspired optimization techniques has gained popularity, since it is versatile and very efficient in solving non-linear design problems like optimization of internal combustion engines [10]. Wong et al. [9] used ELM for predicting the output characteristics of biodiesel engine and optimized the predicted values using PSO, such that the performance increases and emission decreases. To speed up the convergence of optimization techniques two parameters has a vital role to play and they are use of existing information, known as 'exploitation'. Production of new solutions in the search field is termed as 'exploration'. Hybrid optimization makes use of exploration and exploitation in the search field so that the convergence of optimized solution is faster and it avoids premature convergence [11]. Bertram et al. [1] and Jeong et al. [15] used PSO-Genetic algorithm (GA) hybrid optimization for improving output characteristics of a diesel engine, it is been found that hybrid method was more efficient and effective in finding the optimal input parameters for diesel engine, which would increase performance and decrease emissions, when compared to individual optimization techniques. In this work, ELM has been utilized to predict the output characteristics of a CNG-diesel DFE, which is been optimized using a PSO-Firefly (FA) hybrid optimization technique.

2   Experimental Details and Data Generation

In this investigation, the tests were led on a 'naturally aspirated, water cooled, single cylinder, direct injection diesel engine computerized test rig' is being used. Table 1 gives the specification of the engine. The engine were been linked to an 'eddy current dynamometer' for loading and engine was made to run at 1500 rpm, with the help of a governor.

The experiments were conducted at different loads of 0, 3, 6, 9 and 12 kg, at different compression ratios of 18, 17, 16 and 14, along with different flow rates of CNG gas (0, 0.2, 0.4, 0.6 and 0.8 kg/hr.), resulting in 100 sets of sample data.

Table 1. Specifications of engine

Sl. no.	Parameter	Specification
1	Type	Four stroke direct injection single cylinder diesel engine
2	Software used	Engine soft 8.5
3	Injection opening pressure	200 bar
4	Rated power	3.5 kW @1500 rpm
5	Cylinder diameter	87.5 mm
6	Stroke	110 mm
7	Compression ratio	12 to 18
8	Injection pressure/advance	200 bar/23^0 bTDC

3   Diesel Engine Modelling

3.1   Model Parameters

In this investigation, only five input parameters are selected, which were engine load, compression ratio, air consumption, fuel utilization and CNG flow rates. The performance parameters brake thermal efficiency (BTE) and percentage of exhaust emissions (HC, CO, NOx, and smoke) were taken as output parameters. An ELM based Radial basis function (RBF) neural network has been utilized to predict the output parameters, for the given input parameters. It is chosen because its shape is similar to normal distribution and can readily fit scattered and non-linear data, like emission data of an engine. The predicted values of the model is being optimized using PSO, FA and PSO-FA hybrid optimization techniques. To train the model, the input parameters is normalized using Eq. 1, to increase the prediction accuracy and outputs are not normalized since the values of the output are less than one.

$$x^* = \frac{x - x_{\min}}{x_{max} - x_{min}} \tag{1}$$

Where x^* is the normalized parameter, x_{max} and x_{min} are the maximum and minimum values in the dataset.

4   Optimization Techniques

IC engine optimization can be done in two ways; first, one is to do the experiments on the engine and the results are compared and appropriate optimal choice are made which is purely dependent on engineering experience and also it is time consuming and costly. The second approach is using numerical optimization, which is often based on nature-inspired algorithms like PSO, Ant Colony Optimization, FA etc. [11]. The following paragraphs describes PSO and FA methods in more detail.

4.1 Particle Swarm Optimization

PSO is a bio-inspired algorithm, based on flock behavior in nature, such as fish and bird schooling. The flock discovers the best solution or food source, through simultaneous communication among other members, which already have better solution [13].

The regulating variables of PSO algorithm are particles, initial velocity, position, number of iteration, cognitive parameter for local best influence, social parameter for universal best influence, inertia weight w, the universal best particle (*gbest*) helps in faster convergence but there is a chances of prematured convergence where as local best particle (*pbest*) is bit slower but less chances of prematured convergence and inertia weight controls the trade off between local and universal exploration abilities of swarm [15].

Algorithm for PSO:

Step 1: Create initial population of particles
Step 2: Evaluate the OF to get the initial particle position x_{id}^{old} and velocity v_{id}^{old}.
Step 3: Update the local best parameter (pbest) and universal best solution (gbest), if present position is better than previous position
Step 4: Update the particle velocity using Eq. 2
Step 5: Move the particles to the new position using Eq. 3
Step 6: Repeat step 2, till maximum iteration is reached.
Step 7: End [15]

The implementation of PSO algorithm is done in MATLAB 2013a.

$$v_{id}^{new} = w \times v_{id}^{old} + c_1 \times rand \times \left(pbest - x_{id}^{old}\right) + c_2 \times rand \times \left(gbest - x_{id}^{old}\right) \qquad (2)$$

$$x_{id}^{new} = x_{id}^{old} + v_{id}^{new} \qquad (3)$$

where c_1 *and* c_2 are self-recognition and social components respectively.

4.2 Fire Fly Algorithm

FA is metaheuristic algorithm inspired from nature, based on the natural behavior of fireflies. FA has following idealized rules:

- All fireflies are asex, hence they get attracted to one another despite of sex.
- Level of attractivity is corresponding to a firefly's brightness and hence brighter one attracts the less brighter one towards it.
- Attractiveness corresponds to the brightness, but it is inversely proportional to distance between two fireflies.
- A firefly will move randomly if it does not find a brighter than itself.
- The landscape of the OF affects the brightness of the firefly [14].

To obtain optimum value of the OF above rules need to be used.

Fireflies gets attracted to one another due to light intensity, Hence attractiveness

$$\beta = \beta_0 \times e^{-\left(\gamma \times r^2\right)} \tag{4}$$

Where β_0 is the attractivity at distance r = 0. The motion of a firefly i brings near to another due to attractivity, brighter firefly j is determined by

$$x_i^{t+1} = x_i^t + \beta_0 \times e^{-\left(\gamma r_{ij}^2\right)}\left(x_j^t - x_i^t\right) + \alpha \epsilon_i^t \tag{5}$$

Where the secondary expression is due to the attractivity, third expression is randomization, with α being the random number and ϵ_i is a vector of randomly generated numbers are obtained from a gaussian distribution.

The algorithm for FA is as follows is written based on pseudo code given in

Step 1: Produce primary population of firefly

Step 2: Evaluate the OF to generate initial light intensity for the fireflies.

Step 3: Initialize the randomization parameters, attractiveness and light absorption co-efficient (α, β and γ).

Step 4: Compare the firefly's brightness at distance i and j. If the brightness of j is more than i. Then firefly at i will move to j.

Step 5: Update the attractiveness using Eq. 4 and move to the new position using Eq. 5

Step 6: Ranking the fireflies and finding the current best

Step 7: Repeat step 2, till maximum iteration is reached.

Step 8: End [14].

4.3 Hybrid PSO-FA Optimization

In this paper, a hybrid PSO-FA algorithm is implemented for decreasing the emissions and increasing the efficiency of the engine simultaneously. The suggested PSO-FA is a strategy of bringing together the benefit of faster calculation of PSO with steadiness of FA to expand the universal search ability. Initial set of random input values is generated and then the evaluation of fitness function is obtained. The first 50% of the input data is passed to the PSO as particles and another 50% to FA as light intensity of firefly to find the best solution. The output is being combined and sorted in increasing order to find the lowest among the whole set and then the termination criteria, which is number of iteration in this case is checked and then the new set of values generated by the PSO and FA is passed to the OF and the cycle repeats. The lowest value found during sorting is the optimal value of inputs that can result in higher performance and lower emissions. The algorithm is as shown in Fig. 1.
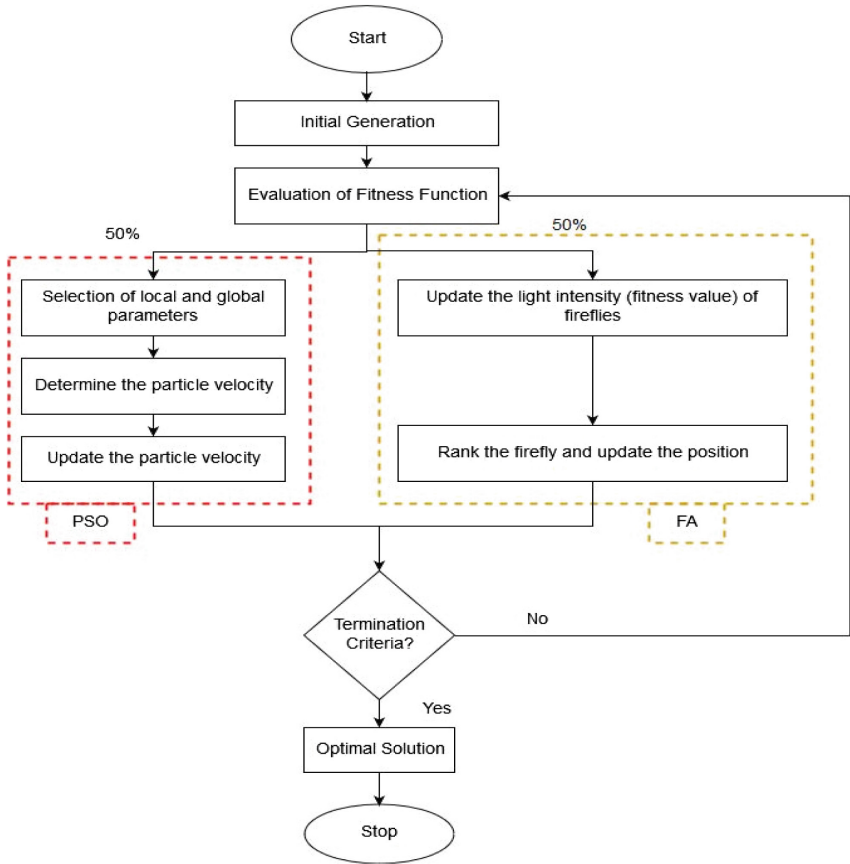
Fig. 1. Flowchart of hybrid PSO-FA algorithm [16]

4.4 Objective Function Development

Objective function is very important in any optimization problem. There are different ways to obtain the OF, when studying a predictive task. Many researchers have used regression equation as OFs for engine based problems. But, it is found that ANN prediction models performance outperforms the regression model and are showed in [17]. It is seen that ELM has a better prediction accuracy for less data [9, 11]. In this work ELM RBF algorithm is used. The training data of 85 dataset has been selected from a total dataset of 100 for training the ELM model, that have been developed by, and is being used to predict the performance parameter BTE and emission characteristics CO, NOx, HC, and smoke. The performance parameter should increase and emission parameters should decrease and is defined in a single OF. To overcome the opposite nature of the output characteristics, the authors have defined a new term that is unused heat energy, which is given as (1-BTE), which should decrease along with emissions. So the following OF is being implemented which is similar to that described in [17].

$$OF(x) = (1 - BTE(x)) + \sum_{i=1}^{4} K_i \times \exp\left[\left(\frac{E(x)_i - MAX_i}{MAX_i}\right)\right] \qquad (6)$$

Where $E(x)_i$ describes the amount of emission at the instance of i for the single value x, MAXi maximum amount for emission parameter in the data set and K_i is a penalty factor and it is taken as small as possible.

5   Results and Discussion

The main purpose of this study to achieve optimal input parameters engine load (kg), compression ratio, Air consumption (mmWC), fuel utilization (kg/hr) and CNG flow rate (kg/hr) for which performance should increase and emissions decrease, using a single OF. Trials were conducted for population size of 10, 20, 30 for both PSO and FA. It was found that the population size had little effect on the convergence and hence it was taken to be 10, so that time consumed due to population size for both FA and PSO is reduced. The timing taken by the three population sizes is shown in Table 2. The time taken for different population size are in seconds. All the input datasets have been normalized using Eq. 1.

Table 2. Execution time for different population size

Optimization techniques	Population size		
	10	20	30
PSO	0.455777	0.486046	0.513664
FA	0.478854	0.480983	0.639680
PSO-FA	0.241338	0.275976	0.309952

5.1   PSO Simulation Parameters

In the present work, the following parameters were set:

Number of iterations = 300
Population size = 10
Self-recognition $c_1 = 1$
Social components $c_2 = 3$
Inertia weight w $= \frac{maxit - itr}{maxit}$
Where $maxit$ = maximum number of iteration and itr = present iteration.

5.2   FA Simulation Parameters

Parameters like number of iterations, population size of fireflies, randomization parameter, light absorption coefficient and attractiveness are required for FA and their values are as given below:

Population size of the fireflies: 10
Attractiveness = 0.1
Randomization parameter = 0.1
Number of iteration = 300
Light absorption coefficient = 1.
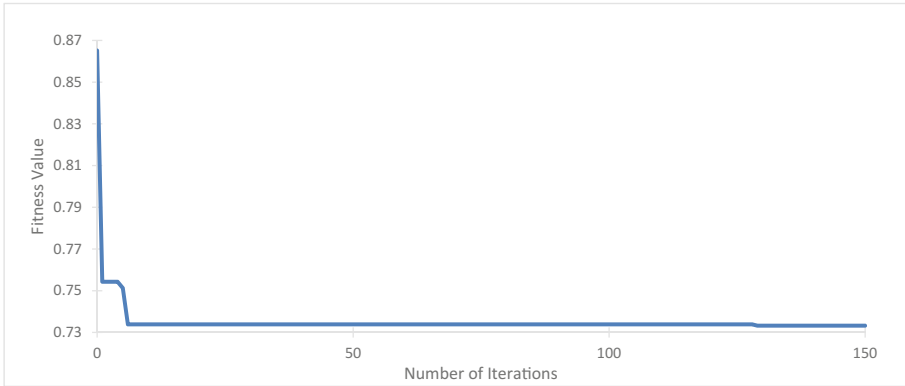
5.3   Hybrid Optimization Parameters

Population size: 10
Number of iterations = 150
All other values is kept same as described above.
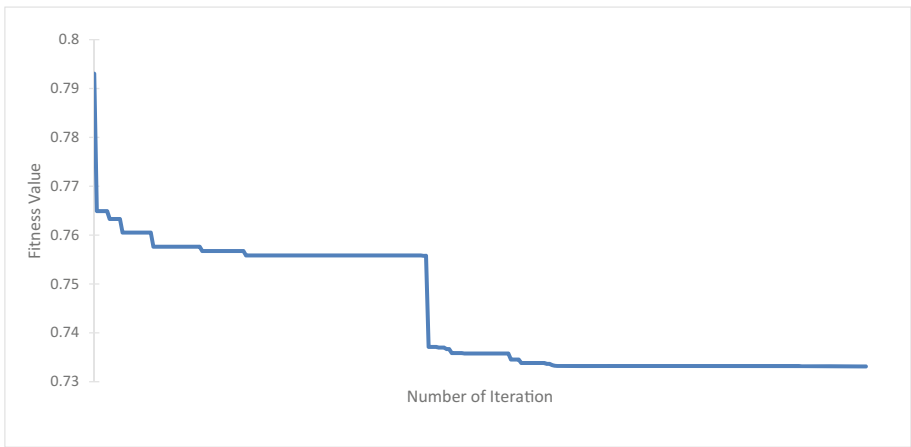
Fig. 2. Fitness value v/s number of iterations for PSO
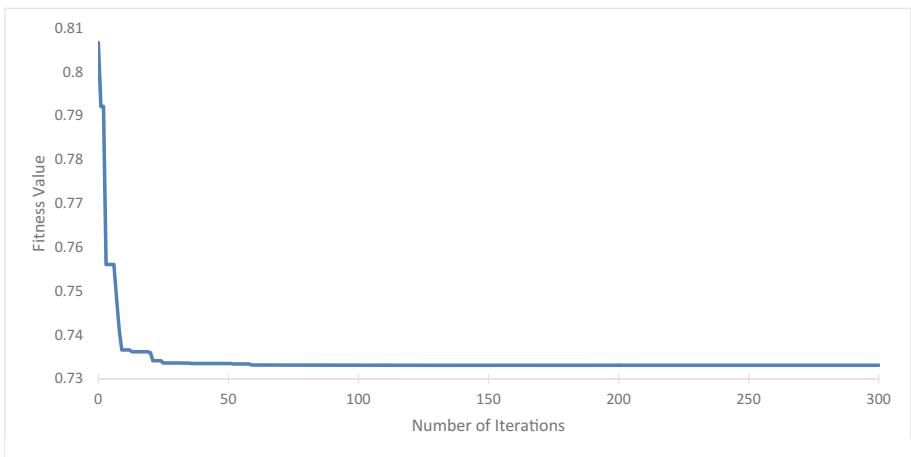
Fig. 3. Fitness value v/s number of iterations for FA

Fig. 4. Fitness value v/s number of iterations for PSO-FA hybrid

6   Comparison of Optimization Techniques

In this work, the target is on hybrid PSO-FA optimization technique. In this algorithm the benefits of both PSO and FA are available. Due to the advantages of faster calculation of PSO with steadiness of FA to expand the global search ability, the Hybrid PSO-FA converges faster like FA and takes less time like PSO. Since all the three optimization techniques give the same optimal input parameter values namely 12 kg engine load, 16 compression ratio, 0.2 kg/hr CNG flow rate, which are controllable parameters, 60.3756 mm WC of air and 0.8556 kg of fuel, which are uncontrollable parameters, for which the BTE is 26.63%, NOx is 352.03 ppm, CO is 0.032%, Smoke is 37.87 HSU and HC is 111.14 ppm. Hence, the comparison is done with respect to number of iterations, time taken by each optimization technique and number of population, which is shown in Table 3. It can be seen from Figs. 2, 3 and 4, the hybrid technique gives faster convergence when compared to individual optimization techniques. It converges quickly at 10th iteration and then remains almost constant, whereas firefly algorithm takes up to 50 iterations and PSO about 170 iterations to converge to the same value. The reason for faster computational speed in hybrid technique is due to attractiveness of the fireflies, which helps in quick convergence, and due to the PSO algorithm, the FA is not getting trapped into local minima [13].

It can be seen that there is a reduction in execution time for the hybrid technique by 47.05% when compared with PSO and 49.6% when compared with FA algorithm and number of iterations taken is 150, which is 50% less compared to the individual optimization techniques.

Table 3.   Comparison of optimization techniques

Optimization technique	Execution time (s)	Population size	Number of iterations
PSO	0.455777	10	300
FA	0.478854	10	300
PSO-FA	0.241338	10	150

7 Conclusion

In the current investigation, CNG-Diesel DFE has been used to obtain experimental dataset, which is used in ANN based model development using ELM-RBF neural network, which is used an OF. The model predicts the output characteristics for different input parameters like engine load, compression ratio, CNG gas flow, air and fuel utilization. In this work, the target is on development of a PSO-FA hybrid technique, which has not been attempted in optimization of engine parameters such that performance increases and emission decreases simultaneously. The conclusions drawn from the present work is as follows:

- ELM-RBF neural network takes less time to model the output characteristics of a CNG-Diesel DFE.
- PSO-FA Hybrid technique has the benefit of faster calculation of PSO with steadiness of FA to expand the global search ability.

References

1. Bertram, A.M., Zhang, Q., Kong, S.-C.: A novel particle swarm and genetic algorithm hybrid method for diesel engine performance optimization. Int. J. Engine Res. **17**, 732–747 (2016). https://doi.org/10.1177/1468087415611031
2. Liu, J., Yang, F., Wang, H., Ouyang, M., Hao, S.: Effects of pilot fuel quantity on the emissions characteristics of a CNG/diesel dual fuel engine with optimized pilot injection timing. Appl. Energy **110**, 201–206 (2013). https://doi.org/10.1016/j.apenergy.2013.03.024
3. Papagiannakis, R.G., Rakopoulos, C.D., Hountalas, D.T., Rakopoulos, D.C.: Emission characteristics of high speed, dual fuel, compression ignition engine operating in a wide range of natural gas/diesel fuel proportions. Fuel **89**, 1397–1406 (2010). https://doi.org/10.1016/j.fuel.2009.11.001
4. Mahla, S.K., Das, L.M., Babu, M.K.G.: Effect of EGR on performance and emission characteristics of natural gas fueled diesel engine. Jordan J. Mech. Ind. Eng. **4**(4), 523–530 (2010)
5. Pirouzpanah, V., Sarai, R.K.: Reduction of emissions in an automotive direct injection diesel engine dual-fuelled with natural gas by using variable exhaust gas recirculation. Proc. Inst. Mech. Eng. Part. D: J. Automob. Eng. **217**, 719–725 (2003). https://doi.org/10.1243/09544070360692104
6. Rai, A., Satheesh Kumar, N., Srinivasa Pai, P., Shrinivasa Rao, B.R.: Fuzzy logic based prediction of performance and emission parameters of a LPG-diesel dual fuel engine. Procedia Eng. **38**, 280–292 (2012). https://doi.org/10.1016/j.proeng.2012.06.036
7. Yusaf, T.F., Buttsworth, D.R., Saleh, K.H., Yousif, B.F.: CNG-diesel engine performance and exhaust emission analysis with the aid of artificial neural network. Appl. Energy **87**, 1661–1669 (2010). https://doi.org/10.1016/j.apenergy.2009.10.009
8. Shivakumar, S., Srinivasa Pai, P., Shrinivasa Rao, B.R.: Artificial neural network based prediction of performance and emission characteristics of a variable compression ratio CI engine using WCO as a biodiesel at different injection timings. Appl. Energy **88**, 2344–2354 (2011). https://doi.org/10.1016/j.apenergy.2010.12.030

9. Wong, P.K., Vong, C.M., Cheung, C.S., Wong, K.A.I.N.: Diesel engine modelling using extreme learning machine under scarce and exponential data sets. Int. J. Uncertain. Fuzziness Knowl.-Based Syst. **21**, 87–98 (2013). https://doi.org/10.1142/s0218488513400187

10. Yang, X.-S.: Engineering Optimization: An Introduction with Metaheuristic Applications. Wiley, Hoboken (2010)

11. Wong, K.I., Wong, P.K., Cheung, C.S., Vong, C.M.: Modeling and optimization of biodiesel engine performance using advanced machine learning methods. Energy **55**, 519–528 (2013). https://doi.org/10.1016/j.energy.2013.03.057

12. Yang, X.-S.: Nature-inspired Metaheuristic Algorithms. Luniver Press, Bristol (2010)

13. Yang, X.-S.: Nature-Inspired Optimization Algorithms. Elsevier, Amsterdam (2014)

14. D'Mello, G., Srinivasa Pai, P., Puneet, N.P.: Optimization studies in high speed turning of Ti-6Al-4V. Appl. Soft Comput. **51**, 105–115 (2017). https://doi.org/10.1016/j.asoc.2016.12.003

15. Jeong, S., Obayashi, S., Minemura, Y.: Application of hybrid evolutionary algorithms to low exhaust emission diesel engine design. Eng. Optim. **40**, 1–16 (2008). https://doi.org/10.1080/03052150701561155

16. Alonso, J.M., Alvarruiz, F., Desantes, J.M., Hernndez, L., Hernndez, V., Molt, G.: Combining neural networks and genetic algorithms to predict and reduce diesel engine emissions. IEEE Trans. Evol. Comput. **11**, 46–55 (2007). https://doi.org/10.1109/tevc.2006.876364

17. Huang, G.-B., Zhou, H., Ding, X., Zhang, R.: Extreme learning machine for regression and multiclass classification. IEEE Trans. Syst. Man Cybern. B Cybern. **42**, 513–529 (2012). https://doi.org/10.1109/TSMCB.2011.2168604

Use of Hybrid Algorithm for Surface Roughness Optimization in Ti-6Al-4V Machining

Grynal D'Mello[(✉)] ⓘ, P. Srinivasa Pai ⓘ, and Adarsh Rai ⓘ

NMAM Institute of Technology, Nitte,
Karkala Taluk, Udupi District 574110, Karnataka, India
grynal@nitte.edu.in

Abstract. In this study, an effort has been made to develop a hybrid optimization algorithm based on Particle Swarm Optimization (PSO), a widely used optimization technique and Bat Algorithm (BA) a newly introduced meta-heuristic algorithm. The machining parameters namely cutting speed (V_c), feed rate (f), depth of cut (d) along with tool wear (VB) and cutting tool vibrations (V_y) are the inputs. Further a Hybrid PSO-BA algorithm has been implemented in order to optimize surface roughness in High Speed Turning (HST) of Ti-6Al-4V. It was observed that the proposed hybrid PSO-BA algorithm increases the accuracy and convergence by 1.92% and 4.17% for R_a and R_t, which are surface roughness parameters compared to BA. Validation experiments have been performed based on PSO-BA predicted values. The experimental values are close to the predictive values with an error of 0.47% for R_a and 1.159% for R_t which is considerably better than that obtained from BA only.

Keywords: Ti-6Al-4V · Surface roughness · PSO · BA · Hybrid PSO-BA

1 Introduction

Surface roughness is an extensively used index to identify the quality of a product. Achieving desired surface roughness is a challenging task in case of any manufacturing process. Several machining parameters affect the quality of the surface and many uncontrollable parameters make it difficult to achieve a high quality surface [1]. Ti-6Al-4V is a widely used titanium alloy. It is an excellent corrosion resistant material and is used in biomedical applications, aerospace etc. It is considered as a difficult to machine due to its high temperature capability, better fatigue strength and reaction with the cutting tools. It a costly material and incur high machining costs and high tool wear rate [2, 3]. Conventional machining consumes lot of time which decreases the productivity. The quality of the surface can be improved by introducing High Speed Machining (HSM) which is considerably faster than the conventional machining processes [4]. High speed cutting of Ti-6Al-4V was presented by Sutter and List to understand the effect of forces and chip formation mechanisms during cutting. From their study, it was concluded that cutting speed plays a vital role in defining the cutting forces and chip formation [5]. Ramesh et al. tried to analyse the quality of surface during machining Ti-6Al-4V. They concluded that feed rate is the most significant parameter affecting the quality of the surface followed by cutting speed, whereas depth of cut showed negligible effect [6]. Cutting tool vibrations and tool wear are considered

as uncontrollable parameters which affect the quality of surface. D'Mello et al. studied the effect of VB and V_y along with machining parameters on surface roughness of Ti-6Al-4V. The cutting tool vibrations were found to be low at higher cutting speeds for uncoated carbide inserts thus proving that high speed machining is suitable for turning Ti-6Al-4V. A direct correlation between surface roughness and vibration was observed, whereas in case of tool wear it was vice versa [7].

Optimization techniques are widely applied to find the best solution for different problems. Implementation of these techniques can drastically reduce the cost, increase the productivity and also achieve better quality of the product. Artificial intelligence techniques like Artificial Neural Network (ANN) [8–10], Fuzzy Logic [11–13], Genetic Algorithm (GA) [14–16], Particle Swarm Optimization (PSO) [17–19], Ant Colony Optimization (ACO) [20–22], Artificial Bee Colony (ABC) [23, 24] and Firefly Algorithm (FA) [25–27] have been widely implemented in various machining applications.

Raja and Baskar applied PSO to find out the suitable machining parameters during turning. Pilot experiments were performed on a CNC turning centre using tungsten carbide inserts. The main objective of this study was to reduce the machining time and surface roughness. They concluded that feed rate was found to have significant effect on machining time and quality of surface [28]. Tanweer et al. developed Self Regulating Particle Swarm Optimization (SRPSO) algorithm and evaluated using CEC 2005 standard benchmarking. The obtained results from SRPSO were compared with the six state of art PSO algorithm and found that SRPSO achieves faster convergence and generates better solutions [29]. Escamilla-Salazar et al. implemented swarm intelligence during machining Ti-6Al-4V. Multi objective optimization process was used in this study to optimize cutting temperature and surface roughness. The objective function was defined from the Back propagation neural network. The study concluded that use of swarm intelligence with different objectives were efficient and effective thereby providing intelligent production planning [30]. Several works have been carried out by the researchers to improve the conventional PSO by trying to overcome the convergence issues [31, 32]. Garg et al. used ANN as a fitness function to optimize using PSO. The objective was to predict the flank wear in drilling using PSO trained neural network. The prediction showed the robustness of PSO in accurately predicting the flank wear [33].

Xing She Yang introduced a new metaheuristic algorithm named Bat Algorithm (BA) based on the echolocation behaviour of bats. Yang applied BA for multi objective optimization problems. This algorithm was validated against different engineering problems and then applied on design problems like welded beam design. Results revealed that multi objective bat algorithm is an effective multi objective optimizer [34]. Yang carried out a literature review and the applications of BA in various fields. It has been revealed that BA is more efficient when compared to other optimization techniques due to its ability of frequency tuning, automatic zooming and parameter control [35].

Many researchers have tried using Hybrid algorithms which uses two optimization algorithms and develop a connection to use the key points of each algorithm in developing a single algorithm. Pan et al. developed hybrid algorithm using PSO and

BA. This produced the best practical results, when compared to only PSO and BA [36]. Yildiz and Ozturk implemented hybrid enhanced genetic algorithm in the selection of optimum machining parameters in turning operations. The proposed technique achieved superior results when compared to other techniques from the literature [37]. Costa et al. applied hybrid particle swarm optimization technique for the optimization of multi-pass turning operations. The results of conventional PSO and hybrid PSO were compared. Hybrid PSO outperformed conventional PSO in the multi pass turning operation [38]. Li and Yin proposed a particle swarm inspired cuckoo search algorithm for real parameter optimization. A comparison between conventional PSO, Cuckoo search (CS) and Particle Swarm inspired Cuckoo Search (PSCS) was done and PSCS outperformed other two techniques [39].

From the literature survey carried out, the hybrid algorithms are gaining more importance and produce more accurate validation results when compared to conventional optimization techniques. PSO has been widely used in all the areas of engineering and has been successfully implemented. There are no efforts towards the use of BA in the area of machining. This study makes use of conventional PSO and BA techniques to optimize the machining parameters namely V_c, f, d followed by uncontrollable parameters VB and V_y to achieve least surface roughness parameters R_a and R_t during high speed turning of Ti-6Al-4V. Further a hybrid PSO-BA algorithm has been implemented to further optimize the parameters and a comparison has been made. A similar study has been made by the authors in [40] for the same application, where the performance of PSO has been made with BA and FA and BA has been established to be the best. Hence in this work, the performance of BA has been compared with the hybrid PSOBA. The experiments have been performed using 883 inserts (uncoated carbide) and the objective function has been derived using ANN model. The results obtained have been validated to confirm the effectiveness of the developed hybrid algorithm.

2 Experimental Investigations

2.1 Work Material

In this study, Grade 5 titanium alloy Ti-6Al-4V has been used as the work material. The material composition in grade 5 titanium alloy include 6.02% Al, 0.03% Cr, 0.13% Fe, 0.04% Mn, 3.85% V and 89.93% Ti. A 50 mm diameter and 200 mm long workpiece is held in a three jaw chuck on a high speed CNC lathe (Make: HMT, Model: Stallion 100 SU). The experiments are conducted under dry conditions.

2.2 Cutting Tool Material

The experiments are performed using uncoated carbide inserts 883 with MR4 chip breaker supplied by SECO tools of ISO grade CNMG 120408 held in PCLNL 2020 K12 tool holder (Make: SECO). Experiments are conducted by keeping a tool overhang length of 60 mm. The geometry of the tool is as follows: back and side rake angle: −6°, end cutting angle: 5°, principal cutting edge angle 75° and nose radius of 0.8 mm.

2.3   Experimental Design

Three level full factorial experimental design has been used to conduct experiments in a systematic manner. Five input parameters namely V_c, f, d, VB and V_y have been used in this study. Experimental design has been developed by considering only V_c, f and d. VB and V_y cannot be considered during experimental design as these are considered as uncontrollable parameters. The experimental design has been done for three levels of cutting speed (150, 175 and 200 m/min), feed rate (0.15, 0.2 and 0.25 mm/rev) and depth of cut (0.8, 1 and 1.2 mm). The design is developed using MINITAB 17 Statistical software [43].

2.4   Signal Measurement

The cutting tool vibrations have been measured online using an accelerometer (Make: ISOTRON) mounted on the tool holder near to the cutting zone. The signals are measured in three perpendicular directions, i.e. depth of cut, cutting speed and feed rate denoted as V_x, V_y and V_z respectively. These signals are sent to the computer using DNA cube hardware using Lab View software and processed further. The vibration signals along cutting speed direction has been found to be more significant than other directions [7]. Hence only V_y has been considered for further evaluation in this work.

2.5   Tool Wear Measurement

The tool flank wear measurement is done offline using a Tool Makers Microscope (MITUTOYO make) having a magnification of 15X. A machining pass of length 48 mm is considered during the experimentation process. After each machining pass, the tool insert is detached from the tool holder and tool flank wear is measured by placing the insert in the microscope. The measurement is done till the tool flank wear reached 0.4 mm as per ISO 3685:1993 consideration standard.

2.6   Surface Roughness Measurement

After each machining pass, the surface roughness is measured using Taylor Hobson Taly Surf 50 by placing the workpiece on a V block. The surface roughness of the turned surface is measured at three different locations and the average value of surface roughness has been considered as the final value for further analysis.

3   Objective Function Development Using ANN

ANN is a set of neurons interconnected to each other which has the ability to learn the multivariable relationships between the process parameters. A network consists of an input, hidden and an output layer. The predicted output by the network is generated in the output layer based on the input and variation of neurons in the hidden layer [41].

This study makes use of feed forward supervised Multi-Layer Perceptron Neural Network (MLPNN) which has an input layer, a hidden layer and an output layer. The output of this network is considered as an objective function for all the optimization algorithms investigated in this study.

The experimental data collected from the study has been divided into training and test data as per the requirement of MLP network. 246 experimental data have been considered as training data and remaining 43 as test data. Only the training data has been considered for the development of the objective function. The network has been trained using *tansig* and *logsig* transfer functions along with *trainlm*, *trainbfg*, *trainscg* and *traincgf* training algorithms available in neural network tool box of MATLAB 2014b [44].

Simulations parameters for MLP considered in this study include the mean squared error (MSE) of 0.001, maximum epochs of 1000 and a learning rate of 0.01. The network has been trained to achieve highest $R2$ (Coefficient of regression) by wavering the number of hidden neurons between 5 and 20. The input parameters include V_c, f, d, VB and V_y and output parameters are R_a and R_t. The stopping criterion of the network is set as either maximum epochs or least MSE reached.

The main goal of the objective function is to minimize the surface roughness parameters R_a and R_t. The objective function can be developed in the form

$$R_a = [V_c \, f \, d \, VB \, V_y \, V_c^2 \, f^2 \, d^2 \, VB^2 \, V_y^2 \, V_c * f \, V_c * d \, V_c * VB \, V_c * V_y \, f * d \, f$$
$$* VB \, f * V_y \, d * VB \, d * V_y \, VB * V_y][C] \tag{1}$$

where V_c = cutting speed (m/min), f = feed rate (mm/rev), d = depth of cut (mm), VB = tool flank wear (mm), V_y = cutting tool vibration in speed direction (g), C = Matrix containing coefficients in terms of equation. This objective function has been derived based on the study conducted by the authors in [40].

The best model output for R_a has been achieved using *logsig* transfer function with *trainlm* (Levenberg-Marquardt) training algorithm producing an $R^2 = 0.991$ for 20 hidden neurons. Hence, this model has been considered for developing the objective function. Similarly, the best model output for R_t has been achieved using *tansig* transfer function with *trainlm* training algorithm producing an $R^2 = 0.975$ for 20 hidden neurons. This has been used as the objective function for all the optimization algorithms investigated in this study.

4 Optimization Algorithms

4.1 Particle Swarm Optimization (PSO)

Particle Swarm optimization is an evolutionary non-traditional optimization technique which deals with an exploration method based on population explained by Kennedy and Russell in 1995 [42]. It is a social behaviour technique where a group of swarm could reach for their food comfortably. PSO has the ability to find the best possible solution in a large space. Best refers to the position of the particle when the objective is

to minimise the output. Each swarm is called as a particle which is allowed to move in a multidimensional space at a certain velocity. All particles undergo a condition to attain best position. 'pbest' is the particle's best position as identified by the same particle and 'gbest' is the global best position achieved by any particle in the population. The details of the PSO algorithm are available in [28].

The particle position and velocity are updated using Eq. (2)

$$v_i^{k+1} = v_i^k + c_1.rand_1^k\left(pbest_i^k - x_i^k\right) + c_2 rand_2\left(gbest_i^k - x_i^k\right) \tag{2}$$

Where v_i^k is the velocity value in the previous iteration c_1 and c_2 are the learning factors. The new position of the particles can be achieved by the expression shown in Eq. (3).

$$x_i^{k+1} = x_i^k + v_i^{k+1} \tag{3}$$

PSO algorithm has been described in this paper in order to present the concept of proposed hybrid PSO-BA optimization technique.

4.2   Bat Algorithm (BA)

Xin She-Yang developed a relatively newer metaheuristic algorithm named Bat Algorithm (BA). Meta means 'beyond' or 'higher level' which perform better than simple heuristics and heuristic means 'to find' or 'to discover by trial and error' [34]. This algorithm is inspired from the echolocation characteristic of bats. This feature helps the bats to navigate and hunt for food.

In this algorithm, the virtual bat movement simulation is expressed as shown in (4)–(6).

$$f_i = f_{min} + (f_{max} - f_{min})\beta \tag{4}$$

$$v_i^t = v_i^{t-1} + \left(x_i^t - x_*\right)f_i \tag{5}$$

$$x_i^t = x_i^{t-1} + v_i^t \tag{6}$$

Further the bats generate new solutions by flying randomly as expressed in Eq. (7).

$$x_{new} = x_{old} + \epsilon A^t \tag{7}$$

Updating of loudness parameters A_i and emission rate r_i is also done after updating the position of the bats which is represented in Eq. (8).

$$A_i^{t+1} = \alpha A_i^t, r_i^{t+1} = r_i^0[1 - \exp(-\gamma t)] \tag{8}$$

The BA implementation is based on the algorithm specified by Xing She Yang [34].

In BA, the following parameters have been set - Number of iterations 100 and the population size has been varied between 20 to 40. Loudness parameter and pulse

emission rate is varied between 0 and 1. Frequency parameters have been defined and the minimum and maximum range has been set as 0 and 2 respectively. The range of cutting speed is between 150 to 200, range of feed rate is between 0.15 to 0.25, range of depth of cut is between 0.8 to 1.2, range of tool flank wear has been 0.07 to 0.51 and the range of V_y 20.233 to 160.4382. These ranges have been kept the same for hybrid PSO-BA optimization technique also.

4.3   Proposed Hybrid PSO-BA

The structure of the hybrid algorithm is developed based on the key points of individual optimization techniques. This algorithm has been derived from the standard PSO and BA algorithms. The main objective behind developing this kind of algorithm is to highlight the strengths of individual algorithm, thereby producing most optimized results. The flow chart of the proposed PSO-BA is shown in Fig. 1.

Initialization: Generate the population. The population size is initialised for PSO, which is common for BA. Set the required parameters for both PSO and BA, in terms of particle and velocity and in terms of frequency, pulse rate and loudness parameter respectively along with the number of iterations.

Evaluation: The fitness function is evaluated which remains common for both PSO and BA.
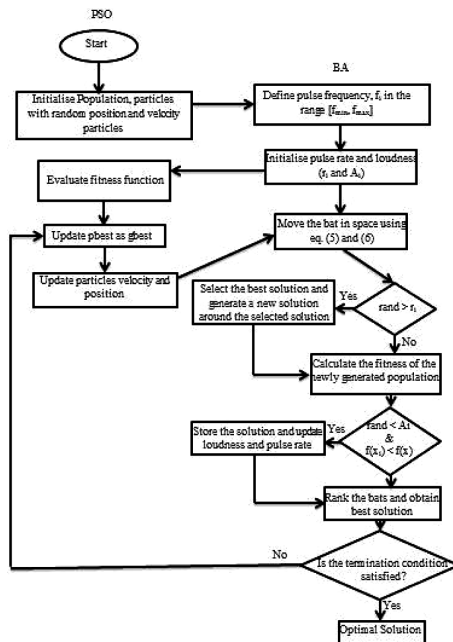
Fig. 1. Flow chart of proposed hybrid PSO-BA

Update: PSO evaluates the gbest and updates the particle velocity and position using Eqs. (2) and (3). Artificial bat movement is generated by using Eqs. (5) and (6). The location and velocity of the Bat is updated to achieve the best fitness value using Eqs. (7) and (8).

Communication Strategy: The best artificial bats obtained from the population are diverted towards the poorer particles of the population from PSO and the population is updated for each iteration.

Termination: If the termination condition is satisfied, then the algorithm stops. If not, the execution is repeated from evaluation to termination stage until the best possible result is obtained for a given number of iterations.

5   Results and Discussion

Machining parameters namely V_c, f, d followed by uncontrollable parameters namely VB and V_y have been considered for the minimization of surface roughness parameters R_a and R_t in turning Ti-6Al-4V using optimization techniques namely BA and Hybrid PSO-BA. From a previous study carried out by the authors, it has been found that BA performed better in terms of prediction and execution time when compared to PSO [40]. This study mainly focusses on the performance of hybrid PSO-BA along with the conventional BA. The individual results of BA are presented followed by that for PSO-BA hybrid.

5.1   BA

The results obtained using BA are given in Table 1. The lowest surface roughness for R_a has been found to be 0.63 µm and 2.6613 µm for R_t. These results have been obtained by setting the loudness parameter at 0.4 and pulse rate at 0.6 with a population of 40 for 100 iterations. The conditions for achieving minimum surface roughness for R_a and R_t are same given in [40] in case of PSO.

Table 1. Optimization results of BA

Roughness parameter	V_c (m/min)	f (mm/rev)	d (mm)	VB (mm)	V_y (g)	Roughness (µm)
R_a	199.51	0.156	0.998	0.501	43.40	0.6300
R_t	152.12	0.151	1.008	0.073	26.56	2.6613

Figure 2 shows the variation of surface roughness with iterations for BA for R_a. It can be seen that there has been step wise behavior in minimization of surface roughness till 70th iteration and it stabilizes after 90th iteration achieving R_a value of 0.63 µm.
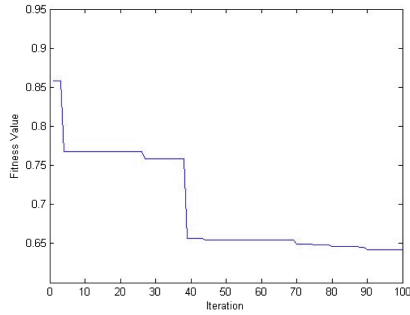
Fig. 2. Surface roughness (R_a) variation with iterations in BA

5.2 Hybrid PSO-BA

Optimal results of hybrid PSO-BA are given in Table 2. The minimum value of surface roughness parameters obtained are R_a = 0.6179 µm and R_t = 2.5501 µm. These results are obtained for 95 iterations with a population size of 45 by setting the learning factors c_1 and c_2 to 1 and 3, loudness and pulse rate has been set to 0.4 and 0.6 respectively.

Table 2. Optimization results for hybrid PSO-BA

Roughness parameter	V_c (m/min)	f (mm/rev)	d (mm)	VB (mm)	V_y (g)	Roughness (µm)
R_a	199.1373	0.153	0.9596	0.4825	27.1345	0.6179
R_t	156.8461	0.1541	1.0963	0.0779	87.8535	2.5501

Figure 3 shows the variation of surface roughness with iterations for proposed PSO-BA. It can be seen that minimization takes place till 8th iteration and remains constant till 95th iteration showing a quick convergence to reach the minimum value of surface roughness of 0.6179 µm.
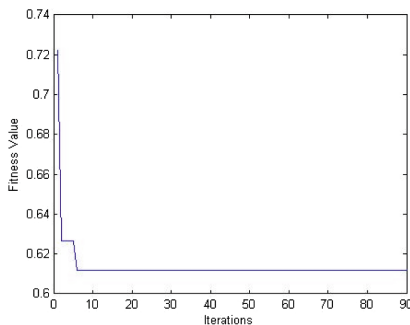
Fig. 3. Surface roughness (R_a) variation with iterations for proposed hybrid PSO-BA

5.3   Comparison of Results

In this study, the main focus is to highlight the effectiveness of hybrid PSO-BA for the optimization of surface roughness by considering cutting speed, feed rate, depth of cut, tool flank wear and cutting tool vibrations and compare it with BA.

Table 3. Comparison of optimization techniques

Comparable techniques	Parameters	Technique	Minimum value (µm)	Better technique	% improvement
BA and hybrid PSO-BA	R_a	BA	0.6300	PSO-BA over BA	1.92
		PSO-BA	0.6179		
	R_t	BA	2.6631		4.24
		PSO-BA	2.5501		

The optimum values of surface roughness parameters R_a and R_t obtained using BA and hybrid PSO-BA are compared in Table 3. The percentage improvement has been calculated for R_a and R_t. The comparison has been done between BA and hybrid PSO-BA. It is seen than hybrid PSO-BA produced an improvement of 1.92% for R_a and 4.24% for R_t over BA. Similar results have been obtained by Pan et al. during the implementation of hybrid particle swarm optimization along with bat algorithm [36]. It is also seen that hybrid PSO-BA took 95 iterations considering lowest population size of 35 to reach the least value of surface roughness with the lowest execution time of 0.62 s thereby establishing that hybrid PSO-BA is fastest in execution. This algorithm replaces the poorer particles of PSO with the best artificial bats and the poorer bats are replaced with the best particles of PSO resulting in easy communication [36].

5.4   Validation Experiments

From the comparison of three techniques, it can be seen that hybrid PSO-BA produced the best possible result when compared to BA. A validation experiment has been performed by setting the parameters obtained from hybrid PSO-BA technique. The experiments have been performed for two different conditions of R_a and R_t respectively. VB and V_y have not been considered as they are uncontrolled parameters and cannot be set by the user. Only machining parameters have been set and a tool insert with a tool wear of 0.38 mm has been taken for experimentation in case of R_a.

It can be seen from Table 4 that the results obtained through validation are very close to the PSO-BA predicted values. A comparison of these results shows that there is an error of 8.426% and 6.74% in case of VB and V_y respectively as they are uncontrollable factors, but the surface roughness parameter R_a shows an error of just 0.47% between the predicted value of 0.6179 µm and actual 0.615 µm. The error of 9.11% and 54.79% was higher in case of R_t for VB and V_y respectively but there was only 1.159% error for the response. This shows the accurate prediction of cutting conditions and other parameters using hybrid PSO-BA technique for obtaining optimal surface roughness values, which is the major goal of any machining operation. And this is particularly important during machining Ti-6Al-4V alloy.

Table 4. Results from validation experiments

Condition	For R_a			For R_t		
	Parameters obtained for PSO-BA	Actual values obtained	% error	Parameters obtained for PSO-BA	Actual values obtained	% error
V_c	199.1373	199.1373		156.8461	156.8461	
f	0.153	0.153		0.1541	0.1541	
d	0.9596	0.9596		1.0963	1.0963	
VB	0.4825	0.445	8.426	0.0779	0.085	9.11
V_y	27.1345	25.42	6.74	44.1619	28.53	54.79
Response						
Surface roughness	0.6179	0.615	0.47	2.5501	2.58	1.159

6   Conclusions

This study deals with the minimization of surface roughness parameters R_a and R_t. To attain this, machining parameters namely V_c, f, d along with VB and V_y have been optimized using PSO, BA and Hybrid PSO-BA. As per the authors' knowledge, Hybrid PSO-BA has not been applied before in any machining applications. The objective function has been derived from the output of ANN model. The following conclusions can be drawn from this study.

1. Hybrid PSO-BA has been found to be better, when compared to BA in terms of quick convergence with least execution time and minimization of surface roughness parameters R_a and R_t.
2. From the results predicted by two techniques, it is preferred to have high V_c, low f, intermediate d, high VB and low V_y for lower value of R_a.
3. R_t can be minimized by maintaining low V_c, low f, intermediate d, low VB and low V_y.

References

1. Benardos, P.G., Vosniakos, G.-C.: Predicting surface roughness in machining: a review. Int. J. Mach. Tools Manuf. **43**(8), 833–844 (2003). https://doi.org/10.1016/s0890-6955(03) 00059-2
2. Bandapalli, C., Sutaria, M.B., Bhatt, V.D.: High speed machining of Ti alloys a critical review. In: Proceedings of the 1st International and 16th National Conference on Machines and Mechanisms (iNaCoMM2013) (2013)
3. Rahman, M., Wong, Y.S., Zareena, A.R.: Machinability of titanium alloys. JSME Int. J. Ser. C Mech. Syst. Mach. Elem. Manuf. **46**(1), 107–115 (2003). https://doi.org/10.1299/jsmec. 46.107

4. Rahman, M., Wang, Z.G., Wong, Y.S.: A review on high-speed machining of titanium alloys. JSME Int. J. Ser. C Mech. Syst. Mach. Elem. Manuf. **49**(1), 11–20 (2006). https://doi.org/10.1299/jsmec.49.11

5. Sutter, G., List, G.: Very high speed cutting of Ti-6Al-4V titanium alloy change in morphology and mechanism of chip formation. Int. J. Mach. Tools Manuf. **66**, 37–43 (2013). https://doi.org/10.1016/j.ijmachtools.2012.11.004

6. Ramesh, S., Karunamoorthy, L., Palanikumar, K.: Surface roughness analysis in machining of titanium alloy. Mater. Manuf. Process. **23**(2), 174–181 (2008). https://doi.org/10.1080/10426910701774700

7. D'Mello, G., Pai, P.S., Puneet, N.P.: Surface roughness evaluation using cutting vibrations in high speed turning of Ti-6Al-4V-an experimental approach. Int. J. Mach. Mach. Mater. **18** (3), 288–312 (2016). https://doi.org/10.1504/ijmmm.2016.076281

8. Pontes, F.J., Ferreira, J.R., Silva, M.B., Paiva, A.P., Balestrassi, P.P.: Artificial neural networks for machining processes surface roughness modeling. Int. J. Adv. Manuf. Technol. **49**(9), 879–902 (2010). https://doi.org/10.1007/s00170-009-2456-2

9. Asiltürk, I., Çunkaş, M.: Modeling and prediction of surface roughness in turning operations using artificial neural network and multiple regression method. Expert Syst. Appl. **38**(5), 5826–5832 (2011). https://doi.org/10.1016/j.eswa.2010.11.041

10. Ozel, T., Correia, A.E., Davim, J.P.: Neural network process modelling for turning of steel parts using conventional and wiper inserts. Int. J. Mater. Prod. Technol. **35**(1–2), 246–258 (2009). https://doi.org/10.1504/ijmpt.2009.02523

11. Ho, S.Y., Lee, K.C., Chen, S.S., Ho, S.J.: Accurate modeling and prediction of surface roughness by computer vision in turning operations using an adaptive neuro-fuzzy inference system. Int. J. Mach. Tools Manuf. **42**(13), 1441–1446 (2002). https://doi.org/10.1016/s0890-6955(02)00078-0

12. Rai, A., Kumar, N.S., Rao, B.S.: Fuzzy logic based prediction of performance and emission parameters of a LPG-diesel dual fuel engine. Procedia Eng. **38**, 280–292 (2012). https://doi.org/10.1016/j.proeng.2012.06.036

13. Jiao, Y., Lei, S., Pei, Z.J., Lee, E.S.: Fuzzy adaptive networks in machining process modeling: surface roughness prediction for turning operations. Int. J. Mach. Tools Manuf. **44** (15), 1643–1651 (2004). https://doi.org/10.1016/j.ijmachtools.2004.06.004

14. Mariajayaprakash, A., Senthilvelan, T., Gnanadass, R.: Optimization of process parameters through fuzzy logic and genetic algorithm–a case study in a process industry. Appl. Soft Comput. **30**, 94–103 (2015). https://doi.org/10.1016/j.asoc.2015.01.042

15. Venkadesh, S., Hoogenboom, G., Potter, W., McClendon, R.: A genetic algorithm to refine input data selection for air temperature prediction using artificial neural networks. Appl. Soft Comput. **13**(5), 2253–2260 (2013). https://doi.org/10.1016/j.asoc.2013.02.003

16. Chaquet, J.M., Carmona, E.J., Corral, R.: Using genetic algorithms to improve the thermodynamic efficiency of gas turbines designed by traditional methods. Appl. Soft Comput. **12**(11), 3627–3635 (2012). https://doi.org/10.1016/j.asoc.2012.06.009

17. Torabi, S.A., Sahebjamnia, N., Mansouri, S.A., Bajestani, M.A.: A particle swarm optimization for a fuzzy multi-objective unrelated parallel machines scheduling problem. Appl. Soft Comput. **13**(12), 4750–4762 (2013). https://doi.org/10.1016/j.asoc.2013.07.029

18. Gao, L., Huang, J., Li, X.: An effective cellular particle swarm optimization for parameters optimization of a multi-pass milling process. Appl. Soft Comput. **12**(11), 3490–3499 (2012). https://doi.org/10.1016/j.asoc.2012.06.007

19. Han, F., Yao, H.F., Ling, Q.H.: An improved evolutionary extreme learning machine based on particle swarm optimization. Neurocomputing **116**, 87–93 (2013). https://doi.org/10.1016/j.neucom.2011.12.062

20. Ghaiebi, H., Solimanpur, M.: An ant algorithm for optimization of hole-making operations. Comput. Ind. Eng. **52**(2), 308–319 (2007). https://doi.org/10.1016/j.cie.2007.01.001

21. Cheng, B., Wang, Q., Yang, S., Hu, X.: An improved ant colony optimization for scheduling identical parallel batching machines with arbitrary job sizes. Appl. Soft Comput. **13**(2), 765–772 (2013). https://doi.org/10.1016/j.asoc.2012.10.021

22. Baskar, N., Asokan, P., Prabhaharan, G., Saravanan, R.: Optimization of machining parameters for milling operations using non-conventional methods. Int. J. Adv. Manuf. Technol. **25**(11–12), 1078–1088 (2005). https://doi.org/10.1007/s00170-003-1939-9

23. Yusup, N., Sarkheyli, A., Zain, A.M., Hashim, S.Z.M., Ithnin, N.: Estimation of optimal machining control parameters using artificial bee colony. J. Intell. Manuf. **25**(6), 1463–1472 (2014). https://doi.org/10.1007/s10845-013-0753-y

24. Das, M.K., Kumar, K., Barman, T.K., Sahoo, P.: Investigation on electrochemical machining of EN31 steel for optimization of MRR and surface roughness using artificial bee colony algorithm. Procedia Eng. **97**, 1587–1596 (2014). https://doi.org/10.1016/j.proeng.2014.12.309

25. Horng, M.H.: Vector quantization using the firefly algorithm for image compression. Expert Syst. Appl. **39**(1), 1078–1091 (2012). https://doi.org/10.1016/j.eswa.2011.07.108

26. Khadwilard, A., Chansombat, S., Thepphakorn, T., Thapatsuwan, P., Chainate, W., Pongcharoen, P.: Application of firefly algorithm and its parameter setting for job shop scheduling. J. Ind. Technol. **8**(1), 49–58 (2012)

27. Yang, X.S., Hosseini, S.S.S., Gandomi, A.H.: Firefly algorithm for solving nonconvex economic dispatch problems with valve loading effect. Appl. Soft Comput. **12**(3), 1180–1186 (2012). https://doi.org/10.1016/j.asoc.2011.09.017

28. Raja, S.B., Baskar, N.: Particle swarm optimization technique for determining optimal machining parameters of different work piece materials in turning operation. Int. J. Adv. Manuf. Technol. **54**(5–8), 445–463 (2011). https://doi.org/10.1007/s00170-010-2958-y

29. Tanweer, M.R., Suresh, S., Sundararajan, N.: Self-regulating particle swarm optimization algorithm. Inf. Sci. **294**, 182–202 (2015). https://doi.org/10.1016/j.ins.2014.09.053

30. Escamilla-Salazar, I.G., Torres-Trevio, L.M., Gonzlez-Ortz, B., Zambrano, P.C.: Machining optimization using swarm intelligence in titanium (6Al 4 V) alloy. Int. J. Adv. Manuf. Technol. **67**(1–4), 535–544 (2013). https://doi.org/10.1007/s00170-012-4503-7

31. Tanweer, M.R., Auditya, R., Suresh, S., Sundararajan, N., Srikanth, N.: Directionally driven self-regulating particle swarm optimization algorithm. Swarm Evol. Comput. **28**, 98–116 (2016). https://doi.org/10.1016/j.swevo.2016.01.006

32. Tanweer, M.R., Suresh, S., Sundararajan, N.: Dynamic mentoring and self-regulation based particle swarm optimization algorithm for solving complex real-world optimization problems. Inf. Sci. **326**, 1–24 (2016). https://doi.org/10.1016/j.ins.2015.07.035

33. Garg, S., Patra, K., Pal, S.K.: Particle swarm optimization of a neural network model in a machining process. Sadhana **39**(3), 533–548 (2014). https://doi.org/10.1007/s12046-014-0244-7

34. Yang, X.S.: Bat algorithm for multi-objective optimisation. Int. J. Bio-Inspired Comput. **3**(5), 267–274 (2011). https://doi.org/10.1504/ijbic.2011.042259

35. Yang, X.S.: Bat algorithm: literature review and applications. Int. J. Bio-Inspired Comput. **5**(3), 141–149 (2013). https://doi.org/10.1504/ijbic.2013.055093

36. Pan, T.-S., Dao, T.-K., Nguyen, T.-T., Chu, S.-C.: Hybrid particle swarm optimization with bat algorithm. In: Sun, H., Yang, C.-Y., Lin, C.-W., Pan, J.-S., Snasel, V., Abraham, A. (eds.) Genetic and Evolutionary Computing. AISC, vol. 329, pp. 37–47. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-12286-1_5

37. Yildiz, A.R., Ozturk, F.: Hybrid enhanced genetic algorithm to select optimal machining parameters in turning operation. Proc. Inst. Mech. Eng. Part B: J. Eng. Manuf. **220**(12), 2041–2053 (2006). https://doi.org/10.1243/09544054jem570
38. Costa, A., Celano, G., Fichera, S.: Optimization of multi-pass turning economies through a hybrid particle swarm optimization technique. Int. J. Adv. Manuf. Technol. **53**(5), 421–433 (2011). https://doi.org/10.1007/s00170-010-2861-6
39. Li, X., Yin, M.: A particle swarm inspired cuckoo search algorithm for real parameter optimization. Soft. Comput. **20**(4), 1389–1413 (2016). https://doi.org/10.1007/s00500-015-1594-8
40. D'Mello, G., Pai, P.S., Puneet, N.P.: Optimization studies in high speed turning of Ti-6Al-4V. Appl. Soft Comput. **51**, 105–115 (2017). https://doi.org/10.1016/j.asoc.2016.12.003
41. Karayel, D.: Prediction and control of surface roughness in CNC lathe using artificial neural network. J. Mater. Process. Technol. **209**(7), 3125–3137 (2009). https://doi.org/10.1016/j.jmatprotec.2008.07.023
42. Kennedy, J., Russell, C.E.: Particle swarm optimization. In: Proceedings of the 1995 IEEE International Conference on Neural Networks, Perth, Australia, pp. 1942–1948 (1995)
43. Minitab, Inc.: MINITAB release 17: statistical software for windows. Minitab Inc., USA (2014)
44. MATLAB and Statistical Toolbox R2013a. The MathWorks Inc., Natick, MA (2013)

Eigenvalue Analysis with Hough Transform for Shape Representation and Classification

Bharathi Pilar[1] ⓘ and B. H. Shekar[2(✉)] ⓘ

[1] University College, Mangalore, Karnataka, India
bharathi.pilar@gmail.com
[2] Mangalore University, Mangalore, Karnataka, India
bhshekar@gmail.com

Abstract. In this work, we present eigenvalue based shape descriptor (EHough) which makes use of small eigenvalue and large eigenvalue along with Hough Transform to obtain the dominant features. The small eigenvalue and large eigenvalue are computed for each pixel associated with a shape boundary. In order to compute eigenvalues, we have taken every pixel associated with a shape boundary and its connected pixels within a window of certain size. Each pixel under processing is replaced by these eigenvalues which results in two matrices. These two matrices capture the structure of a shape. It is well known fact that the Hough transform is region based and is well suited under noise conditions. Hence, we perform Hough Transformation on these two eigenvalue based matrices to obtain compact representation of the shape and these features are matched using Euclidean Distance. We have performed decision level fusion of proposed approach with blockwise binary pattern (BBP) to enhance the classifier accuracy. Extensive experimental results on the publicly available shape databases namely, Kimia_99 and Kimia_216 and MPEG_7 data sets demonstrate the accuracy of the proposed method. The results of the experiments exhibit the success of proposed approach, in comparison with well-known algorithms from the literature.

Keywords: Eigenvalue · Hough Transform · Blockwise binary pattern
Earth movers distance · Histogram matching · Euclidean distance
Decision level fusion · Shape representation · Shape classification

1 Introduction

The research work in Computer vision and Machine Intelligence has lead to significant advances in most of the automated intelligent systems. The main component in intelligent system is representation of the objects. Several algorithms have been proposed to represent objects effectively. These algorithms take color, texture, spatial layout and shape as the feature. The shape based methods play a major role and are popular in most of the image analysis application due to their resemblance to human perception of recognition. The applications of shape based descriptors include detection of malignant tissues in medical field, to confirm the presence of all the components on a printed circuit board, analyzing the fossils in forensics by means of partial match etc.

The major challenges in shape representation and classification is incomplete boundary and boundary noise. It is a well-known fact that spectral descriptors are good enough techniques to handle challenges due to noise. Hence, we have attempted to explore Spectral descriptors. There are many spectral transforms available in the literature which includes Fourier transform [29], Hough transform [11], Wavelet transform [28] and Generic Fourier transform [30], Radon transform etc. We have explored Hough [11] based descriptor for the accurate representation and classification of the shapes. The problem of representation and classification of a shape has many challenges. Apart from geometric transformation, the existence of large interclass similarities and intra-class dissimilarities of the shape classes, nonlinear scaling, deformation, occlusion and variation in the shape due to motion of the object makes it very challenging and leads to misclassification of the shapes. Though there are algorithms which capture local information [27], global information, structural [5, 19, 24] and regional information [3, 7, 10] still those methods fail to reach the accuracy needs of the present-day system. There is a class of algorithms which does the feature fusion or decision level fusion [20–22] of one or more features to arrive at the better accuracy. We attempted to do the fusion of Hough transform and binary based features to enhance the accuracy.

2 Proposed Approach

This paper presents a shape descriptor namely, Eigenvalue Analysis with Hough Transform (EHough) and a decision level fusion of EHough and Blockwise Binary Pattern as the combined classifier.

2.1 Eigenvalue Matrix Computation

The eigenvalue matrix is computed for the shape boundary. The small eigenvalue and large eigenvalue are computed for every pixel on shape boundary. We have taken every boundary pixel and its $N - 1$ connected pixels within a window of certain size, in order to compute eigenvalues. The positions are recorded for each pixel under processing and the pixels connected to the pixel under processing results in an $N \times 2$ matrix. This $N \times 2$ matrix is subjected to eigenvalue analysis to obtain small eigenvalue and large eigenvalue. Each pixel under processing is replaced by these eigenvalues which results in two matrices. The structural features of shape are captured by these two matrices. Let LMat be the Matrix which is computed by considering the largest eigenvalue replacing the boundary pixel and SMat be the matrix which is computed by considering smallest eigenvalue replacing the boundary pixel. We perform Hough transform on these two eigenvalue based matrices to obtain feature vectors of the shape.

The steps involved in the process of eigenvalue Matrix computation are as follows.

   i. The contour of preprocessed shape from the training set of standard dataset is extracted.

  ii. For each pixel p on the boundary extracted in the step i, we consider $N - 1$ neighboring boundary pixels for the processing as shown in Fig. 1.

iii. The positional parameters of these points are recorded by taking the x and y coordinates of these N pixels. This will result in N × 2 Matrix, say cMat = (x₁, y₁), (x₂, y₂),…, (xₙ, yₙ).

iv. Calculate the largest and smallest eigenvalue of cMat. Replace the pixel p by largest eigenvalue and store as LMat and replaced p by smallest eigenvalue and store as SMat.

Repeat the above process for all the boundary pixels. This will result in two matrixes LMat and SMat which are of the same dimension as that of the shape matrix. Now, Hough transform is applied on these two matrices, LMat and SMat to obtain the feature vectors. The steps involved in the process of obtaining the features are pictorially presented in Fig. 1.
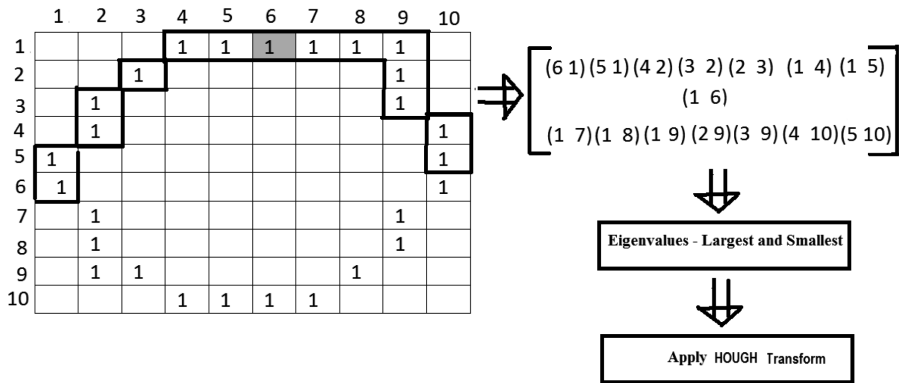
Fig. 1. The process of obtaining feature vectors

2.2 Hough Transformation

Hough transform [11] is used to extract the important structural characteristics of an image through the process of line detection. A point in an image plane is transformed to a line in a parameter space with a slope intercept form and is transformed to a curve in parameter space with rho-theta form. The process of Hough transform is explained very briefly in the following session. Let (x, y) be a point in an image plane. Let us consider n − 1 points which are collinear with (x, y). When these points are transformed to parameter space with slope intercept form, we get a set of concurrent straight lines. The problem with this form is, in case of vertical lines the slope m tends to infinite.

In the parametric space with rho-theta plane these straight lines will be transformed to a set of sinusoidal curves is presented in Fig. 2.

The rho-theta plane is subdivided into accumulator cells, with (ρ_{max}, ρ_{min}) and (Θ_{max}, Θ_{min}) being the expected ranges of $\rho\Theta$ plane. In this plane, each point on image plane that is transformed to a line in parametric space with slope-intercept form will have a specific accumulator cell as shown in Fig. 3.
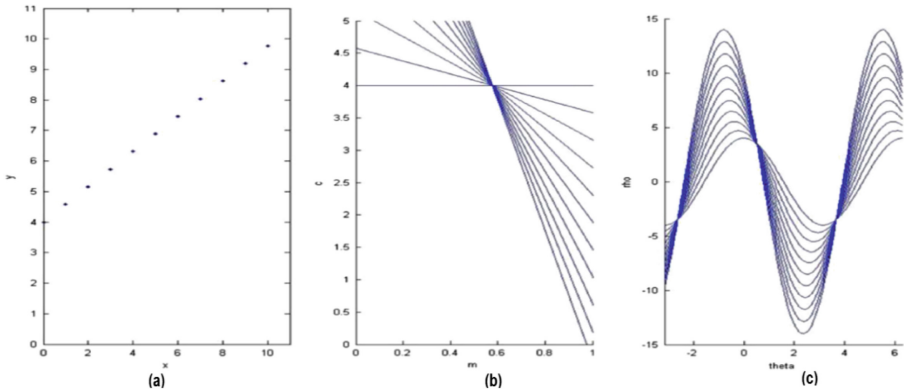
Fig. 2. (a) A set of collinear points in image plane; (b) concurrent straight lines with a slope intercept form in parametric space; (c) a set of sinusoidal curves in $\rho\Theta$ plane, image curtesy [15]

The accumulator cells are initialized to zero. The k collinear points lying on a line $x\cos\Theta j + y\sin\Theta j = \rho_i$ are transformed to k sinusoidal curves with intersection point at (ρ_i, Θ_j) in the parametric space. The Θ is incremented and the value for ρ is computed. This is the value for the accumulator cell (ρ_i, Θ_j).
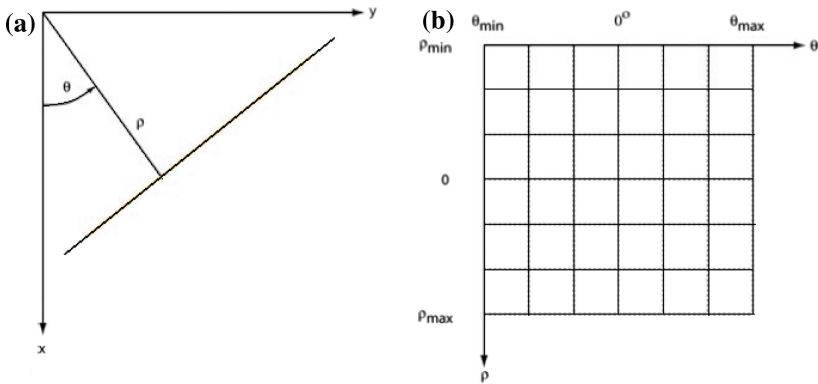
Fig. 3. (a) Line representation; (b) accumulator cells in $\rho\Theta$ plane

The steps involved in the process of finding Hough Transformation of the image is given below.

   i. Set all cells equal to zero
  ii. Repeat the following steps for every (x_k, y_k)
 iii. Let Θ = every subdivision on the Θ-axis
 iv. Calculate $\rho = xk\cos\Theta + yk\sin\Theta$
  v. Round off ρ to the nearest allotted value on the ρ-axis
 vi. Increment accumulator cell (ρ, Θ) with 1.

The effect of rotation and scaling due to Hough transform on some sample image is shown Figs. 4 and 5 respectively.
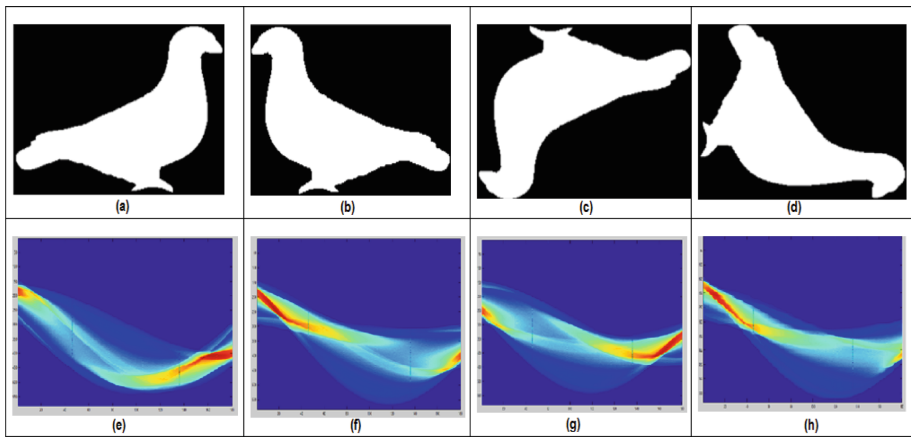
Fig. 4. Variation in Hough Transformed image due to rotation (a) original shape; (b) flipped shape (c) rotated by 180°; (d) rotated by 90°; and (e), (f), (g) and (h) are Hough Transformed images of (a), (b), (c) and (d).

2.3 Feature Extraction by Hough Transform

The dominant features are obtained from the shape using Eigen based Hough transform is presented below.

In the pre-processing step, a rectangular box is fitted to the given shape to obtain only relevant part of the shape.

   i. The shape boundary is obtained. For each boundary pixel, 15 neighboring boundary pixels are considered and the large and small eigenvalues are computed.
  ii. The matrices LMat and SMat are obtained corresponding to large and small eigenvalues respectively.
 iii. For each matrix apply Hough transform and obtain the coefficients of the Hough transform.
 iv. The feature vectors of LMat are combined with feature vectors of SMat to form resultant feature vector.

Repeat the above process for every training shape and store the feature vectors in the knowledge base for the further processing. The Fig. 6 presents the various steps involved in obtaining the feature vector.

Size 150 x 250

Size 250 x 350
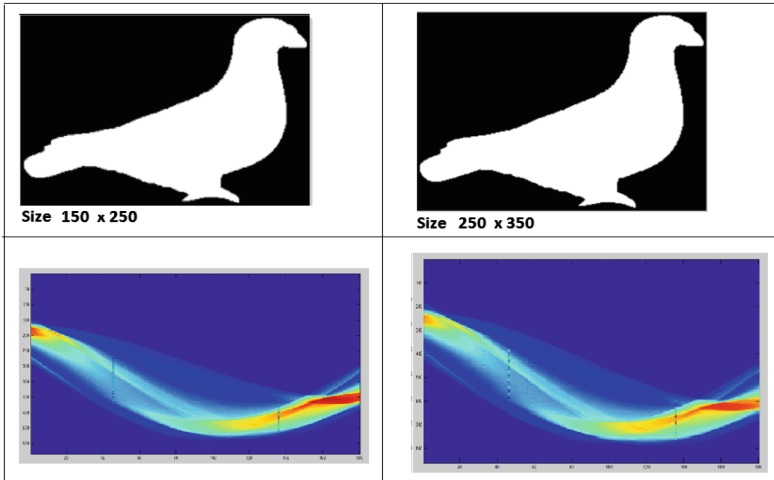
Fig. 5. Variation in Hough Transformed image due to Scaling.

(a)

(b)

(c)

(d)

(e)
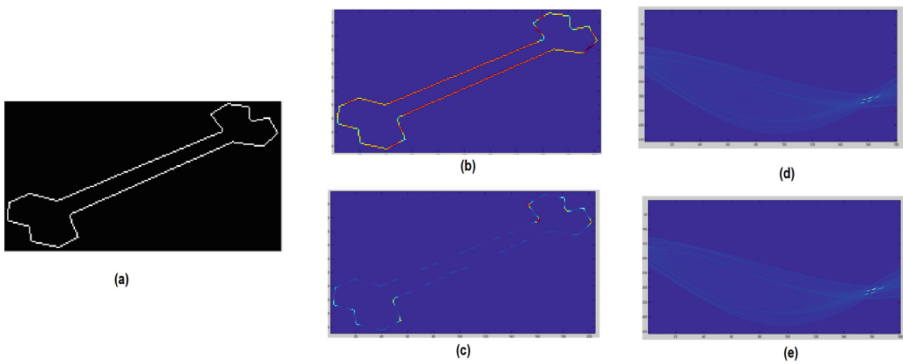
Fig. 6. The process of obtaining feature vectors; (a) the boundary of a given shape (b) and (c) large Eigen matrix and small Eigen matrix respectively; (d) and (e) the Hough transform of (b) and (c) respectively.

The process of Blockwise binary pattern based feature extraction is done as follows. The shape is considered as 3×3 blocks. Let x be the central pixel under processing. In each block of 9 pixels, we compute a new value for the x based on other surrounding 8 neighbors. The process of computing the new value for x is explained in detail in our paper [17]. The 8 neighboring pixels are traversed in clockwise direction

forming a binary stream of 8 bits as shown in the Fig. 7. The decimal equivalent of the binary stream is computed. This value is rotation variant and position dependent. In order to make it position independent and rotation invariant, one-bit right shift operation is performed on the binary stream then we get another decimal value equivalent to this binary stream. The one-bit right shift operation is performed and the decimal equivalent of binary stream is recorded. The process is repeated till we get 8 decimal equivalents of binary stream. The minimum value which is rotation invariant will be the computed value for x. The 3×3 block is replaced by single pixel x with the computed value. This process is iterated for every 3×3 neighborhood in the training shape. Now the size of the new shape will be reduced to one third of the original shape size, thus reducing the number of feature values to be processed. The shape matrix now consists of decimal values. This is converted to a 10-bin histogram and is stored for the further processing. Every shape in the dataset undergoes similar procedure in order to yield BBP feature vectors.
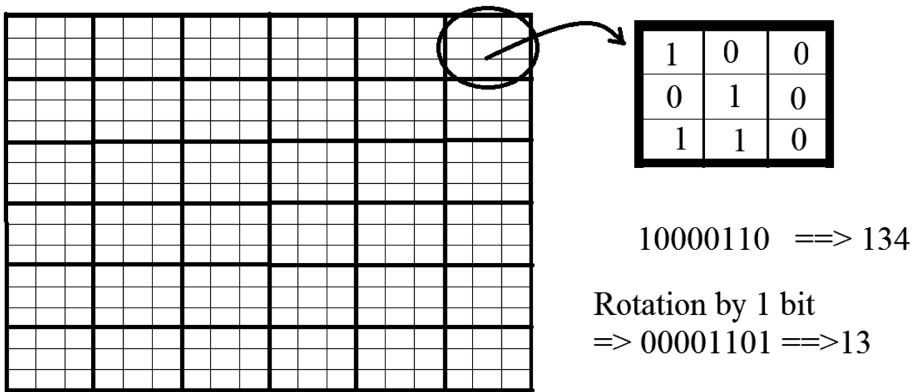
Fig. 7. Block based binary pattern on a shape.

2.4   Classification

We have used decision level fusion strategy to combine EHough and BBP based feature vectors. The Euclidean distance and Earth Movers Distance [18] metric is used to match EHough and BBP features respectively. Let D_H and D_B be the distance matrices obtained after matching test samples with training samples using EHough and BBP respectively. The decision level fusion to compute the resultant matrix is done as follows.

$$DR = D_H + \beta\,D_B \qquad (1)$$

A set of experiments are performed on given standard dataset to fix the value of for β that dataset.

For training and testing process we have used leave one strategy. Every shape in dataset is compared with remaining N − 1 shapes in the dataset to obtain distance matrix. If there are N shapes in the dataset then we get N × N distance matrix with main diagonal zero representing self-match.

3   Experimental Results and Discussions

The Experimental results on MPEG_7, Kmia_216 and Kimia_99 are presented here. The Bull's eye score, Top N retrieval and Precision-Recall graphs are used as the evaluation metrics. The Bull's eye is calculated by considering the top T retrievals, where T = 2 * Number of class samples. The leave one strategy is used for testing.

The experimental results on Kimia_99 dataset are presented below. The retrieval results obtained due to the proposed method is shown in Table 1. In addition, we have also made a comparative study with similar approaches and presented in Table 2.

Table 1. The top 10 closest matching shapes on Kimia's 99 dataset.

Closest retrievals	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	Sum
Proposed EHough	99	98	95	95	88	80	84	84	71	62	856
Proposed EHough+BBP	99	99	99	96	95	93	93	87	84	76	921

Table 2. The top 10 closest matching shapes on Kimia's 99 dataset – A comparison

Closest retrievals	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	Sum
SC [5]	97	91	88	85	84	77	75	66	56	37	756
CPDH+EMD (Eucl) [24]	96	94	94	87	88	82	80	70	62	55	808
CPDH+EMD (shift) [24]	98	94	95	92	90	88	85	84	71	52	849
Gen Model [26]	99	97	99	98	96	96	94	83	75	48	885
Learned manifold [6]	99	99	98	98	98	96	95	89	80	65	917
Proposed EHough+BBP	**99**	**99**	**99**	**96**	**95**	**93**	**93**	**87**	**84**	**76**	**921**

The results of Kimia_216 dataset is presented below. The retrieval results obtained due to the proposed method is shown in Table 3. We have also made a comparative study with similar approaches and presented in Table 4.

Table 3. Top 11 closest matching shapes on Kimia'-216 dataset.

Closest retrievals	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	Sum
Proposed EHough	211	209	208	201	195	197	184	169	174	147	136	2031
Proposed EHough+BBP	215	215	211	207	205	204	203	194	185	184	157	2180

Table 4. Top 10 closest matching shapes on Kimia-216 dataset – a comparison

Closest retrievals	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	Sum
SC [5]	214	209	205	197	191	178	161	144	131	101	78	1809
CPDH+EMD (Eucl) [24]	214	215	209	204	200	193	187	180	168	146	114	2030
CPDH+EMD (shift) [24]	215	215	213	205	203	204	190	180	168	154	123	2070
Proposed EHough	**211**	**209**	**208**	**201**	**195**	**197**	**184**	**169**	**174**	**147**	**136**	**2031**
Proposed EHough+BBP	**215**	**215**	**211**	**207**	**205**	**204**	**203**	**194**	**185**	**184**	**157**	**2180**

The results of MPEG_7 dataset is given below. It consists of 1400 shape from 70 classes with 20 shape samples. The Fig. 8 shows different class samples of MPEG_7. The class-wise retrieval rate is presented in Fig. 9 and is also given in tabular form in Table 5. The retrieval rate in the form of Bull's eye score on MPEG_7 is presented in Table 6. We have also listed the results of some well-known methods in literature. The Top-12 closest retrievals are also presented in Table 7.
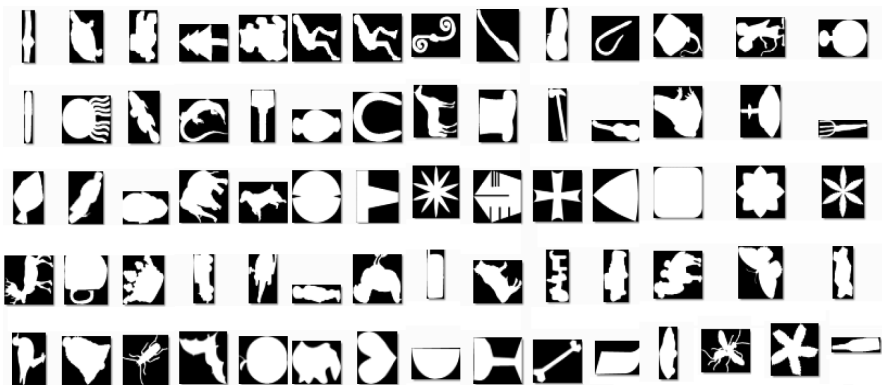
Fig. 8. The 70 different classes in MPEG-7 dataset.

The Precision-Recall graph of well-known method IDSC and our approach is presented in Fig. 10, which clearly shows the success of proposed approach. The performance of the proposed eigenvalue based Hough Transform as a standalone method as well as a combined approach integrated with BBP for various shape datasets is shown in Table 8.

Fig. 9. The retrieval rate for 70 classes in MPEG_7.

Table 5. The retrieval rate for 70 classes in MPEG_7.

Octopus	Pencil	Device0	Device6	Classic	Device2	Lizard	Elepht	Device	Bird
0.4425	0.4625	0.47	0.54	0.545	0.5725	0.615	0.6175	0.635	0.6625
Device7	Spoon	Device4	Spring	Chicken	Horse	Beetle	Turtle	Fork	Device3
0.735	0.7375	0.7425	0.78	0.795	0.7975	0.88	0.8225	0.845	0.8625
PerCar	Hat	Watch	Device5	Tree	Cattle	Camel	Fly	Misk	Cellphoe
0.8875	0.8925	0.895	0.895	0.905	0.9075	0.91	0.915	0.93	0.9375
Jar	HorseShoe	Dog	Deer	Fish	Seasnake	Key	Device9	Device8	Bat
0.94	0.94	0.94	0.945	0.945	0.9525	0.9625	0.9675	0.975	0.98
Shoe	Butterfly	Comma	Ray	Crown	ImFish	HCircle	Pocket	Bell	Bone
0.9825	0.9875	0.99	0.99	0.995	0.995	0.995	0.9975	0.9975	1
Glass	Heart	Apple	Bottle	Brick	Car	Carriage	Child	Chopper	Cup
1	1	1	1	1	1	1	1	1	1
Face	Flatfish	Fountain	Frog	Guitar	Hammer	Rat	Stef	Teddy	Truck
1	1	1	1	1	1	1	1	1	1

Fig. 10. Precision-Recall graph of IDSC and proposed approach on MPEG_7.

Table 6. The Bull's eye Score in comparison with few well-known methods on MPEG_7.

Approach	MPEG-7
Proposed EHough+BBP	**88.73**
Learned manifold [6]	88.52
Two Strategies [25]	88.39
Aspect shape context [13]	88.30
Hierarchical Parts [16]	88.30
Shape tree [9]	87.70
TAR+Shape complexity global [1]	87.23
TAR+Shape complexity [2]	87.13
SC+DP [4]	86.80
HPM [14]	86.35
Symbolic representation [8]	85.92

Table 7. Top 12 closest retrievals on MPEG_7.

Retrieval	IDSC [12]	Proposed EHough+BBP
R0	1400	1400
R1	1375	1383
R2	1342	1370
R3	1315	1345
R4	1256	1311
R5	1235	1293
R6	1209	1254
R7	1188	1228
R8	1134	1186
R9	1108	1160
R10	1084	1158
R11	1045	1118

Table 8. Bull's Eye Score on standard shape datasets.

Approach/dataset	MPEG-7	KIMIA-216	KIMIA-99
Proposed EHough	77.26	93.60	95.87
Proposed EHough+BBP	88.73	95.14	98.62

4   Conclusion

In this paper we propose eigenvalue and Hough based shape descriptor (EHough) for shape representation and classification. The structural features of the shape are captured by local eigenvalues. The Hough transform estimates the parameters of a shape from its boundary as well as region points, capturing region information and is hence robust to noise. The proposed integrated approach, which is obtained by the fusion of EHough and binary pattern, shows significant improvement in the classifier accuracy. This has been well demonstrated experimentally on Kimia99, Kimia-216 and MPEG-7 data sets. The performance is presented in the form of Bull's eye score, Top N retrievals and P-R graph demonstrating the success of our approach.

References

1. Alajlan, N., El Rube, I., Kamel, M.S., Freeman, G.: Shape retrieval using triangle-area representation and dynamic space warping. Pattern Recogn. **40**(7), 1911–1920 (2007)
2. Alajlan, N., Kamel, M.S., Freeman, G.H.: Geometry-based image retrieval in binary image databases. IEEE Trans. Pattern Anal. Mach. Intell. **30**(6), 1003–1013 (2008)
3. Bai, X., Donoser, M., Liu, H., Latecki, L.J.: Efficient shape representation, matching, ranking and its applications. Pattern Recogn. Lett. **83**(P3), 241–242 (2016)
4. Bai, X., Wang, B., Yao, C., Liu, W., Tu, Z.: Co-transduction for shape retrieval. IEEE Trans. Image Process. **21**(5), 2747–2757 (2012)

5. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. **24**(4), 509–522 (2002)
6. Chahooki, M.A.Z., Charkari, N.M.: Learning the shape manifold to improve object recognition. Mach. Vis. Appl. **24**(1), 33–46 (2013)
7. Chauhan, P.C., Prajapati, G.I.: 2D basic shape detection and recognition using hybrid neuro-fuzzy techniques: a survey. In: 2015 International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO), pp. 1–5. IEEE (2015)
8. Daliri, M.R., Torre, V.: Robust symbolic representation for shape recognition and retrieval. Pattern Recogn. **41**(5), 1782–1798 (2008)
9. Felzenszwalb, P.F., Schwartz, J.D.: Hierarchical matching of deformable shapes. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8. IEEE (2007)
10. Hasegawa, M., Tabbone, S.: Amplitude-only log radon transform for geometric invariant shape descriptor. Pattern Recogn. **47**(2), 643–658 (2014)
11. Hough, P.V.: Method and means for recognizing complex patterns. Technical report (1962)
12. Ling, H., Jacobs, D.W.: Shape classification using the inner-distance. IEEE Trans. Pattern Anal. Mach. Intell. **29**(2), 286–299 (2007)
13. Ling, H., Yang, X., Latecki, L.J.: Balancing deformability and discriminability for shape matching. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010. LNCS, vol. 6313, pp. 411–424. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15558-1_30
14. McNeill, G., Vijayakumar, S.: Hierarchical Procrustes matching for shape retrieval. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 885–894. IEEE (2006)
15. Mukhopadhyay, P., Chaudhuri, B.B.: A survey of Hough Transform. Pattern Recogn. **48**(3), 993–1010 (2015)
16. Payet, N., Todorovic, S.: Matching hierarchies of deformable shapes. In: Torsello, A., Escolano, F., Brun, L. (eds.) GbRPR 2009. LNCS, vol. 5534, pp. 1–10. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-02124-4_1
17. Pilar, B., Shekar, B.H.: An integrated approach of radon transform and block-wise binary pattern for shape representation and classification. In: 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 1976–1982. IEEE (2016)
18. Rubner, Y., Tomasi, C., Guibas, L.J.: The earth mover's distance as a metric for image retrieval. Int. J. Comput. Vis. **40**(2), 99–121 (2000)
19. Sebastian, T.B., Klein, P.N., Kimia, B.B.: On aligning curves. IEEE Trans. Pattern Anal. Mach. Intell. **25**(1), 116–125 (2003)
20. Shekar, B.H., Pilar, B.: Shape representation and classification through pattern spectrum and local binary pattern a decision level fusion approach. In: 2014 Fifth International Conference on Signal and Image Processing (ICSIP), pp. 218–224. IEEE (2014)
21. Shekar, B.H., Pilar, B.: Discrete cosine transformation and height functions based shape representation and classification. Procedia Comput. Sci. **58**, 714–722 (2015)
22. Shekar, B.H., Pilar, B., Kittler, J.: An unification of inner distance shape context and local binary pattern for shape representation and classification. In: Proceedings of the 2nd International Conference on Perception and Machine Intelligence, pp. 46–55. ACM (2015)
23. Shen, W., Wang, X., Yao, C., Bai, X.: Shape recognition by combining contour and skeleton into a mid-level representation. In: Li, S., Liu, C., Wang, Y. (eds.) CCPR 2014. CCIS, vol. 483, pp. 391–400. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-45646-0_40
24. Shu, X., Wu, X.J.: A novel contour descriptor for 2D shape matching and its application to image retrieval. Image Vis. Comput. **29**(4), 286–294 (2011)

25. Temlyakov, A., Munsell, B.C., Waggoner, J.W., Wang, S.: Two perceptually motivated strategies for shape classification. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2289–2296. IEEE (2010)
26. Tu, Z., Yuille, A.L.: Shape matching and recognition – using generative models and informative features. In: Pajdla, T., Matas, J. (eds.) ECCV 2004. LNCS, vol. 3023, pp. 195–209. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-24672-5_16
27. Wang, Z., Liang, M.: Locally affine invariant descriptors for shape matching and retrieval. IEEE Sig. Process. Lett. **17**(9), 803–806 (2010)
28. Wunsch, P., Laine, A.F.: Wavelet descriptors for multi-resolution recognition of hand printed characters. Pattern Recogn. **28**(8), 1237–1249 (1995)
29. Zahn, C.T., Roskies, R.Z.: Fourier descriptors for plane closed curves. IEEE Trans. Comput. **100**(3), 269–281 (1972)
30. Zhang, D., Lu, G.: Shape-based image retrieval using generic Fourier descriptor. Sig. Process. Image Commun. **17**(10), 825–848 (2002)

Event Data Analysis in Large Virtualized Environment

M. B. Bharath[1](✉) [iD] and D. V. Ashoka[2] [iD]

[1] EMC Software and Services India Pvt. Ltd, Bangalore, India
Bharath.Basavarajappa@emc.com
[2] JSS Academy of Technical Education, Bangalore, India
dr.ashok_research@hotmail.com

Abstract. Monitoring large virtualized datacenter is a daunting task. Around 3 to 3.5 million of events/alerts are received annually from 90+ customers, managed by our organization. Global command centers process these events in real-time and takes an appropriate necessary action for each of those alerts. One of the challenging tasks is to extract useful analytical details out of this large dataset. Attempts to run analytics on this event data poses a problem like heterogeneous source, unstructured data without any centralized repository to collect these events etc. These three issues are the same classic issues face by any "Big data" analysis. To address this issues a novel unified framework is built. This paper focus on big data problems and the solution proposed to address event data analysis in large virtualized environment as a use case. Along with this detailed design and implementation of the proposed solution with resulting report details are also discussed.

Keywords: Big data · Data analytics · Virtualized datacenter
Event monitoring · Events reporting

1 Introduction

The datacenter managed service business comprises two primary task one managing customer day-to-day administrative task and the second one is monitoring datacenter elements for performance and fault events. Most of the production setup will have adequate redundancy to handle any single point of failure. Even then most of the failures need to assessed and rectified in a time bound manor to avoid any business impacts. In this context monitoring becomes critical aspect of the service offering.

Most of the datacenters monitoring is done at three distinct layers as listed here,

- Storage hardware's layer
- Virtualized infrastructure/Hyper Converged Infrastructure
- SAN fabric connecting to storage array to virtual infra.

On a day to day basis it's normal to collect massive amount of details from datacenter elements. Most of these data is processed and any performance and fault issues will be report as an event to Global Command Center (GCC). GCC is 24 × 7 operations which monitors any faults/events and take corrective action based on the

situation. On a per day basis it is common to get around 7000 to 10,000 alerts in our GCC; which gets acted by our L0/L1 team. Apart from this initial action not much is done with event data. There is no post-processing analysis on this massive amount RAW event data. These events are from heterogeneous source, like multi-vender storage arrays, converged and hyper converged infrastructure etc. Each with their own property template and values. Any analysis of such varying dynamic data is similar to that of "Big Data" analytics.

Big Data refers to large amount of unstructured data from variety of source – which outpace your current ability to process it by existing application or processes. In recent years there has been a dramatic explosion of unstructured data in digital world. IDC and EMC report state that world data will reach 40 Zeta bytes at rate for double every two years [1]. This massive amount of data pose good amount of technical changes. At the same time it also opens door to new opportunity in digital world which cannot be foreseen before. Because of this every enterprise is trying to harness this new asset through data analytics, by build capabilities to collect, organize and process and then extract useful details from this massive data. With this pretext applying analytics on top of large virtualized infrastructure raw event data will give insights that were not perceived before. This analysis can help in tackling the issues like noise suppression and identification of problematic assets/device within the environment. Coupling this with machine learning will enable automated dynamic root cause analysis and isolation of devices/systems from symptom pool.

2   Related Works

Continued innovation in the field of processor technology and adaptation of Internet of things (IOT) has produced the exponential growth in digital data. This massive amount of digital data provides good opportunities to extract additional useful information, hence creating a large Big data science field. John Marshey made a first reference to 'Big Data' in his talk in Silican Graphics [2]. There were lot of activates in early 80s and 90s on data mining that eventually transformed into big data analytics as it matured. Weiss and et book [3] on data mining had a details on 'Big Data' problems and some of early methods used for those use-cases. Even though data mining was a popular topic in academia; more research papers started appearing after year 2000 onwards. One of such early paper was "Big Data: Dynamic factor model for macroeconomic measurement and forecasting" [4]. Diebold and team used data analytics to forecast the economic growth parameters using analytical model. In today's world major chunk of data comes from the Smart phones users and these individual are creating more digital content than any other organization. Companies like Facebook, google, flipkart, amazon, alibaba, twitter, Netflix etc., use these vast amount of data to identify any user patterns or customer preferences that can be identified. So that they can have a targeted sales pitch to the individual customer based on their preferences. On a similar line MIT lab [5] has done the research on people behavior pattern on social media. Fayyad [6] has given some interesting number on digital data usage and their growth. She also tried to formally define the term 'Big Data' in her talk in BigMine workshop in 2012. First reference to 3V's can be traced to Laney [7] and later this has been enhanced and adopted with

additional 2V's by Gartner [8]. Abhilash explains the different challenges faced during virtualized environment monitoring [9] in their publication. As big data filed grow, its presence and use can be felt across multiple fields [10, 11]. Each trying to address different type of technical problem or Business use case.

3 Defining Big-Data

In beginning Big data is defined through 3V's, in early 2001 by Gartner. Later as it matured it add two more additional dimension to it. Now it's very common to articulate big data problem in terms of 5V's, they are identified as Volume, Value, Veracity, Variety and Velocity. Volume is the primary driver of big data which limit your ability to comprehensively handle the incoming data. If you couple that with multiple source of inputs, along with varying degree of live stream-data will complicate the solution. This multi-source and dynamic stream-data is represented by Variety and Velocity. Any big data problem allows a possibility to explore new dimensions of looking at existing data, at the same time it poses good amount of technical challenges. Following Fig. 1 abstractly represents the five dimensions of Big data analytics.

Fig. 1. Five dimensions of any big data problem

4 Problem Definition

Our managed service business handles around 90+ customers virtualized data center operations. This contract primarily include both management and monitoring. Where management includes administrative tasks like adding new device, migrating data from

old device. Decommissioning of old devices, updating redundancy path, making provision for new expansion of disk utilization and so on. Most of these task are done based the request/ticket raised by the customer. Whereas monitoring includes constant validation of performance alerts & fault events – using different enterprise monitoring tools like EMC SRM, DPA and NimSoft etc. Which collects the device performance data and fault events. Monitoring for fault is one of the major task, any failure/faults need be restored within the stipulated time period. So that component failures will have minimal or no impact of the business operations. Most of the accounts have strict SLA agreements with availability number of 99.99%.

In order to meet this challenge, different operational process have been built, using centralized global command center. Which receives faults/failure events from multiple customers (and locations). All these are processed and normalized and then forwarded to event management software called Smarts. This smarts live console is monitored by the GCC on 24 × 7 bases and each event will be forwarded to appropriate team for future action. Apart from this catch and dispatch action, these event data is not used for further analysis. Following challenges were faced, when this data is used for any analytical processing. They can be listed as follows.

- Heterogeneous source
- Large dynamic streaming data
- Unstructured event data.

Heterogeneous source: We get the events from multiple application and devices. Most of them from different venders like EMC, IBM, HP, Cisco, Brocade etc. Each monitoring application used by a specific vender has its own template for event data. Most of the time faults/events from monitoring application and devices are converted as SNMP traps; which are forwarded to centralized processing center and then used by GCC team. Due to nature of the business – heterogeneous source is implicit fact that cannot be avoid in this scenario.

Large dynamic streaming data: Most of the events are near to real time. Any failure on customer site will be reported to command center console with 3 to 4 min. Since it's based on the SNMP traps; it's hard to predict the flow size. Even though we get around 7 K to 10 K alerts per day; a bust of 5 K alerts with an hour is not uncommon. This unpredictable dynamic flow of events makes it had to build any conventional RDBMs system to deal with this RAW data. We need a stable system which can handle the sudden surge of data along with relatively acceptable performance level.

Unstructured event data: Here is sample example of an event from Smarts archive log (as shown in the Fig. 2). This is an example of VNX array Link Down alert. It is interesting to note that this archive log entry has a unstructured data and details are in property value pairs. Left side of the equal is property and right-hand side is its value. This has to be converted into a format that is more manageable, like table format or.csv structure. Hence enabling the data feed to any analytics tools.

```
"April 21, 2016 2:58:44 AM EDT" ICS_Notification::NOTIFICATION-EMC_20VNX-Block-Port_DC__SHAP__1128-F
{
    ClassName="EMC VNX-Block-Port";
    InstanceName="DC_SHAP_1128-Port-4-LinkStatus";
    EventName="VNX-Block PortLink StateChange";
    ClassDisplayName="EMC VNX-Block-Port";
    InstanceDisplayName="DC_SHAP_1128-Port-4-LinkStatus";
    EventDisplayName="VNX-Block PortLink StateChange";
    SourceDomainName="Trap Processor";
    Active=FALSE;
    OccurrenceCount=1;
    FirstNotifiedAt="April 20, 2016 10:10:13 PM EDT";
    LastNotifiedAt="April 20, 2016 10:10:13 PM EDT";
    LastClearedAt="April 20, 2016 10:57:10 PM EDT";
    LastChangedAt="April 21, 2016 2:54:48 AM EDT";
    IsRoot=TRUE;
    Acknowledged=TRUE;
    EventType="DURABLE";
```

Fig. 2. Unstructured RAW event data

5   Data Lifecycle and Proposed Solution

Any data analysis process goes through four major stages as shown in the Fig. 3. Big data processing needs contextual management, since problem or business objective that you try to address is directly related to, how that data is contextualized (useful information is extracted and categorize). Most of the time it requires building on existing capabilities. That is to augment the existing process to enable the business objective. In our case using existing old event data collection process and adding new feature like tagging each custom data with information like customer name and data center was done before starting analysis process. Also normalizing data from different sources & create centralized repository for raw event data was built for this purpose.

5.1   Data Acquisition and Warehousing

This phase logically have two intervened steps one, building master data repository and second Data filtering/reduction. One critical part of building robust Data Warehousing platform is to consolidate data across several interrelated sources. Hence creating an effective repository of master data, becomes relatively easier. This centralized master data can provide coherent information access to data analytics process. On the other hand data reduction is science and art that need good amount of research and analysis. Building effective intelligent process that will filter out the data such way that no valuable information is missed once it's in master repository. Normally this filtering is done in real-time to avoid expense of storing all data and then applying filtering.

5.2   Data Extraction and Structuring

Data extraction is a process of collecting relevant information from multiple disconnected sources and then normalize it to a format that can be consumable by analysis process. Two main type of data that are collected are – structured monitoring data and

unstructured log data. Structured data is format represented in the ordered fashion which can be easily consumable. Like rows and columns – where every entry is a unique entity for that table of rows and columns. This format is relatively easy to store and analyze. These are regular traditional transactional data; which constitutes only 10 to 15% in today's world. On the other hand unstructured data is a free form, behavioral in nature with multiple properties associated with it. There is no predefined format or template and doesn't adhere to traditional structured data format. It's heterogeneous in nature and comes with dynamic variables in multiple formats. Unstructured data is growing in much faster phase, contributing more than 90% of today world data (IDC 2011). Our event data was unstructured format with property value pair as shown in the Fig. 2.
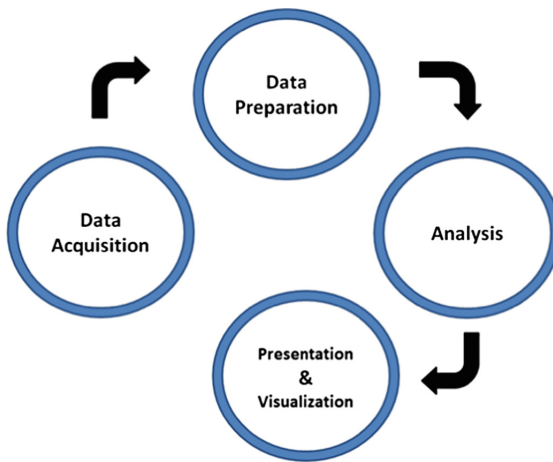
Fig. 3. Four data lifecycle phases

Most valuable information is embedded in these unstructured data. So it needs skillful art to extract, useful details for analytics. These are the hidden nuggets that are most ignored in today's world. It's very common to use ETL (Extract Transform and Load) job or transformation sequence to filter massive raw data and then extract useful details and transform it to a semi structured format which can be used by any analytical algorithms.

5.3   Data Modeling and Data Analysis

Once the repository of cleansed data is created next stage is complex iterative stages of data analysis. Data analysis is not only about extracting, identifying and correlating and interpreting scenarios. Now it's becoming very common to have fully automated framework. Which process data from multiple sources and transform it into structured template, used to extracting semantics and characteristics using advance computer intelligence. The Fig. 4, summaries this process abstractly.
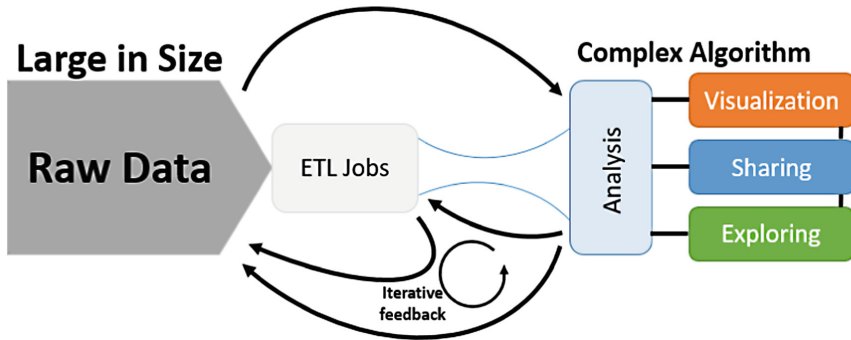
Fig. 4. Unstructured data modeling and analysis

New advance in memory processing technology are accelerating the unstructured data processing. Apache spark demonstrated this capability by sorting 1 Peta bytes (PB) of data in 234 min on 190 nodes [12]. This was thought impossible couple of years back. Availability of multiple open source tools and relatively cost effective commodity hardware are giving big boost to the advanced data analytics fields of research.

5.4 Data Interpretation

True success of any data analytics is measured through how well it represent the value of your analysis. Irrespective of complexity involved in extracting details, is of no used, unless it is represented in proper format. Which convey the results effectively. Using simple structured format like tables, may not be suitable in most of the cases. Using more interactive visual representation will be more effective.

6 Implementation Details

The framework built to automate the event data analytics, collects the RAW events details from two production servers. Which are configured in active passive configuration. This extracts 26 properties from every event object. Which are then converted to input.csv file. This input file is pre-processed by removing any duplicates or incomplete entries. On successful filtering of this data will be feed to the SRM collectors. Which reads this sanitized input data and convert it to a time series data. Which are used for building preconfigured report with default templates. These reports are configured through scheduled jobs to generate daily, weekly or monthly reports. Figure 5 shows abstracted high level data flow in this framework.
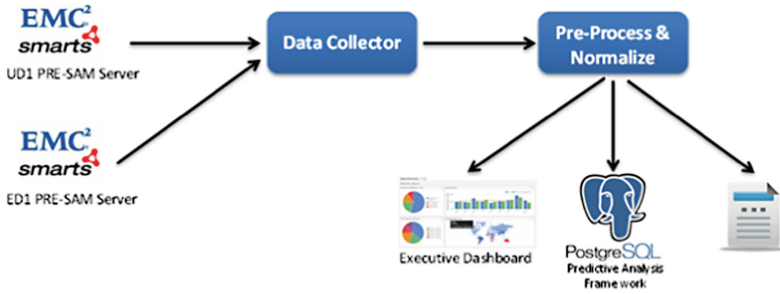
Fig. 5. Normal event data flow in analytics framework

Algorithm used for these calculation:
Let's use the following symbols to represent property of raw event data

$\xi_{n_i} = EventNameofthealert$
$\xi\zeta_i = EventClassName$
$\delta_{n_i} = DeviceName$
$A_i = AccountNameorCustomerName$
$S_i = Severityvalue$

Aggregated trend is calculated using formula 1, show in trend report result Fig. 6.

$$\alpha_i = \sum_{n=0,i} \left(\xi_{n_i} + \beta_i \right) \tag{1}$$

Aggregated severity count is calculated using formula 2

$$\beta_j = \sum_{n=0,i} \left(\xi_{n_i} + S_i \right) \tag{2}$$

Number of top event count at account level is calculated using formula 3, as shown in the report Fig. 9.

$$\gamma_k = \sum_{n=0,i} \left(\xi_{n_i} + \xi\zeta_i + A_i \right) \tag{3}$$

Event count at device level are calculated using following formula 4

$$\theta_l = \sum_{n=0,i} \left(\delta_{n_i} + \xi_{n_i} + A_i \right) \tag{4}$$

7   Test and Results

Following section shows couple of automated reports of our event data analytics. Figure 6 shows the alert trending over period of months. Whereas Fig. 7 shows the percentage of alert count from different accounts (or customers). Due to space constrains, we have shown only subset of original reports build using this research work. Figure 8 shows the daily aggregated alert count trends in a month. Whereas Fig. 9 shows the SRM application event distribution across event types.
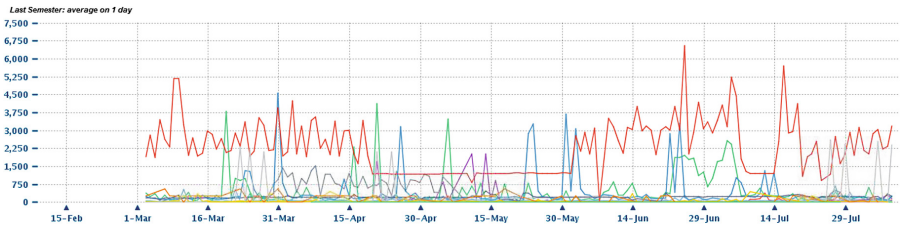
Fig. 6. Alert trending data for last 6 months
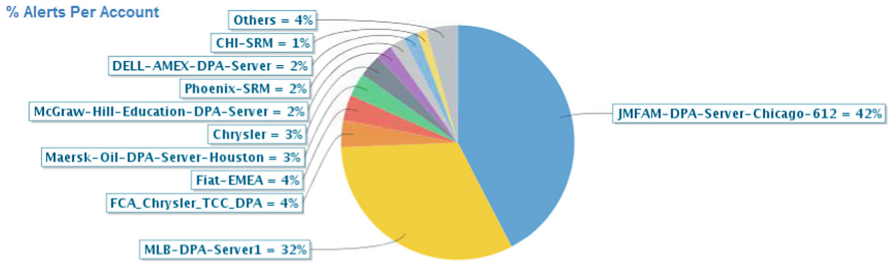
Fig. 7. Percentage of alert count from different sources
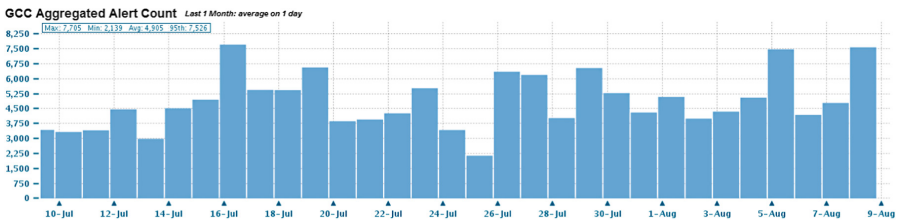
Fig. 8. Aggregated alert trend in last one month.
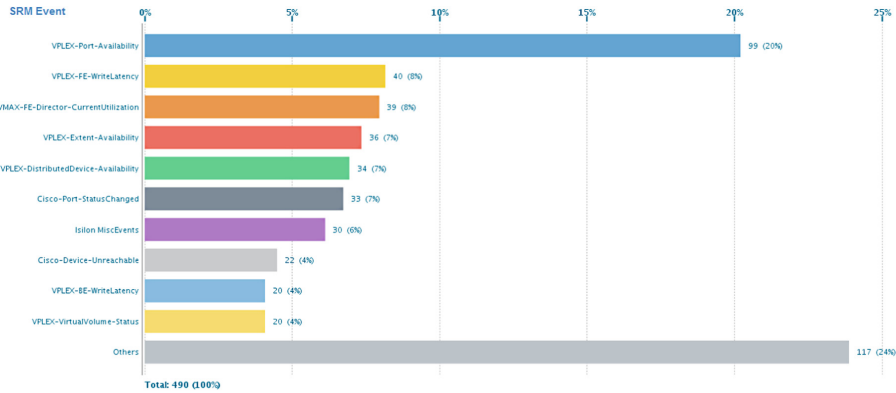
Fig. 9. Top 10 noise alerts from SRM application

8   Conclusion

Monitoring large datacenter is a complex afire, on top of it a very large heavily virtualized setup makes it, even more arduous task. It's very normal to get multiple source of inputs and alerts in large environment. As any problem will have its traces across multiple elements within a datacenter. Collection of these symptoms and processing them for extracting usefully administrative inputs are interesting research topic. Our attempt was regarding this path – running analytics on raw event data collected during datacenter management and finding the useful details like number of alerts, repeated noise alerts, alerts from a device, alert trending patterns etc. These will help in identifying the generic patterns, which will help in debugging complex performance issues. Event analytics gives additional level of insight which were never utilized otherwise. Our attempt to build analytic framework on top of event management has given good value addition to customer by identify problematic device and helping them to predict the event flow pattern in near future. One can easily conclude that big data analytics on top of existing data in large datacenter is going to norm in future.

References

1. Gantz, J., Reinsel, D.: Extracting value from chaos. IDC iView **1142**, 1–12 (2011)
2. Diebold, F.: On the origin(s) and development of the term "Big Data". In: Pier Working Paper Archive 12-037, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania (2012)
3. Weiss, S.M., Indurkhya, N.: Predictive Data Mining: A Practical Guide. Morgan Kaufman Publishers Inc, San Francisco (1998)
4. Diebold, F.: "Big Data" dynamic factor models for macroeconomic measurement and forecasting. In: Discussion Read to the Eighth World Congress of the Econometric Society (2000)
5. Petland, A.: Reinventing society in the wake of big data. Edge.org (2012). http://www.edge.org/conversation/reinventing-society-in-the-wake-of-big-data

6. Fayyad, U.: Big Data Analytics: Applications and Opportunities in On-line Predictive Modeling (2012). http://big-data-mining.org/keynotes/#fayyad
7. Laney, D.: 3-D data management: controlling data volume, velocity and variety. META Group Res. Note **6**, 70 (2001)
8. Gartner. http://www.gartner.com/it-glossary/bigdata
9. Abhilash, C.B., Ashoka, D.V.: A survey on operating system virtualization methods and challenges. i-Manag. J. Inf. Technol. **5**(1), 28–33 (2016)
10. Intel. Big Thinkers on Big Data (2012). http://www.intel.com/content/www/us/en/bigdata/big-thinkers-on-big-data.html
11. Aggarwal, C.C. (ed.): Managing and Mining Sensor Data. Advances in Database Systems. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-6309-2
12. First Public Cloud Petabyte Sort (2014). http://databricks.com/blog/2014/11/05/spark-officially-set-a-new-record-in-large-scale-sorting.html

Noisy Speech Recognition Using Kernel Fuzzy C Means

H. Y. Vani[(⊠)] and M. A. Anusuya

Sri Jayachamarajendra College of Engineering, Mysore, India
{vanihy,anusuya_ma}@sjce.ac.in

Abstract. In the area of voice recognition, soft computing technique is a prominent method to identify and cluster speaker variability's in the speech signal. But whenever the signal is convoluted by a noisy signal standard FCM method fails to give the good results. To overcome this, Kernel FCM (KFCM) is used in this paper. PCA helps in reducing the features of convoluted signal. The recognition results are compared with and without applying PCA using KFCM function and the same is presented for word recognition rate.

Keywords: Cepstral Coefficients (MFCC) · Soft computing technique
Additive noise · Convolution noise · Computation Complexity (CC)
Principal Component Analysis (PCA) · Kernel Fuzzy C Means (KFCM)

1 Introduction

For current Wi-fi environment speech recognition is used to instruct and answer systems in a hand free environment. It is a method to send instant message, to keep real-time transcripts during conversations and hence acts as a linguistic user interface. Many application services are under development to identify a word that is spoken in mind. Hence speech recognition is a link between a mind and a computer than conventional methods. Hence this paper is towards increasing the robustness of the noisy speech signal for the recognition application in the noisy environment.

Feature extraction and feature selection phases are important for clustering. Feature extraction is done using MFCC for a signal. The signal is corrupted by the noise introduced by surrounding or by the components used during speech recording process. High number of co-efficient results in consuming high memory and processing time by reducing recognition rate. To increase the recognition rate optimal features are necessary. To obtain the optimal features PCA is performed.

The paper presents The literature in Sect. 2. Sections 3 and 4 presents database and method of implementation. In Sect. 5 results are discussed. Last part of the paper presents conclusion and future enhancement.

2 Literature Survey

Automatic Speech Recognition System (ASR)
Figure 1 shows the traditional Automatic Speech Recognition [1] system. In the first phase optimal features are selected by applying feature extraction and feature selection

algorithm at the front end. In the second phase of the signal processing the extracted features from the front end are used to match the features with the reference model. Since this paper is towards achieving the robustness it highlights on front-end and clustering phases. A brief review of literature is presented in the below sections.
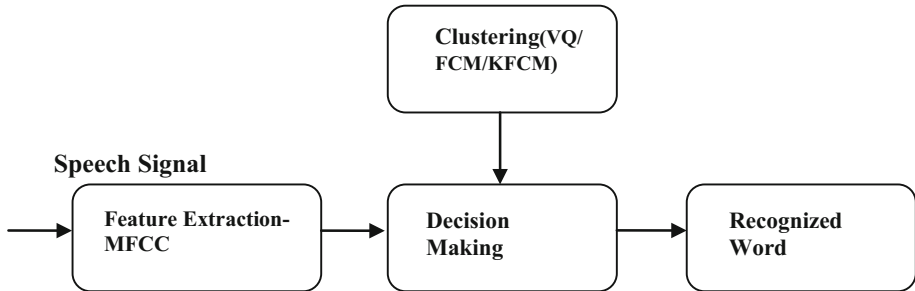
Fig. 1. Speech recognition system

2.1 Feature Extraction for Noisy Speech Signal

Computation of Cepstral Coefficients. [2] is the common technique for extracting spectral features of clean and noisy signals. The analysis is performed by applying Fast Fourier Transform (FFT) on the speech frame sequences in order to obtain parameters.

2.2 Feature Selection Method

Principal Component Analysis
[2] Uncorrelated variables for correlated features of the speech signals are identified and computed by Principal Component Analysis (PCA).This is used to reduce the uncorrelated variables.
Importance of PCA application:

- Reduced number of variables
- Decreases the size of the solution space
- Less features i.e. it reduces the feature dimension
- Selects only highest correlation features

  These steps are used to compute the components

1. The data is normalized.
2. By creating the covariance matrix Eigen values and Eigen vectors are derived.
3. Eigen values are sorted according to descending order and choose k-Eigen vectors that match to the k largest Eigen values. The value k identifies the reduced dimension of the original data.

2.3   Clustering Methods

This section presents hard clustering(VQ) and soft clustering(FCM) methods that are used for clustering procedure.

   (i) K-means [3]:

   It is an simplest method of clustering as it groups the data into 'C' clusters using Euclidian distance as the distance measure. The objective function is as shown with Eq. (1).

$$V = 1/C_i \sum_{1}^{C_i} X_i \tag{1}$$

   (ii) Fuzzy C Means (FCM) [4, 5]:

   It is an effective method for data clustering. The objective function for partitioning the data into clusters is given by

$$J_m(U, \lambda; X) = \sum_{i=1}^{C} \sum_{t=1}^{T} u_{it}^m d_{it}^2 \tag{2}$$

   (iii) Kernel Fuzzy C-means (KFCM)

   The point of belongingness is identified by Euclidian distance function in FCM but kernel function $\Phi(x)$ is used by KFCM proposed by Zhang [6, 7]. The objective function of KFCM is given by

$$J = \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij}^m \left\| \Phi(x_j - \Phi(v_i)) \right\|^2 \tag{3}$$

Where ϕ is an implicit nonlinear map, and

$$\left\| \Phi(x_j - \Phi(v_i)) \right\|^2 = K(x_j, x_j) + K(v_i, v_i) - 2K(x_j, v_i) \tag{4}$$

Where $K(x, y) = \Phi(x)^T \Phi(y)$ is the inner product kernel function.
If we adopt the Gaussian RBF kernel i.e.
$\frac{-|x-y|^2}{\sigma^2}$, then k(x, x) = 1. The simplified objective function becomes

$$J = 2 \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij}^m (1 - K(x_i, v_i)) \tag{5}$$

where u_{ij} is membership function and v_i is the cluster center. These variables are computed using the following formula.

$$u_{ij} = \frac{\left(1 - K\left(x_j - v_i\right)\right)^{-1/(m-1)}}{\sum_{k=1}^{c}\left(1 - K\left(x_j - v_k\right)\right)^{-1/(m-1)}} \tag{6}$$

$$u_i \frac{\sum_{k=1}^{N} u_{ij}^m K\left(x_j, v_i\right) x_j}{\sum_{k=1}^{N} u_{ij}^m K\left(x_j, v_i\right)} \tag{7}$$

The loop gets terminated when $\{J(i) - J(i - 1)\} < \varepsilon$, where ε is criteria for termination.

3 Database Construction

For training and testing the signals are recorded using software. The recording frequency is 8 kHz with mono channel. 10 female and male speakers signals for isolated kannada words are recorded. Each speaker uttering 15 times for one word. Totally, 15 recordings are taken for each word from both speakers. The same is used for simulation.

Two types of signals are considered i.e., noisy and clean speech signals. The signals with unwanted noise and unvoiced are preprocessed manually. Clean signals are captured in close room. The noisy signals are created by adding the babble noises for various decibels. The clean signals are convoluted with the noisy signals. The Table 1 shows the database used for the simulation.

Table 1. Speech database

Number	Kannada word	Symbol used in the paper
1	ONDU	One
2	ERADU	Two
3	MOORU	Three
4	NALKU	Four
5	AIDU	Five
6	AARU	Six
7	ELLU	Seven
8	ENTU	Eight
9	OMBHATHU	Nine
10	HATHU	Ten

4 Proposed Methodology

This section explains the procedure used to extract prominent features by feature extraction and feature selection using MFCC and PCA techniques. The reduced set of features from MFCC and PCA are transferred to another high dimension space.

After extracting the features from MFCC and PCA the selected features are mapped to high dimension space. The proposed work is carried out for various combinations of feature extraction, selection and clustering methods. They are (Fig. 2)

  i. MFCC-PCA-FCM
 ii. MFCC-PCA-FCM, MFCC-PCA-KFCM
iii. MFCC-PCA-VQ.
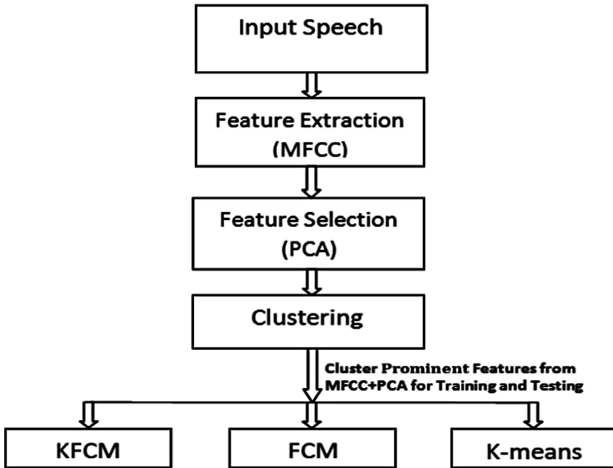
Fig. 2. Data flow diagram

5 Results and Discussions for Additive and Convolution Noise

(i) Convolution Noise

Convolution is a formal mathematical operation, just as multiplication, addition, and integration. The robustness of the speech recognition system is evaluated for all the three methods with the various noise types. Tables 2 and 3 shows recognition accuracies for different type of noise signals.

Table 2. Recognition rate for convoluted noise

Clustering method	Gender	Word recognition accuracies		
		SNR (dB)		
		15 dB	10 dB	Clean
KFCM	Male	60	55	90
	Female	50	50	85
FCM	Male	60	55	90
	Female	50	40	85
K-Means	Male	46	45	80
	Female	40	40	78

Table 3. Additive noise recognition rate

Clustering method	Gender	Word recognition accuracies		
		SNR (dB)		
		15 dB	10 dB	Clean
KFCM	Male	58	50	90
	Female	50	50	85
FCM	Male	55	50	90
	Female	50	40	85
K-Means	Male	50	45	80
	Female	40	40	78

From the tables it is clear that KFCM performs better in average for both speakers. Major observations from the above results are as listed below:

 (i)   Feature selection procedure reduces the time complexity.
 (ii) Transforming into kernel space and applying Gaussian kernel as an alternate for

       Euclidian distance measure improves the recognition accuracy.

 (iii) Since the features are put into kernel space, the non-linearity is reduced.
 (iv)   KFCM with PCA performs better for noisy speech signal

From the above observations it is clear that selecting reduced and prominent features helps in better modeling of the speech signals consuming less time and memory for training. Hence this paper is attempted mainly towards achieving good recognition accuracy.

6   Conclusion and Future Enhancements

When compared to FCM and VQ the KFCM performs better towards the increase of robustness of the speech recognition system under various noisy conditions like convoluted or additive noises. From the simulation results we have achieved 60% of the accuracy for the isolated words in average for the acoustic approach. It is clear that Gaussian kernel functions performs better under various noisy conditions.

The proposed system can be further enhanced for various types of noise signals like car noise, street noise, pink noise etc., for larger data sets. This can also be tested for speaker dependent and independent data sets for obtaining less error rates.

References

1. Rabiner, L., Juang, B.H.: Fundamentals of Speech Recognition, vol. 103. Prentice Hall Englewood Cliffs, New Jersy (1996)
2. Anusuya, M.A., Katti, S.K.: Front end analysis of speech signal processing: a review. Int. J. Speech Technol. **14**, 99–145 (2011). https://doi.org/10.1007/s10772-010-9088-7
3. Linde, Y., Buzo, A., Gray, R.: An algorithm for vector quantizer design. IEEE Trans. Commun. **28**, 84–95 (1980)
4. Tran, D.-T.: Fuzzy approaches to speech and speaker recognition, thesis submitted to University of Canberra, May 2000
5. Bezdek, C.: A convergence theorem for the fuzzy ISODATA clustering algorithms. IEEE Trans. 0162-8828/80/0100-0001
6. Wang, X., Wang, Y., Wang, L.: Improving fuzzy c-means clustering based on feature-weight learning. Pattern Recogn. Lett. **25**, 1123–1132 (2004). https://doi.org/10.1016/j.patrec.2017.02.015
7. Zhang, D.: Kernel-Based Associative Memories, Clustering Algorithms and their Applications, Nanjing University of Aeronautics and Astronautics (2004)

An Enhanced Water Pipeline Monitoring System in Remote Areas Using Flow Rate and Vibration Sensors

Praveen M. Dhulavvagol (ORCID), K. R. Ankita(✉) (ORCID), G. Sohan(✉) (ORCID), and Renuka Ganiger(✉) (ORCID)

KLE Technological University, Hubballi, India
praveen.md@bvb.edu,
ankitakirgundal4@gmail.com, sohanvg@gmail.com,
renukaganiger@gmail.com

Abstract. Currently in India and most of the other countries, if there is any leakage of water pipes in remote areas then it's difficult to trace and repair the leakage within a short span of time, it takes some amount of time usually two to three days to locate the leakage and to repair it. During this time a lot of water is wasted. In this paper we discuss the techniques to detect the leakage and locate the leak point on the pipeline. In the proposed approach we discuss efficient techniques for monitoring, detecting leakage in pipes and checking the durability of the pipes using flow rate and vibration sensors. The objective is to build a prototype model for detecting leakage in water pipeline using flow rate sensors and also to predict the durability of the pipe using vibration and pressure sensors. Data from remote sensors are collected through wireless medium and are monitored in a server. If there is any leakage in the pipelines then the server sends the sms message to the water board authorities and notifies the leak point in the pipeline system. Further the durability of pipe can be predicted using pressure sensor so that the authorities can be intimated in prior about chances of breakage in pipelines. Thus the proposed system can save huge amount of water by intimating the authorities in time. Thus the experimental result validates the effectives and reliability of the proposed system.

Keywords: Water leakage · Detection · Monitoring · Leak detection

1 Introduction

The scarcity of drinking water is one of the critical issues that everyone is facing in their day today life. According to the sources there is only 2.5% of fresh water available among which only 1.5% is used for drinking purpose. In future if proper water pipelines monitoring system is not adopted means, this water scarcity of drinking water will lead into a biggest challenge for mankind survivability. The water leakage is mainly due to poor monitoring, maintaining and aging of pipes, to prevent future water losses many techniques have been proposed for detecting leakage in pipeline which has some limitations as discussed in Table 1.

Table 1. Pipeline leakage detection techniques

Detection parameters considered	Methods followed
Acoustics	Sound produced by leak is sensed using hydrophone device
Optical	Fiber optical temperature profile sensing
Cable sensing	Electrical cables are used that when come into contact with hydrocarbons of soil
GPR (Ground penetrating radar)	Electrical radiation is used in microwave band
Vapour sampling	Vapour monitoring uses monitoring wells as testing locations for leaking product
Video inspection	Real-time event is detected using digital video inspections

In conventional water pipeline maintenance system a robust and efficient techniques lack, due to which a huge volume of water is getting wasted. So we need to overcome these techniques by adopting new latest technologies which are very robust and reliable in detection and maintenance of pipeline leakage in remote areas. If any leakage is detected in pipes which are located in remote locations then it will be difficult to trace the location of pipeline leakage, sometimes the concerned authorities may take several days to race the location and repair the pipe where the leakage is happening. So this delay causes large quantity of water loss. So this conventional approach need to be automated by adopting sensors based technology for identifying the leakage and tracking the location were exactly leakage is happing so that within short period of time leakage can be repaired or prevented so that there is no much loss of water.

The main objective of proposed system is to survey the existing methods and solutions and propose an enhanced method for detection and monitoring of pipelines in remote locations using different sensors.

There are mainly four objectives which describes the complete working of the system.

1. To detect water leakage using flow rate sensors.
2. Monitoring pipeline system using Vibration sensors.
3. Sending notification message to concerned water management system authority when leak is detected.
4. To develop an efficient pipeline system to prevent leakage of water.

2   Related Work

In today's world drinking water is essential for every leaving organism, so we need proper water distribution mechanisms so that the drinking water reaches to all the people with required quantity. The need of sustainable water supply system is ever increasing. In the conventional approach the leakage of water is detected manually. As discussed in Table 1 various different methods are used to detect pipeline leakage [1].

Monitoring leakage of underground pipelines now has become one of the important and challenging tasks. Acoustic emission effects and wireless connection for communication makes the system easy to deploy. Acoustic emission based deployed systems are inexpensive and gives accurate results [2]. The main objective is to detect the underground pipeline leakage which will be monitored in the server. The system is designed and developed using different types of sensors such as temperature, acoustic, flow rate and pressure sensors, the readings are continuously recorded from these sensors and are processed in the server to detect and locate the exact leakage point in the pipelines [8].

2.1 Classification of Leak Detection Techniques

According to the survey the leak detection techniques can be classified into two main categories i.e. hardware and software based methods. The Fig. 1 shows the classification of leak detection techniques.
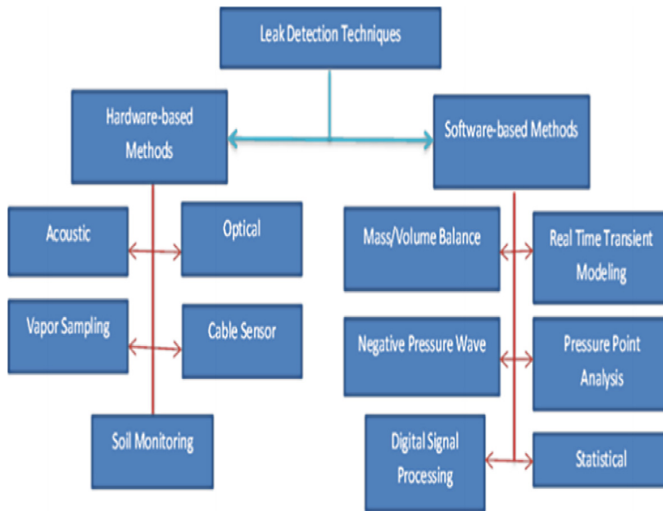
Fig. 1. Classification of leak detection techniques

In case of hardware based leak detection the leak is detected externally outside the pipeline by visualization or through some devices [3]. These methods have potentially good sensitivity to detect leak and are very accurate in tracing the leak location.

In software based leak detection different software techniques are used, programs are written to monitor and different algorithms are used to continuously monitor the water pipeline considering the following metrics such as flow rate of the water within the inlet and outlet of pipes, temperature, vibration and pressure of the water flowing through the pipe. The above discussed methods need power and continuous monitoring service to maintain the pipeline system in good condition [5]. Along with these

techniques there are few more techniques which are efficient to maintain and monitor the pipelines in the rural and remote areas where the water board authority workers find it difficult to continuously walk around and monitor so to overcome these issues and challenges the water board management authority can use latest and efficient techniques such as ground penetrating radar (GPR), pressure measurements and acoustic measurements, these techniques helps to identify the leakage and can be used to maintain the good condition of the pipeline system.

The leak detection systems discussed in Table 1 can be broadly classified into non-invasive and invasive. Non-invasive leak detection techniques are used to detect external leak of pipes using sensors or visual inspection [8]. In case of Invasive leak detection we continuously monitor the pressure, temperature and flow rate of the water. Invasive leak detection techniques require drilling of holes or pipe cutting to mount the device inside the pipe, so it's not that flexible and often disrupt the pipeline mechanism.

2.2   Cable Sensing Method

This method involves installation of cable through the entire pipeline. Here the sensor consists of two electrical circuits, one circuit is used to monitor the cable life, and the second circuit is used to sense the leakage in pipes. Whenever there is a leakage in the pipe the spilled water will get in touch with the embedded circuit board and this intern may lead into short circuits and breakage of complete system. So the cable sensing method basically uses wired networks which have major drawbacks such as cost, maintenance, installation. And if there is damage in the cable then it will directly affect the network performance.

Maintenance of these wires is difficult and we need to pay for labour and repair cost is high [7]. A wireless sensor network seems to be a better solution compared to traditional wired networks. In fact most of the researchers have given more focus on wireless sensor network considering feasibility, high performance, low cost and easy deployment for monitoring the pipeline detection system.

2.3   Leak Edge Detection Method

In leak edge technique we usually measure the pressure using the acoustic based methods. The acoustic based pressure measurement techniques are very costly and consume lot of time for the deployment, repair and maintenance of the pipeline system [9]. In leak edge detection technique the major focus is on steady state, but we need to migrate this from steady state analysis to transient state analysis.

According to the survey there are so many limitations of existing system such as

- It needs more human resource.
- There can be human error.
- It is costly.
- It is time consuming. During that time there can be more water leakage.
- Inspection and replacement of pipelines as quickly as possible.

3   System Design

The overall block diagram of the enhanced pipeline monitoring system is as shown in Fig. 2, the system mainly consists of sensors, server, GSM, GPS and microcontroller.
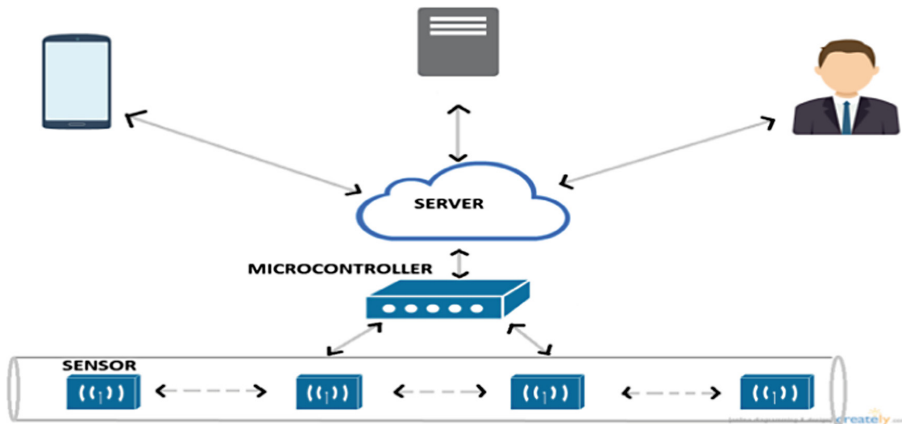
Fig. 2. System architecture

The proposed system model mainly consists of two modules, the transmitter module and the receiver module. In the transmitter model the flow rate sensors and vibration sensors are deployed to capture the flow rate of water within the pipe, the sensor generates the pulses. A microcontroller used to process the pulses and provides the desire results. A wireless module is mounted at the transmitter end to transmit the sensed pulse vales to the server; these pulse values will be continuously monitored.

3.1   Transmitter Module

The Fig. 3 shows the representation of transmitter module. Flow rate sensors are used to get the flow rate of water within the pipe. These sensors are installed at the inlet and outlet of the pipes the values of flow rate sensors are captured and compared with the threshold value. A threshold value is fixed considering the flow of the water within the pipe. If the difference of the flow rate of water is greater than the threshold value, then a notification message about the leak detection will be sent to the concerned authority of water board management.

The main goal here is to reduce leakage and save water, so just informing concerned authority about leakage can work to certain extent, but informing them about the break in the pipe well in hand, can help them to take corrective measures before any pipeline breakage. To measure the durability of the pipe vibration sensors are used, if the pipe is too old and about to break then also a notification message will be sent to the concerned authority so that pipe leakage can be avoided.
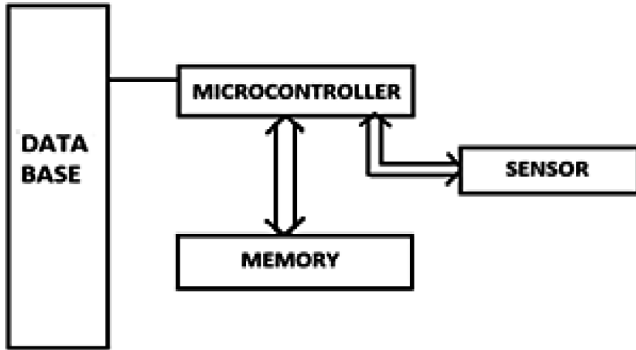
Fig. 3. Transmitter module

3.2   Receiver Module

The Fig. 4 shows the receiver module setup, the receiver model consists of web server which receives the data from the transmitter module and monitors that data continuously. If there is leakage detection or the pipe is too old with cracks, then the receiver module will send the notification message to the concerned authority.
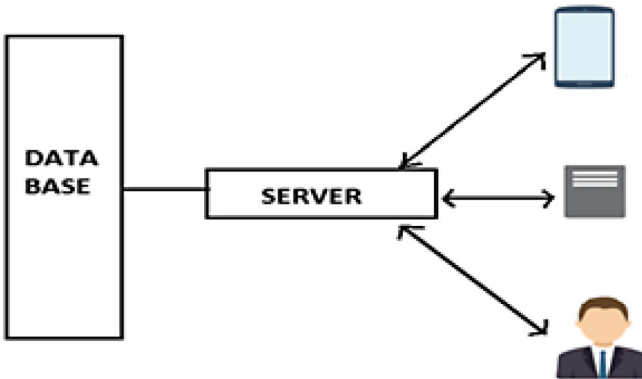
Fig. 4. Receiver model

3.3   Flow Rate Sensor

The flow rate sensors are used to measure the flow rate of water within the pipe at both inlet and outlet of the pipeline, the Fig. 6 illustrates the setup of Water pipeline monitoring system using flow rate sensor, the prototype of the proposed system mainly consists of a pipeline of 3 feet length with the two valves and the sensors which are mounted on the two ends of the pipe i.e. one at the input end from where the water enters the pipe and another one is the output end from where the water comes out of the pipe, if there is a leakage in-between the pipe then the flow rate sensor values will be

processed in the microcontroller, if the flow rate sensor reading are different at both inlet and outlet ends, and if this difference is more than the threshold value then it indicates that there is leakage in the pipeline. Two sensors are mounted on pipe one at the inlet and other at outlet end, both the sensors value are compared to infer the status of the pipe. The flow rate sensors are usually made up of plastic valve body Hall Effect sensor with a water rotor. When the water passes through the rotor, the speed of the rotor changes along with the flow rates of the water. The hall-effect sensor provides the corresponding output pulse signals (Fig. 5).
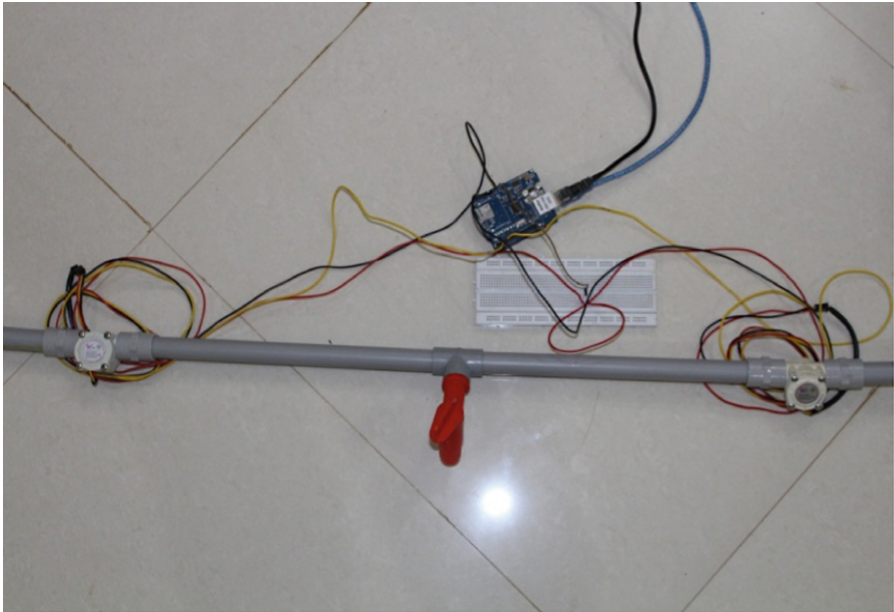
Fig. 5. Pipeline setup for leak detection using flow rate sensor

The parameters considered for flow rate sensors are

- Working range: 1–30 L/Min
- Water pressure: ≤ 1.75 MPa.

3.4   Vibration Sensor

A 6DOF sensor – MPU6050 using accelerometer sensor is used to identify the leak of pipe, The sensor values are continuously transferred to the server system at an interval of 1 s over an Wi-Fi network, the sensor values is analysed at the server system to identify the leak-aged pipe or normal pipe.
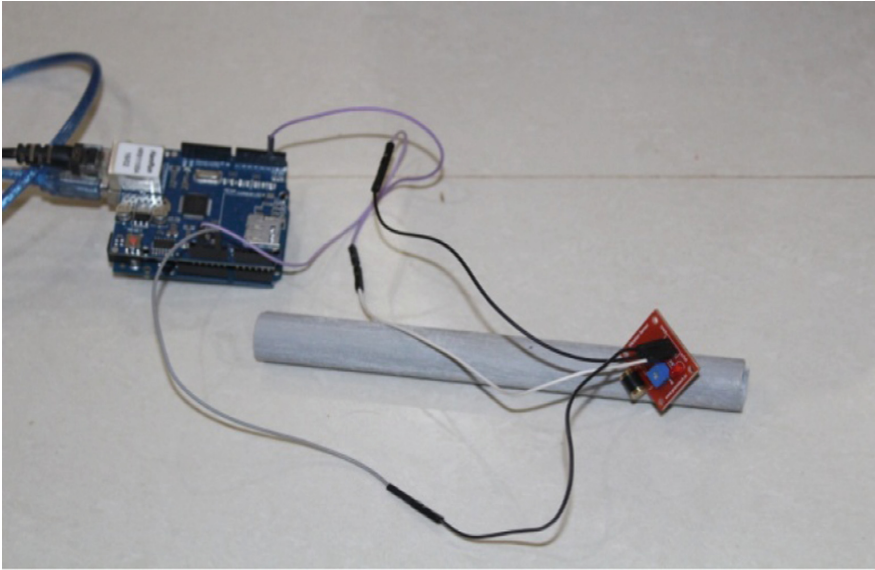
Fig. 6. Pipeline setup for leak detection using MPU6050 sensor

The Table 2 shows the MPU6050 sensor configuration and sensitivity.

Table 2. MPU6050 sensor configuration and sensitivity

Accelerometer full range scale	Accelerometer sensitivity (LSB/g)
±2 g	2048
±4 g	4096
±8 g	8192
±16 g	16384

The Fig. 6 shows the setup of water pipeline monitoring system for leak detection. The experiment was carried out considering two sensors, flow rate and vibration sensors, these sensors are connected to arduino board and are continuously monitored. The arduino board is programmed and set a sensitivity of MPU6050 sensor. The sensitivity of MPU6050 sensor is set to ±15 g and ±20000/s.

3.5 Arduino Controller Board

Arduino UNO micro-controller is used for sensing data and to control physical devices. The arduino board is based on ATmega328 (datasheet). It basically consists of 14 digital input/output pins among which 6 pins are used for PWM output, 16 MHz crystal oscillator, a USB connection, 6 analogue inputs, a power jack and an ICSP header (Fig. 7).
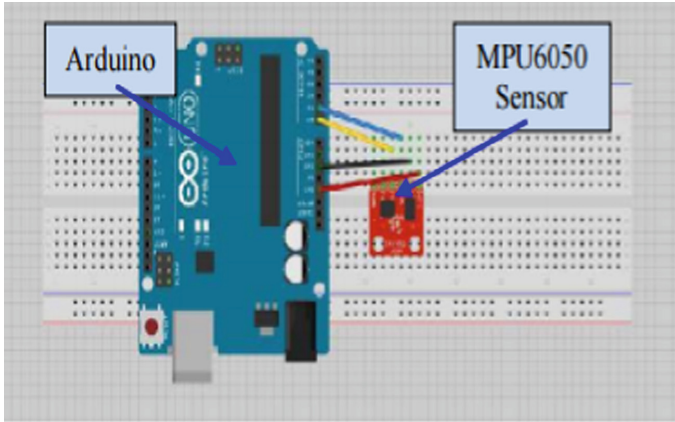
Fig. 7. Arduino Board with MPU6050 sensor

4 Implementation

4.1 Leak Detection Algorithm

Algorithm 1 Leak Detection

1: Read flow rate sensor values.
2: Store data into database.
3: Calculate the difference in the flow rate sensors.
4: **if** the difference is greater than 1 l/min **then**
5:     Enter into For loop.
6:     **for** i less than 15 **do**
7:         read sensor value
8:     **end for**
9:     if Still the difference is greater than 1 l/min **then**
10:         Inform authority about the leakage
11:     **end if**
12: **end if**

4.2 Flow Chart for Leak Detection

The Fig. 8 describe the step by step operations carried out to detect the leakage, the first step is to read the flow rate sensor values and transfer these values to the server system database using a wireless network. In second step calculate the difference in the flow rate values of two sensors if the difference is less than the threshold values then send a notification message to the concerned authority. The threshold value is fixed considering 15–20 readings after the configuration.
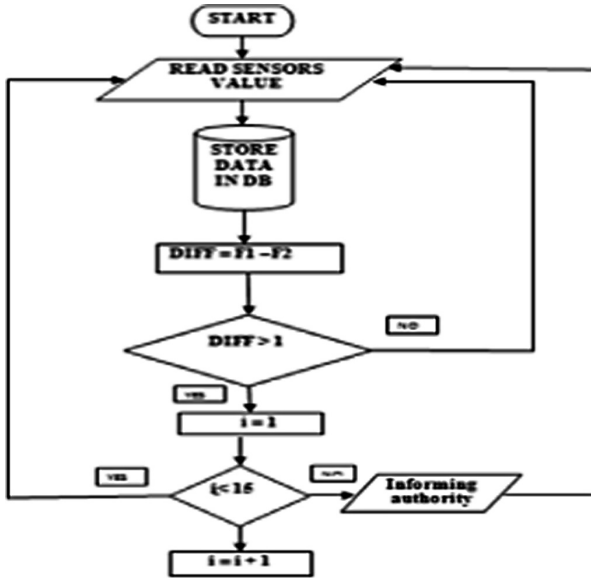
Fig. 8. Flow chart for leak detection

4.3   Flow Chart for Damage Alert

The Fig. 9 describes the step by step operations carried out to detect the leakage using vibration sensor and indicating the damage alert. The first step is to read the values
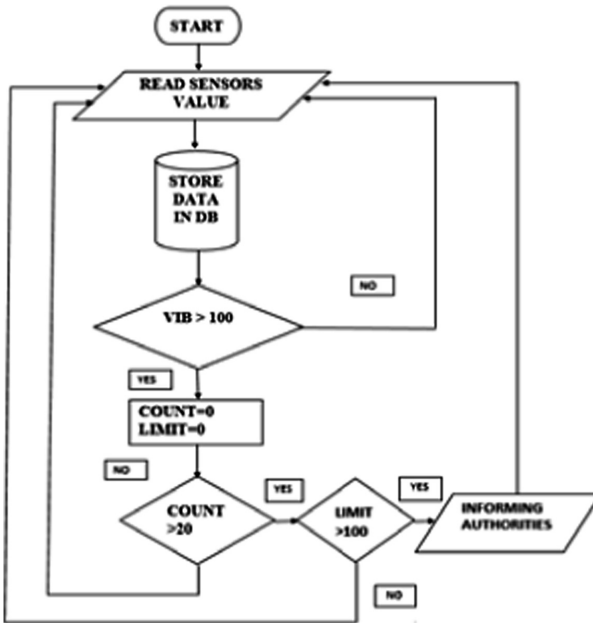
Fig. 9. Flow chart for damage alert

from MPU6050 sensor and transfer these values to the server system database using a wireless network to confirm the damage in pipe, following 20 readings are taken, and a counter is incremented whenever it crosses the threshold. If this counter is greater than 10, then there is a high probability that there is damage and a notification message will be sent to the authority.

4.4 Algorithm for Damage Alert

Algorithm 2 Damage Alert

1: Read vibration sensor values.
2: Store data into database.
3: **if** vibration is greater than 100 **then**
4:      Enter into For loop.
5:      **for** count less than 20 **do**
6:          read sensor value
7:      **end for**
8:      **if** count is equal to 20 and limit is greater than 10 **then**
9:          Warn authority about the break
10:      **end if**
11: **end if**

---

5 Results and Discussion

5.1 Capturing Flow Rate Sensor's Values and Processing

The Fig. 10 shows the result analysis of the two flow rate sensors which are mounted at inlet and outlet of the pipe continuously the readings are captured from both the sensors and difference is being calculated which is shown in the below graph.

5.2 Threshold Value

The pipe which was used for the experiment is 1 inch in diameter and 1 m in length. After taking 25–30 readings for observation the average value of the readings was calculated to fix the threshold value, i.e. 1 L/min was set as threshold for leakage

The Fig. 10 shows the difference value computed considering the readings from flow rate sensors. If water flows uniformly then there will be a linear graph which indicates that there is no leakage in the pipeline. And suppose if there is any variation in the flow rate of water because of leakage in the pipe then there won't be a linear growth in the graph.
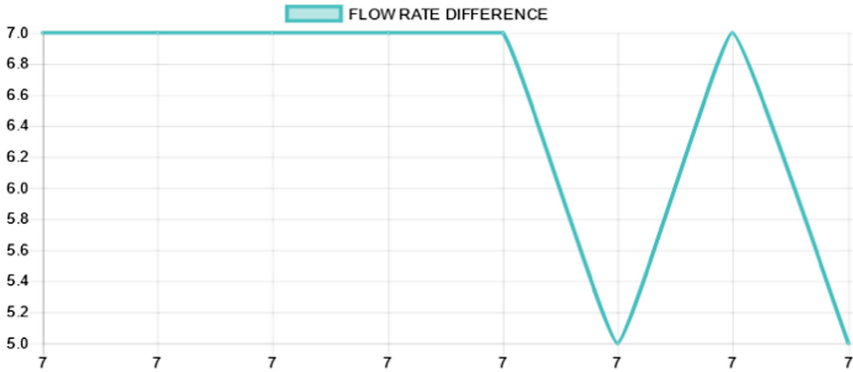
Fig. 10. Graph for monitoring flow rate sensor readings – (flow rate1 vs. flow rate2)

5.3   Vibration Sensor's Readings in the Database

The Table 3 shows the vibration sensor data value stored in the database. If there is any disturbance in the pipe, the vibration increases. If the disturbance remains for about 2 min, then there are chances of pipe being broken.

Table 3. Vibration sensor values in the database

Time stamp	Vibration sensor value
2016-12-26 14:49:16	0
2016-12-26 14:49:20	56
2016-12-26 14:49:22	38
2016-12-26 14:49:24	110
2016-12-26 14:49:26	40
2016-12-26 14:49:28	52
2016-12-26 14:49:30	0
2016-12-26 14:49:32	50
2016-12-26 14:49:34	25
2016-12-26 14:49:36	33
2016-12-26 14:49:38	32
2016-12-26 14:49:40	0
2016-12-26 14:49:42	103

5.4   Vibration Sensor's Values and Processing

To analyse the leakage in pipe, vibration sensors can also be used. The Fig. 11 shows the vibration sensor readings vs. the timestamp graph, again if the readings are more than the threshold value then a leak is detected.
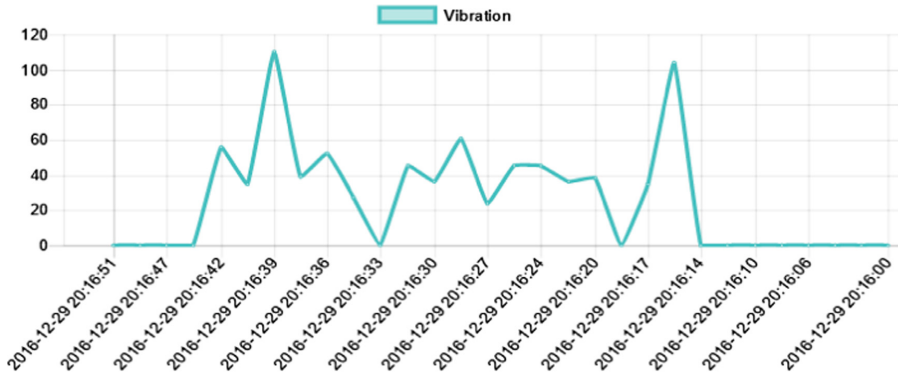
Fig. 11. Graph for monitoring vibration sensor readings

5.5   Sending Notification Message

The two sensors will be continuously sending there readings if the sensor reading is more than the threshold value then a notification message will be sent to the concerned water board authority so that quick action is initiated to repair the pipe. The Fig. 12 shows the message that is received whenever there is leakage or break in the pipes, to transmit these messages we are using wireless medium which will be sent to the water board authority through the server system.
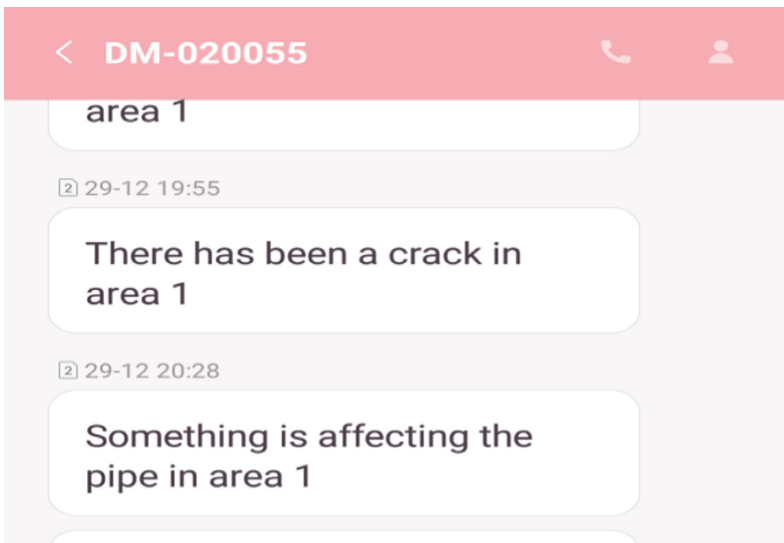
Fig. 12. Notification message

6 Conclusion

In today's life style water is essential and the crucial element for human beings. In majority of the places water is wasted in large volumes due to poor and low quality pipelines which result in breakage of pipes and no proper and quick methods are adopted to repair the pipes, which basically lead into water wastage. So to minimize public health risk and economic loss we need to manage and maintain the pipeline systems, water board authorities regularly audit their Pipeline distribution systems and perform pipeline detection survey. In the remote locations its very challenging and difficult to monitor and repair the pipes due to unavailability of the resources and we won't get to know the leakages so to address these challenges, we designed and developed the smart pipeline monitoring system in which flow rate and vibration sensors are embedded on the pipe to detect leakage in pipe if any. As soon as the leakage is detected a notification message is sent to the concerned authorities so that a quick action is taken to repair the pipe and stop the water leakage. Experimental observations show that flow rate and vibration sensors are best suited to detect the leakage in water pipeline monitoring system.

References

1. Jawhar, I., Mohamed, N., Mohamed, M.M., Aziz, J.: A routing protocol and addressing scheme for oil, gas, and water pipeline monitoring using wireless sensor networks. In: 5th IFIP International Conference IEEE Wireless and Optical Communications Networks, WOCN 2008 (2008)
2. Hieu, B., Choi, S., Kim, Y.U., Park, Y., Jeong, T.: Wireless transmission of acoustic emission signals for real-time monitoring of leakage in underground pipes. KSCE J. Civ. Eng. 15, 805–812 (2011)
3. Liu, Z., Kleiner, Y.: State-of-the-art review of technologies for pipe structural health monitoring. IEEE Sens. J. 12(6), 1987–1992 (2012)
4. Gao, Y., Brennan, M., Joseph, P., Muggleton, J., Hunaidi, O.: On the selection of acoustic/vibration sensors for leak detection in plastic water pipes. J. Sound Vib. 283, 927–941 (2005). Sensors 14(3), 5595–5610 (2014). https://doi.org/10.3390/s140305595
5. Geiger, G., Werner, T.: Leak detection and location - a survey. In: Proceedings of the PSIG Annual Meeting, Bern, Switzerland, pp. 1–11 (2003)
6. Sadeghioon, A.M., Metje, N., Chapman, D.N., Carl, J.: Smart pipes: smart wireless sensor networks for leak detection in water pipelines. Sens. Actuator Netw. 3, 64–78 (2014). https://doi.org/10.3390/Jsan3010064
7. Ria, S., Manjit, K., Hemant, K.: Design and development of automatic waterflow meter. Int. J. Comput. Sci. Eng. Appl. (IJCSEA) 3, 49–59 (2013)
8. JayaLakshmi, M., Gomathi, V.: An enhanced underground pipeline water leakage monitoring and detection system using wireless sensor network. In: 2015 International Conference on Soft-Computing and Networks Security (ICSNS), pp. 1–6 (2015)
9. Misiunas, D., Vitkovsky, J., Olsson, G., Simpson, A., Lambert, M.: Pipeline break detection using pressure transient monitoring. J. Water Resour. Plan. Manag. 131(4) (2005). https://doi.org/10.1061/(ASCE)0733-9496(2005)131:4(316)

10. Obeid, A.M., Karray, F., Jmal, M.W., Abid, M., Qasim, S.M., BenSaleh, M.S.: Towards realisation of wireless sensor network-based water pipeline monitoring systems: a comprehensive review of techniques and platforms. IET Sci. Meas. Technol. **10**(5), 420–426 (2016)
11. Rani, M.U., Kamalesh, S., Preethi, S., Shri, C.K.C., Sungaya, C.: Web based service to monitor water flow level in various applications using sensors. Int. J. Biol. Ecol. Environ. Sci. (IJBEES) **2**, 119–122 (2013)
12. Demma, A., et al.: The reflection of the fundamental torsional mode from cracks and notches in pipes. J. Acoust. Soc. Am. **114**(2), 611–625 (2003)

Smart Helmet with Cloud GPS GSM Technology for Accident and Alcohol Detection

Praveen M. Dhulavvagol[(✉)] , Ranjitha Shet , Prateeksha Nashipudi ,
Anand S. Meti , and Renuka Ganiger

KLE Technological University, Hubballi, India
{praveen.md,anandsmeti}@bvb.edu, ranvishet@gmail.com,
prateekshanashipudi96@gmail.com, renukaganiger@gmail.com

Abstract. Today in India and most other countries, accident is considered as an unexpected and unintended event. According to the survey road accidents lead to hospitalization, injuries and disabilities, considering these facts the safety of the riders has become a most crucial and important issue and concern in majority of the countries. The usage of two wheeler's vehicles is increasing day by day for transport convenience we need to have proper safety and efficient measures to handle accidents. In the proposed paper we discuss smart helmet system using cloud and GPS technology for accident detection and location tracking, GSM is used to send a notification message of accident location to the concerned people to avoid major casualty/life of the person can be saved. MQ3 sensor is also embedded in the helmet to check whether the motorcyclist has consumed alcohol or not. The sensor data captured is stored in the cloud so that the accident information can be fetched anywhere and anytime at a faster rate. Considering the above facts, the prototype of smart helmet is designed and implemented to enhance the safety of the motorcyclist. This new idea will reduce the risk of motorcyclist life. The experimental results confined that the developed prototype model gives 84% accurate result of the accident detection and concerned persons will receive notification message.

Keywords: Mem's sensor · MQ3 sensor · GPS technology · Latitude
Longitude

1 Introduction

The road accidents are one of the major problems all over the world. Around 80,000 people are killed due to rash driving, drunken driving and less usage of helmets. The influence of accidents involving motorcyclist death, fatality and disabilities may be due to over speed, not wearing helmets and not following the road safety measures [1]. During such situations we need to take appropriate measures quickly so that the motorcyclist life can be saved and proper medication is provided quickly to reduce the death rate of motorcyclist. Now a day many new techniques have been adapted in a motorbike such as Anti Breaking System (ABS), Anti Collision System (ACS) to avoid accidents, though latest techniques are adapted in motorcycles still the accidents are happening and people are losing their lives. This might be due to various reasons such as the

motorcyclist is not given proper medication as soon as the accident happens, delay in getting him to the hospital through ambulance due to heavy traffic, accident details are not conveyed to the concerned persons and this delay increases the death rate of the people [5]. Nowadays in India, it is being made mandatory for the motorcyclist to wear helmet while driving vehicles [2].

In this paper the main aim is to enhance the safety of the rider by implementing the following objectives:

1. To detect accident using MEMS sensor.
2. Alcohol detection using MQ3 sensor.
3. To capture the data from sensor and store it in a cloud.
4. To trace the accident location using GPS.
5. Sending the accident location to the concerned people using GSM Modem.

2 Related Work

According to survey details around 80,000 people are killed due to rash driving, drunken driving and less usage of helmets [3]. In India, as we know that helmet is compulsory for a rider to wear but most of the people wear low quality and cheap helmets to avoid paying for traffic police officers and not for their life and body protection [6]. And also the mortality rates can be reduced if proper medical aid reaches the motorcyclist at the earliest. Considering all these reasons the conventional helmet need to be redesigned and developed with latest technologies to overcome the above discussed problems [4]. The survey has been carried out to find out the best strategy available by referring journals, existing system and analysing the results.

According to the reference paper [1] the author discusses a prototype of a smart helmet which provides a means for detecting and reporting accidents. Here the helmet detects the accident and intimates the accident information to the emergency contacts list. Here the author has developed smart helmet using 3-axis accelerometer to monitor the head movement and orientation of motorcyclist, helmet position. Considering these values we can estimate the chances of accidents. Once the accident happens the exact location of accident spot is traced using GPS system and this location is shared to emergency contact peoples. The system also prevents false positives from being triggered. False positives are nothing but, if the helmet drops from certain heights from his hand or it is being thrown by mistake then also the system will send notification message which results in faulty functioning of the system. So to avoid such false positives, accelerometer values need to be monitored appropriately.

In Paper [2] the author discusses a Helmet based on IoT (Internet of Things), powered with Arduino Controller and Wi-Fi unit. The benefits include:

 (i) Tracking the vehicle location using the Wi-Fi network.
(ii) Communication with pre-programmed mobile users during an emergency. The receiver setup collects and stores data from the wireless transmitters, and uploads the Real-time sensor data to the Cloud Services.

In Paper [3] the author discusses a Helmet which is safety equipment for driver. In this system the accident is detected using the vibration or MEMS sensor which uses GPS to trace the accident location and latter transmits a SMS message to the ambulance service and hospital. This system reduces the accident death rates by providing proper medication within the time frame.

In Paper [4] the author discusses on the negligence of the motorcyclist for not wearing the helmet. Road accidents are increasing day by day due to which many people face disabilities and there is no proper mechanism or technology which detects the accidents and location of accident spot to save the motorcyclist.

Considering all the above facts and the challenges, we need to design and develop a smart system which can save the life of motorcyclist. The proposed system is an intelligent helmet. Additional feature of accident detection module is if the accident is detected, the detected location will be notified immediately to the traffic control office. If the accident is minor then the motorcyclist can avoid sending notification message by pressing abort switch.

3 System Design

The main functionalities of the proposed system are:

1. The system should be able to detect the cause for the accident. The end user should be able to decide the cause for the accident and this can be done using MEMS sensor and Alcohol sensor. The MEMS sensor detects the accidents based on the pressure applied on the helmet. The alcohol sensor tells whether the rider was in normal state or drunken state.
2. The system should be able to detect the accident spot based on latitude on longitude values.
3. The system should be able to send the text message to the emergency canter and family members.

The block diagram of smart helmet system is as shown in Fig. 1. It mainly consists of five major components i.e. The MEMS sensors and MQ3 sensor are fixed in the helmet which are in turn connected to the Arduino board. The fixed threshold values of MEMS sensor are stored in Arduino-Mega. Whenever the MEMS sensor crosses this threshold value, accident is detected. MQ3 sensor is embedded inside the helmet which is used to trace weather the motorcyclist has consumed alcohol or not by checking the breath of the motorcyclist [7].

Arduino board will be continuously sensing the sensor values and it stores these sensor values in cloud for faster processing and computation. To track the accident location GPS tracker is used to trace the exact location of the accident. Once the location is traced the system will send the notification message to the concerned persons using a GSM modem. With the help of GSM modem a notification message consisting of location link is sent to the concerned person so that major casualty/life of a person can be saved.

The Proposed system is divided into mainly two modules the transmitter module and the receiver module.
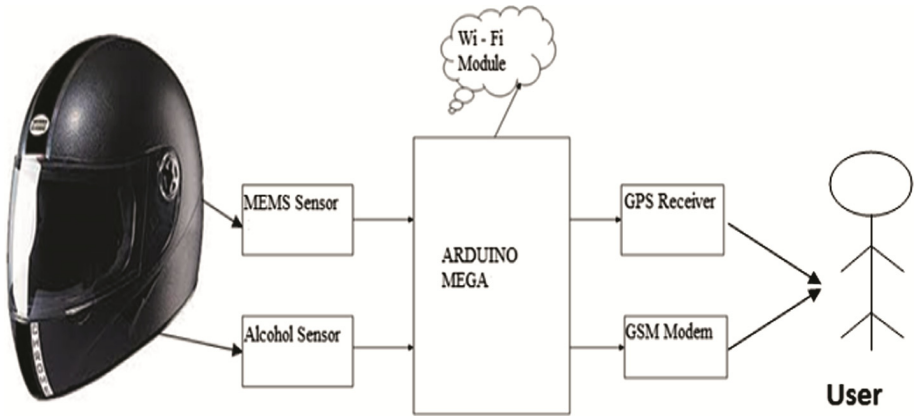
Fig. 1. Smart helmet block diagram

3.1 Transmitter Module

The above Fig. 2 shows the working of transmitter module; here the MEMS sensors and alcohol sensors are connected to the arduino-Mega board. Both the sensors will be continuously reading the data and the read data will be stored in the cloud for faster access and maintenance.
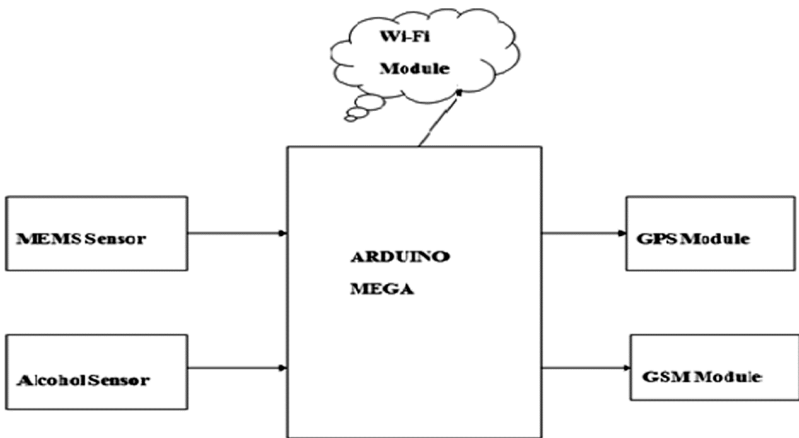
Fig. 2. Transmitter module

3.2 Receiver Module

The Fig. 3 shows the flow diagram of receiver module which is mainly used to receive the sensor information processed by the arduno board. A threshold value is fixed in MEMS sensor. If the sensor value is more than the threshold value then accident will

be detected [4]. When the rider wears the helmet and if he is in drunken state then the alcohol sensor will detect the presence of alcohol.
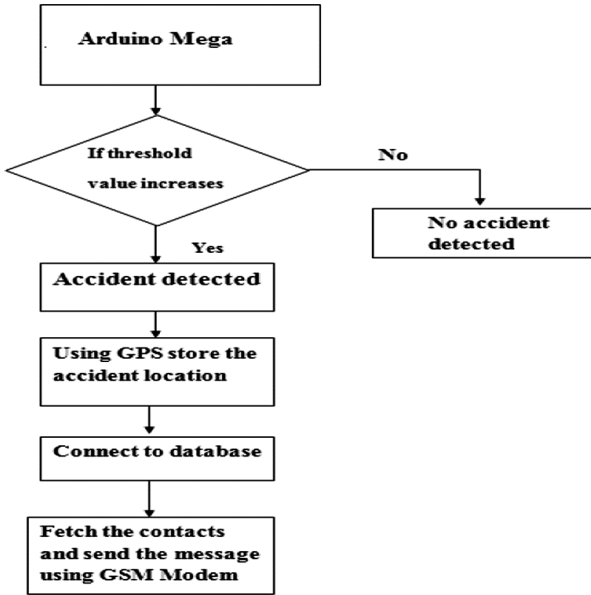
Fig. 3. Receiver module

4 Implementation

4.1 MEMS Sensor

The MEMS sensor is shown in Fig. 4, which is used to detect the accident based on threshold value fixed in the Arduino and this sensor is mounted on top of the helmet.

Fig. 4. MEMS sensor

4.2   MQ3 Sensor

The MQ3 sensor is shown in Fig. 5, which is basically used to detect weather the motor-cyclist has consumed alcohol or not, MQ3 is an alcohol sensor which detects the content of alcohol. The sensor needs 2 to 3.3 V power supply to function and the sensor is around 16.8 mm diameter with 9.3 mm height without pins. This sensor is helpful in identifying the motorcyclist who drink and drive, and also helpful in tracing drink and drive cases for the traffic police.

Fig. 5.   MQ3 sensor

4.3   Arduino-Mega

The Fig. 6 shows the block diagram of arduino board which consists of 14 input and output pins, in this 14 pins 6 pins are used for PWM outputs, a ceramic resonator of 16 MHz which are used for processing the sensed data. Both the MEMS and the MQ3 sensor are connected to the Arduino - Mega which has some fixed threshold values of
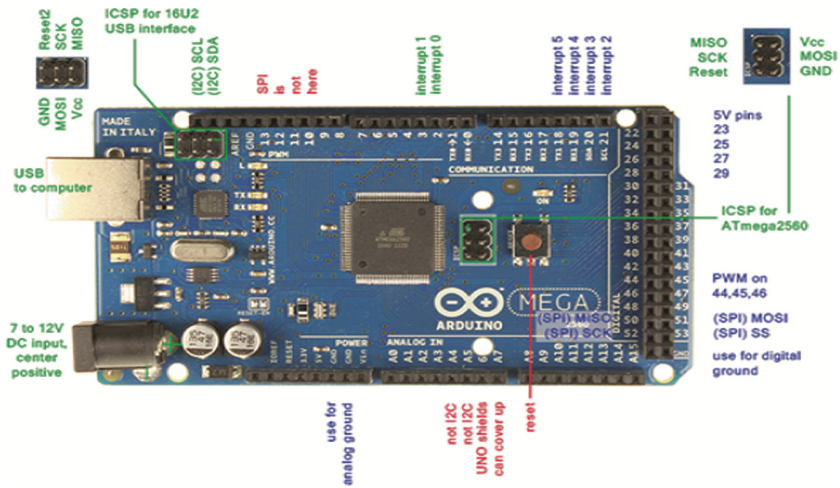
Fig. 6.   Arduino board

MEMS sensor. When the MEMS value crosses the threshold value, accident has been detected.

4.4 Global Positioning System (GPS)

The GPS system is shown in the below Fig. 7, the main functioning of GPS is to trace the accident location; GPS uses the radio signals to send and receive the location information. GPS is mainly used to fetch the current vehicle location where the accident has taken place. In GPS system there are 3 pins receiver pin is connected to the transmitter pin of GSM, and the transmitter pin is connected to receiver pin of GSM and the last pin is connected to Vcc. It gives the latitude and longitude values of a particular location and these values are sent to the GSM module which sends notification to the emergency contacts. GPS works in all weather condition.
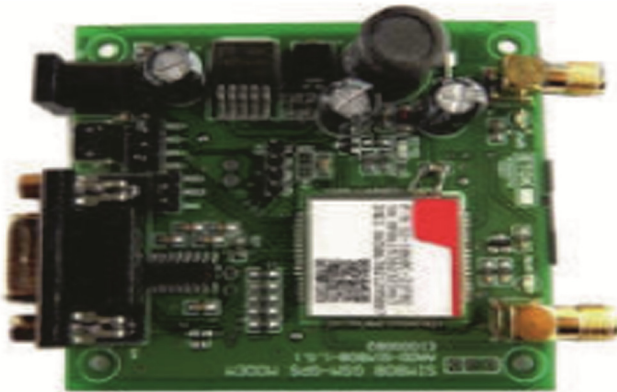
Fig. 7. GPS system

4.5 GSM Modem

The Fig. 8 shows the GSM modem which is used for sending the notification message to the concerned persons, it uses digital cellular technology for transmitting voice and data services.

All the data received from the sensors are stored and maintained in the cloud. The accident location information is recorded considering latitude and longitude values.

Fig. 8. GSM modem

5   Results and Discussion

The prototype model of the circuit is shown in Fig. 9. Here we have used MEMS sensor to detect the accident based on threshold value and alcohol sensor i.e. MQ3 sensor detects weather the motorcyclist has consumed alcohol or not. These two sensors are connected to Arduino- Mega board. Accident location is detected using GPS receiver and once the accident location is detected the GSM modem will send
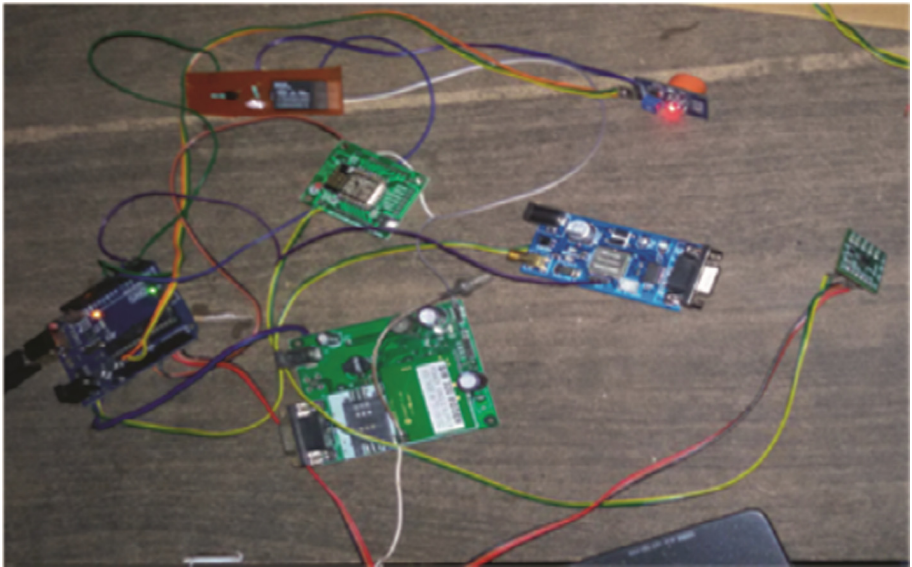
Fig. 9. Setup of prototype model

this accident location to the concerned people in human readable form. All the captured data from the sensor are stored in cloud using Wi-Fi module (Fig. 10).
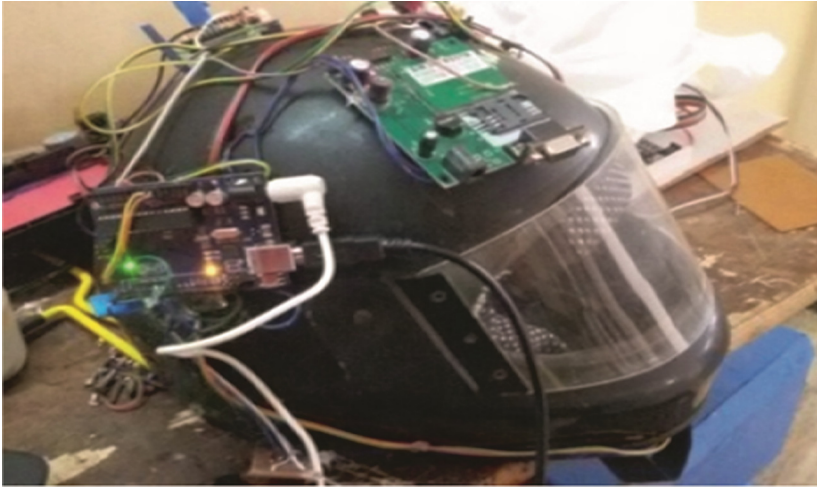
Fig. 10. Smart helmet prototype model

5.1   Accident Detection Using MEMS

The Fig. 11 shows the readings of MEMS sensor
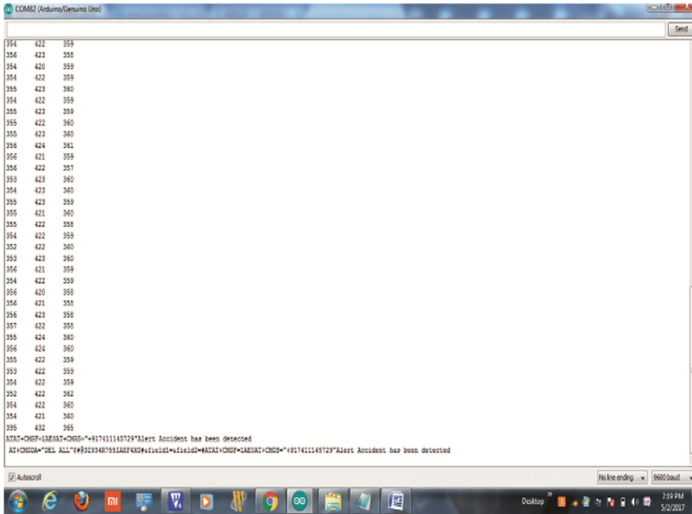
Fig. 11. MEMS sensor readings

If the motorcyclist meets with an accident, then the location of the accident is traced using MEMS sensor, MEMS sensor works on the principle of angular momentum for measuring orientations. If the frequency value is greater than threshold value (greater than 380) a notification message will be sent. Once the accident is detected the accident spot will be traced using GPS tracker and notification message will be sent to concerned persons.

5.2   Sending a Notification Message

See Fig. 12.
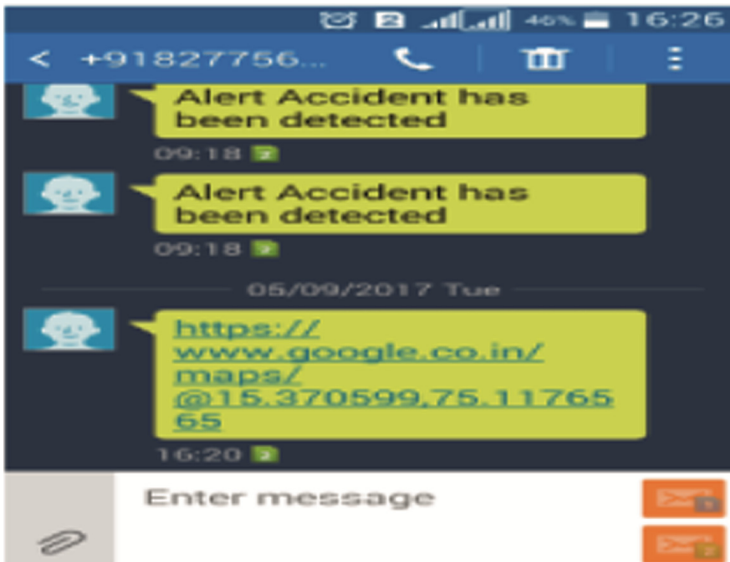
Fig. 12.   Notification message

5.3   Tracking Accident Location

Once the accident is detected the accident location link will be traced using GPS tracker and notification message will be sent to concerned persons (Fig. 13).

Fig. 13. Accident location

The MQ3 sensor is used to sense the presence of alcohol in the motorcyclist breath. If sensitivity of Alcohol sensor is more than 0.04 mg/l of alcohol then "Alcohol Detected" message will be displayed. The detected location will be in the form of latitudes and longitude values. Further this location will be sent to the concerned persons and the nearby emergency care centre.

5.4   Test Cases

5.4.1   Test Case 1

Input: Identify Accident location Using GPS and send the accident location to the concerned person using GSM Modem.
Expected Output: Accident location sent to the concerned person.
Output: Accident location is human readable format sent to the concerned person.

5.4.2   Test Case 2

Input: GSM sends Accident Location to concerned people and Emergency Center.
Expected Output: Message should be delivered in human readable format consisting of latitude and longitude values.
Output: Message delivered to specified numbers and Emergency Center.

5.4.3   Test Case 3

Input: Accident location in the form of latitude and longitude values need to be stored in cloud.
Expected Output: Values in human readable form.

Output: User will be able to access the values from the cloud in human readable format.

6 Conclusion

Motorcyclist accident detection and location tracking system is designed and implemented using technologies like GPS and GSM. MEMS sensor is used for detection of accident; MQ3 sensor detects whether the motorist has consumed alcohol or not which helps in keeping a check on drunk and driving cases, GPS modem is used to find the location of accident spot, the GPS values are recorded in terms of latitude and longitude, and GSM modem is used for transmitting the message to the concerned persons. All the information collected from these are stored in cloud for future references. The above design satisfies all the objectives. This proposed idea can be further enhanced by sending the present location and the vehicle number of the motorist who has consumed alcohol and violated traffic rules.

References

1. Chandran, S., Chandrasekar, S., Elizabeth, N.E.: Konnect: an Internet of Things (IoT) based smart helmet for accident detection and notification. In: 2016 IEEE Annual India Conference (INDICON), pp. 1–4. IEEE, December 2016
2. Srikrishnan, A., Kumar, K.S., Ravi, S.: Cloud incorporated smart helmet integrated with two-wheeler communication setup. IJCTA **9**(4), 2025–2035 (2016)
3. Lakshmi Devi, P., Bindushree, R., Deekshitha, N.M., Jeevan, M., Likhith, M.P.: Helmet using GSM and GPS technology for accident detection and reporting system. Int. J. Recent Innov. Trends Comput. Commun. **4**(5), 18–21 (2016). ISSN: 2321-8169
4. Wei, W., Hanbo, F.: Traffic accident automatic detection and remote alarm device. In: 2011 International Conference on Electric Information and Control Engineering (ICEICE), pp. 910–913. IEEE, April 2011
5. Jung, S., Shin, H., Yoo, J., Chung, W.: Highly sensitive driver condition monitoring system using nonintrusive active electrodes. In: Proceedings of the IEEE ICCE, Las Vegas, NV, USA, pp. 305–306 (2012)
6. Abduljalil, F.M.: A framework for vehicular accident management using wireless networks. In: 2012 IEEE 13th International Conference on Information Reuse and Integration (IRI), pp. 727–729. IEEE, August 2012
7. Donnelly, B.R., Schabel, D.G.: Method and apparatus for automatic vehicle event detection, characterization and reporting. U.S. Patent No. 6,076,028. U.S. Patent and Trademark Office, Washington, DC (2000)

Text-Dependent Speaker Recognition System Using Symbolic Modelling of Voiceprint

Shanmukhappa A. Angadi[1] and Sanjeevakumar M. Hatture[2(✉)]

[1] Visvesvaraya Technological University,
Belagavi 590018, Karnataka State, India
`vinay_angadi@yahoo.com`
[2] Basaveshwar Engineering College, Bagalkot 587103, Karnataka State, India
`smhatture@yahoo.com`

Abstract. Speaker recognition system automatically recognize/identify a speaker by their combined behavioral and physiological characteristics. A symbolic inference system for text-dependent speaker recognition system by exploring the physiological characteristics embedded in the user utterance is presented in this paper. The inter-lexical pause position, complementary spectral features such as spectral centroid, spectral entropy and spectral flatness, loudness, pitch and formants features are extracted from the voiceprint and symbolic data object is constructed. These features are explored in this work as inter-lexical pause position provides the articulation capability of user vocal tract. The functional properties of the human ear is modelled with spectral characteristics and loudness feature provides the strength of ear's perception. The relation between physical and perceptual properties of sound is estimated through pitch whereas formants provide the acoustic reverberation of the human vocal tract. The variability in features of user/speaker utterance of words is represented with symbolic data. The speaker identification is performed using modified span, content and position symbolic similarity measures [3], modified for the current work. The proposed method is evaluated on 100 users of voice corpus of VTU-BEC-DB multimodal biometric database and achieves an overall identification rate of 90.56%.

Keywords: Speaker identification · Symbolic object · Voice biometric
Complementary spectral features · Symbolic similarity measure

1 Introduction

Speaker recognition system identify an individual by analyzing the human voice. Voice is a biometric trait which exhibits combined behavioral and physiological characteristics used to alleviate the problem of spoof attack in biometric system. Voice of an user is unique as the knowledge/codes used in the utterances are user specific. The physical differences in users' voiceprints are characterized by measuring the amplitude, frequency, duration and spectral distribution. The information embedded in the voice signal is extracted at six different levels such as spectral, prosodic, phonetic, idiolectal (i.e. syntactical), dialogic and semantic [1]. The usage of the text in the testing phase of the speaker recognition system categorize it as text-dependent and text-independent

system. Text-dependent systems are further classified into fixed-phrase and prompted-phrase systems. The phrase employed for training is also used for testing in Fixed-phrase system whereas the word sequence (phoneme sequence) is generated by the system and ask the user/claimant to utter the same in Prompted-phrase systems. Further the speech employed for testing in text-independent system is unconstrained. In voice biometrics the speaker voiceprint may vary due to variations in the health, environmental conditions and additive noise. The features extracted in such situations form the speaker voiceprints are varying in nature. The symbolic object representation is employed to represent such variability in the features of voiceprints during speaker recognition in a robust manner.

Symbolic objects are extensions of classical data types. The real world objects are better described with symbolic objects [2, 3]. The feature extracted from the real world objects are usually represented by complex data. The knowledge embedded in the complex data is easily extracted by representing them into symbolic data structure. In pattern recognition, the variability inside classes of individuals is easily expressed by symbolic data. Symbolic objects are of three different types, assertion object, hoard object and synthetic objects [4]. In this work voiceprints are represented as assertion symbolic object. This representation of the symbolic object accommodates the variability in features of speaker voiceprint and is one of the novel contribution of the proposed speaker recognition system.

In the proposed work, the text-dependent speaker recognition system is presented, in which the speaker utterance is represented as symbolic object. The object will cover the features of the speaker voice utterance such as inter-lexical pause position, complementary spectral features such as spectral entropy, spectral centroid and spectral flatness, pitch, loudness and formant frequencies. The intra-speaker variations in the features are captured in a symbolic data structure. This representation of speaker utterance into symbolic objects is a novel technique used by the proposed system. The symbolic knowledge bases for the phrases of English number utterance namely "Twenty One (21)" to "Twenty Nine (29)" are constructed separately for 100 users. Further, speaker identification is performed using span, content and position symbolic similarity measures adopted from [2–4]. The experimentation is performed on voice corpus of VTU-BEC-DB multimodal biometric database. The experimental results show that the proposed method offers a overall correct identification rate of 90.56% for person recognition using voice trait.

The rest of the paper is organized as follows: Sect. 2 presents the recent developments in text-dependent speaker recognition approaches. Section 3 describes the proposed model of user identification using voice symbolic objects. The experimental results and analysis are provided in Sect. 4. Finally, Sect. 5 concludes the work and enlists the future directions.

2   Related Work

Voice is versatile, simple to use, non-intrusive and has high user acceptance as a biometric trait. The voice-tract and accent characteristics are difficult to duplicate even if obtained from a recorded voice. Some of the works related to the voice biometric are

described in the following. For text-dependent applications, whole phrases or phonemes may be modeled using multistate left-to-right hidden morkov models (HMM) [5]. The mel-frequency cepstral coefficients (MFCC) feature has been used for designing a text dependent speaker identification system. The MFCC features of a speaker are quantized to a 23 number of centroids using vector quantization algorithm [6]. The log Mel spectrum is converted into time domain using discrete cosine transform (DCT) coefficients referred as acoustic vectors [7]. The use of discrete wavelet transformation (DWT) in multi-level decomposition helps to represent the meaningful features of the human voice, in low size coefficients [8]. Also Linear Prediction Coefficients (LPC), LPC-Cepstral (LPCC) and perceptual linear prediction (PLP) coefficients have been used to extract the features [9]. The gammatone-filter-based method is employed for auditory feature extraction and the features are tested under white noise, babble noise, tank noise and F-16 cockpit noise mismatched conditions, at different SNR levels [10].

A joint factor analysis (JFA) for text-dependent speaker recognition with random digit strings is presented in [11]. The logistic regression method and extraction of phonetically aware Baum–Welch statistics are employed for speaker recognition. An utterance-level speaker representation is performed with attention network by combining the frame-level features extracted with convolution neural network in [12]. The use of i-vector/ Probabilistic linear discriminant analysis with HMM for text-dependent single utterance task is presented in [13]. A lexicon-based local representation algorithm for text-dependent i-vector speaker verification system is presented in [14]. Gaussian mixer model-based support vector machine (GMM-SVM) and the nuisance attribute projection technique for channel compensation is employed in [15].

The different key challenges in voice biometrics are robustness, adaptation, portability, language modeling, out-of-vocabulary words, confidence measures, prosody, spontaneous speech and modeling dynamics. In order to develop a robust speaker recognition system there is a need of accurate and efficient representation techniques for recognizing voice biometric patterns. Hence, there is scope to develop an efficient representation technique for voiceprint features. In the proposed speaker recognition system the speaker utterance is represented as symbolic object to express variability in the features of voiceprints. The detailed description of the proposed methodology is given in the next sections.

3   Proposed Model for Speaker Recognition

In this work a new model for text-dependent speaker recognition using symbolic object modelling and similarity is proposed. The steps involved in the speaker recognition using proposed method are, voice signal pre-processing i.e. framing, windowing and voice activity detection, feature extraction, symbolic object representation, construction of the knowledge base and finally recognition with the help of symbolic similarity measure. The overall methodology is depicted in Fig. 1. Firstly, the voiceprints are pre-processed in which the input user utterance is segmented with framing and windowing technique and the silence portion is detected and eliminated by analyzing the segmented frames by voice (speech) activity detection. To characterize the speaker,
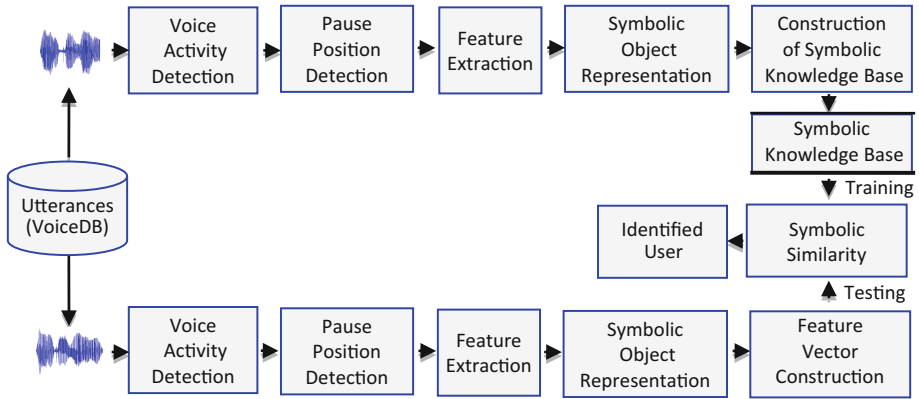
Fig. 1. Speaker recognition system using symbolic modeling of voice features

features are extracted from the silence removed signal. The inter-lexical pause position i.e. start and end points of pause, is located by tracing the amplitude spectrum of the silence removed signal. The complementary spectral features are computed from the power spectrum of the silence removed signal. The loudness feature is computed from the silence removed signal. The utterance of the keywords "21" to "29" of English language from 100 user's of VTU-BEC-DB voice corpus is considered for experimentation. The extracted features are used to construct symbolic data structure. Further, the symbolic knowledge base is constructed for 100 users by selecting five utterances of each English number from every user. Finally, the identification of the user is performed with the help of span, content and position symbolic similarities. The detailed description of each module of the proposed system is presented in the following subsections.

3.1   Description of VTU-BEC-DB Voice Corpus

The proposed speaker recognition system uses voice utterances from VTU-BEC-DB Voice Corpus. The voice corpus contains total of 31000 number utterance of zero (0) to Thirty (30) collected separately in English and Kannada languages from the 100 persons (i.e. 36 males and 64 females) with age range between 18 and 50 years collected in two sessions over a period of one year. The voice samples are recorded with sampling frequency of 44.1 kHz stereo in Linear PCM (LPCM) format with bit rate of 16 bit wave files. The Microphone Sensitivity is set to Medium and the NCF (Noise Cut) recording filter is used. Further, the silence removed signal is obtained by processing the collected voiceprint with steps described in next subsection.

3.2   Voice Detection and Silence Portion Elimination

During the acquisition of voiceprint sometimes the silence portion will occur due to delayed utterance of the speaker or due to improper handling of the acquisition device. The first step in audio signal processing and speaker recognition is to eliminate the

unwanted/silence information from the voiceprints. In the proposed work, to extract the stationary information of the input user voiceprint, it is partitioned into 'N' number of frames with each frame of 20 ms length, by using framing and windowing technique. The presence of silence portion in the frames is detected with the help of voice activity detection i.e. speech activity detection or speech detection technique [16] using Eq. (1).

$$V_i(n) \underset{1 \le i \le N}{=} \left(\sum_{k=1}^{n} | F_i(k) - \left(\sum_{i=1}^{N} \sum_{k=1}^{n} F_i(k) \right) / N | \right) / \sigma \qquad (1)$$

where $F_i(k)$ is the kth sample in the ith frame, N = Total number of frames (N = Duration of Utterance/Duration of the Frame), and σ is standard deviation of entire voice signal. The frame is identified as either voiced and unvoiced frames based on the volume-threshold as given in Eq. (2)

$$V(i) = \begin{cases} Voiced & if\ V(i) \ge volume - threshold \quad 1 \le i \le N \\ UnVoiced & Otherwise \end{cases} \qquad (2)$$

Based on the experimentation the volume-threshold is taken as 5 mv. Finally all the voiced frames are combined to obtain the silence removed signal, SR(n). Further, the SR(n) is used to extract the features as described in next subsection.

3.3   Pause Position Detection and Feature Extraction

The pauses in user utterance will occur due to weak respiration, low muscular tone and slow articulatory rate. The pauses are broadly classified into physical and linguistic and psychological and psycholinguistic [17]. Inter-lexical pauses will appear between utterance of two words in the voice signal. The articulation information embedded in the user utterance is explored by detecting the pause position in the frames of the voice signal. In the proposed work, in order to locate the inter-lexical pause position the SR (n) is partitioned in to 'M' (i.e. M = 40) non-overlapping frames as each frame contains the stationary information of SR(n). The number of samples per frame 'n' in each frame is computed by sampling frequency of 'fs' using

$$n = fs/M \qquad (3)$$

Further, the sliding window (w) with a size 'wl' is devised to locate the pause portion in the frame according to the Eq. (4)

$$wl = \lceil (n/10) \rceil \qquad (4)$$

The sliding window is initially positioned at the left of the ith frame (i = 1) to be analyzed and sum of the amplitude of the overlapped samples of the frame is computed. The process is repeated over the entire frame by shifting the sliding window right by one sample each time till the window reaches to the right edge of the frame. This traversing of the sliding window is repeated for all remaining M − 1 frames. The sum

of the amplitude (AW) of the overlapping sliding window is computed for M frames as given in the Eq. (5)

$$AW(i) = \sum_{j=1}^{j+wl-1} F_i(j) \quad \forall \ 1 \le i \le M \tag{5}$$

The corresponding starting and ending position of the sliding window over the frames is calculated by Eqs. (6) and (7) and are stored for further processing.

$$Win\_Pos\_S(i) = \sum_{j=1}^{j+wl-1} (1/fs) * j \quad \forall \ 1 \le i \le M \tag{6}$$

$$Win\_Pos\_E(i) = \sum_{j=1}^{j+wl-1} (1/fs) * (j+wl-1) \quad \forall \ 1 \le i \le M \tag{7}$$

The pause portion is identified by finding the minimum value of the AW(i) as given by Eq. (8)

$$Min\_AW(i) = min(AW(i)) \quad \forall \ 1 \le i \le M \tag{8}$$

The starting and ending position values associated with frame samples with Min_AW(i) viz. Win_Pos_S(i) and Win_Pos_E(i) are used to trace the start point (S) and end point (E) of the inter-lexical pause as shown in Fig. 2. The frames with minimum amplitude in SR(n) are shown in Fig. 3.

Further in order to extract the complementary spectral features and pitch features the amplitude spectrum (i.e. FFT) of the SR(n) is normalized by dividing each value of the spectrum by the maximum value of the spectrum. The normalized amplitude spectrum AS(k) is partitioned into P (i.e. P = 40) overlapping frames as each frame contains the stationary information of the amplitude spectrum where $1 \le k \le r$ i.e. r = number of spectral samples per frame (i.e. $r = fs/P$).
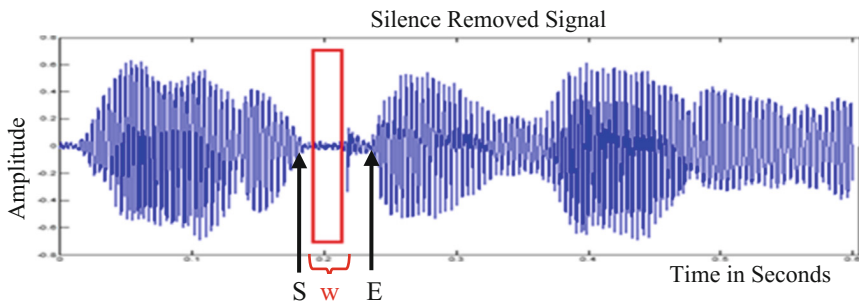
Fig. 2. Pause portion detection with sliding window
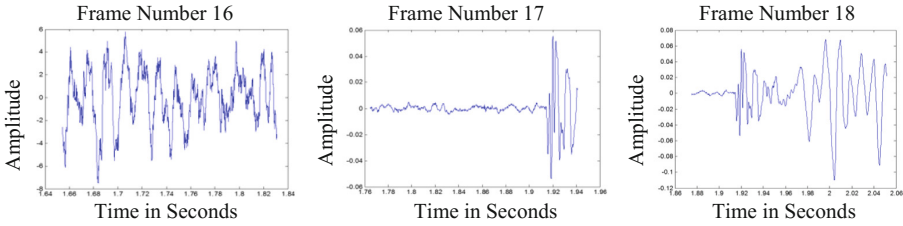
Fig. 3. Frames with minimum amplitudes for pause portion detection

Let $AS_i(k)$ be the amplitude spectrum of the ith frame in spectral domain. The power spectrum of ith frame $PS_i(k)$ is obtained by Eq. (9)

$$PS_i(k) = \sum_{k=1}^{r} |AS_j(k)|^2 \tag{9}$$

Let, the power spectrum of ith frame $PS_i(k)$ be segmented into 'R' non-overlapping sub-bands/blocks to consider each block as regularly repeating function with shorter period of time and spectrum becomes locally stationary (i.e. R = 10 in this work). Where each block 'b' with a size 'bw' is computed using Eq. (10)

$$bw = \lceil (r/10) \rceil \tag{10}$$

Further, complementary spectral features for i^{th} frame is computed from the power spectrum $PS_i(k)$ as described in the following.

3.3.1 Spectral Entropy or Shannon Entropy
Spectral entropy or Shannon entropy (SE) describes irregularity, complexity or level of uncertainty of a signal [18]. The power spectral density of each block of the ith frame of power spectrum is calculated by normalizing according to Eq. (11)

$$PSi, b\_Norm(k) = PS_i(k)/R \tag{11}$$

The normalized energy of the each block of the ith frame of the power spectrum is treated as a probability distribution $Pr_{i,b}(k)$ for calculating entropy by Eq. (12)

$$Pr_{i,b}(k) = PS_{i,b\_Norm}(k) / \sum_{b=1}^{R} PS_{i,b\_Norm}(k) \tag{12}$$

The spectral entropy of the power spectrum of SR(n) is calculated by Eq. (13)

$$SE = -\sum_{i=1}^{P} \sum_{b=1}^{R} Pr_{i,b}(k) \; log_2 \; Pr_{i,b}(k) \tag{13}$$

The description of the spectral centroid feature is given below.

3.3.2 Spectral Centroid

The spectral centroid (SC) measures the spectral shape and median (i.e. centre of gravity) of a sound spectrum [19]. The spectral centroid of each block of the ith frame of the power spectrum is calculated by Eq. (14)

$$SC_{i,b} = \sum_{k=1}^{bw} k * PS_i(k) / \sum_{k=1}^{bw} PS_i(k) \tag{14}$$

The spectral centroid of the power spectrum of SR(n) is calculated by Eq. (15)

$$SC = \sum_{i=1}^{P} \sum_{b=1}^{R} SC_{i,b}(k) \tag{15}$$

Further, the description of the spectral flatness feature is given below.

3.3.3 Spectral Flatness or Wiener Entropy

The spectral flatness or Wiener entropy estimates the uniformity in the frequency distribution of the power spectrum [20]. The spectral flatness of each block of size 'bw' in the ith frame of the power spectrum is calculated by Eq. (16)

$$SF_{i,b} = (\prod_{k=1}^{bw} PS_i(k))^{1/bw} / \frac{1}{bw} \sum_{k=1}^{bw} PS_i(k) \tag{16}$$

The spectral flatness of the power spectrum of SR(n) is calculated using Eq. (17)

$$SF = \sum_{i=1}^{P} \sum_{b=1}^{R} SF_{i,b}(k) \tag{17}$$

After extracting the complementary spectral features, the pitch frequency is estimated which is described in the following.

3.3.4 Pitch

The pitch (F0) estimates the relation between physical and perceptual properties of sound [21]. The auto-correlation sequence of the entire frames of the normalized amplitude spectrum of the signal is computed by Eq. (18)

$$AC = \sum_{y=0}^{P-1} \sum_{x=1}^{P} AS_i(x) * AS_i(x + y) \tag{18}$$

The peak of the auto-correlation(AC) sequence is computed and its associated sample number (Max$_{\_AC\_Sam}$) is obtained to estimate the pitch frequency (PF).

$$PF = fs / Max_{\_AC\_Sam} \tag{19}$$

The pitch also depends on loudness and spectrum. The computation of loudness feature is described in the following.

3.3.5   Loudness

Loudness describes the strength of the ear's perception of a sound. Different models exist in literature to estimate loudness based on the type of sound i.e. stationary, non-stationary and impulsive sounds [22]. The transmission of the acoustic signal through the outer and middle ear by calculating the basilar membrane excitation by critical bandwidth over 24 critical bands is modelled by Zwicker [23]. A partial loudness (N') in Sone/Bark is calculated with a power law as in Eq. (20).

$$N' = c * (E_{Th})^\alpha * [(\frac{1}{2} + \frac{1}{2}(E_{Stim})/(E_{Th}))^\alpha - 1] \tag{20}$$

where, 'c' is the constant and 'α' is the exponent (i.e. $\alpha < 1$) both are independent form frequency, E_{Stim} is the excitation produced by the input stimulus and E_{Th} is the excitation at threshold in quiet. Based on the experimentation $c = 0.1$, $\alpha = 0.25$ and $E_{Th} = 0.32$ are selected in this work. The final value for loudness in 'sones' is calculated using Eq. (21)

$$Loudness = \sum_{z=0}^{24} N' \tag{21}$$

Where z is the critical band rate (measured in Bark).

3.3.6   Formants

Formants are characterized as the spectral peaks of sound range, of the voice, of a person. The Linear predictive coding system (LPC) has been utilized for estimation of the formant frequencies [24]. Linear prediction estimates the vocal tract filter and shows the resonances that shape the formants of the speech signal. The order of the LPC prediction coefficients is about 4 + fs/1000 (i.e. 48 for this work) for all-pole filter (i.e. autoregressive) model of the voice source [25]. The frequencies of first three consecutive peaks of log spectra are extracted to represent the formants F1, F2 and F3 respectively as they contain sufficient information of the vowels and voiceprint. The variations in the extracted features of user voice is represented using symbolic object as described in the next subsection.

3.4   Symbolic Object Representation of Voice Signal

Symbolic objects, offer a formal methodology to represent the variability in features of user/speaker utterance. The features extracted from the user utterance are described as a assertion symbolic object as given by Eq. (22).

$$Assertion\,Object = \begin{array}{l} [\{a_{min}, a_{max}\}, \{b_{avg}\}, \{c_{1min} - c_{1max}\}, \{c_{2min} - c_{2max}\}, \\ \{c_{3min} - c_{3max}\}, \{d_{min} - d_{max}\}, \{e_{avg}\}, \{f_{min} - f_{max}\}, \\ \{g_{1min} - g_{1max}\}, \{g_{2min} - g_{2max}\}, \{g_{3min} - g_{3max}\}] \end{array} \tag{22}$$

The knowledgebase is constructed separately for the English numbers utterances of the phrases "Twenty One" to "Twenty Nine" by selecting the five voiceprints for each phrase from every user. The symbolic object representation associated with each phrase of every individual 'x_i' associated with the ith utterance i.e.$1 \leq i \leq 5$ is given in Table 1.

Table 1. Symbolic object representation of voice features

$a_{min} = min\{Pause\_Position\ (x_i)\}$	$a_{max} = max\{Pause\_Position\ (x_i)\}$
$b_{avg} = mean\{Spectral\_Entropy\ (xi)\}$	$c_{1min} = min\{StdSCmin\ (x_i)\}$
$c_{1max} = max\{StdSCmax\ (x_i)\}$	$C_{2min} = min\{SCmin(xi)\}$
$c_{2max} = max\{SCmin\ (x_i)\}$	$c_{3min} = min\{SCmax\ (x_i)\}$
$c_{3max} = max\{SCmax\ (x_i)\}$	$d_{min} = min\{SF(x_i)\}$
$d_{max} = max\{SF(x_i)\}$	$e_{avg} = mean\{Loudness\ (xi)\}$
$f_{min} = min\{PF(x_i)\}$	$f_{max} = max\{PF(x_i)\}$
$g_{1min} = min\{F1(x_i)\}$	$g_{1max} = max\{F1(x_i)\}$
$g_{2min} = min\{F2(x_i)\}$	$g_{2max} = max\{F2(x_i)\}$
$g_{3min} = min\{F3(x_i)\}$	$g_{3max} = max\{F3(x_i)\}$

The description of the symbolic data structure for the English number utterance of the phrase "TwentyTwo" is depicted in Table 2.

Table 2. Sample symbolic data structure for number utterance "22"

Symbolic object	$\{a_{min,}$ $a_{max}\}$	b_{avg}	$\{c_{1min}-$ $c_{1max}\}$	$\{c_{2min}-$ $c_{2max}\}$	$\{c_{3min}-$ $c_{3max}\}$	$\{d_{min}-d_{max}\}$	e_{avg}	$\{f_{min}-$ $f_{max}\}$	$\{g_{1min}-$ $g_{1max}\}$	$\{g_{2min}-$ $g_{2max}\}$	$\{g_{3min}-$ $g_{3max}\}$
Assertion object	0.1769– 0.2322	1.3253	0.0096– 0.01815	0.0334– 0.0532	0.0758– 0.0882	5.0709e−5– 1.3421e−4	42.3441	228.497– 238.378	271.36– 328.51	441.52– 750.36	1967.87– 2287.01
Measuring unit	Sec	–	–	–	–	dB	Sone	Hz	Hz	Hz	Hz

Further, for every user nine assertion objects are formed separately for the English number utterances from "21" to "29" respectively and the symbolic knowledge base is constructed for 100 users according to the described symbolic data structure. This symbolic knowledge base is then used for user identification and the methodology is described in the following sub-section.

3.5 Methodology for Speaker Recognition

In the proposed work, the symbolic data object is constructed by computing the quantitative features like interval values and discrete values of the voiceprints. The symbolic similarity measure i.e. content, span and position similarity measures modelled as in [2–4] are computed to evaluate the nearest category of the user for identification. The symbolic similarity between the test feature vector (TFV) constructed for testing voiceprint and the ith user in the symbolic knowledge base (SKB$_i$) is written as Eq. (23)

$$S(TFV, SKB_i) = S_{Span}(TFV, SKB_i) + S_{Content}(TFV, SKB_i) + S_{Position}(TFV, SKB_i)$$
(23)

The similarity component due to 'span' indicates the relative sizes of the feature values without referring to common parts between them. In order to compute the similarity component between the TFV and the ith user object in the SKB$_i$ due to span, the following values are calculated,

$$\left.\begin{aligned}
A_l &= min(SKB_i) \\
A_u &= max(SKB_i) \\
B_l &= min(TFV) \\
B_u &= max(TFV) \\
L_a &= |A_u - A_l| \\
L_b &= |B_u - B_l| \\
L_s &= |max(A_u, B_u) - min(A_l, B_l)|
\end{aligned}\right\}$$
(24)

$$S_{Span}(TFV, SKB_i) = (L_a + L_b)/2L_s$$
(25)

Where, A_l is the lower limit of symbolic knowledge base of ith user, A_u is the upper limit of symbolic knowledge base of ith user, B_l is the lower limit of test feature vector, B_u is the lower limit of test feature vector, L_a is the length of interval of feature vector, and L_b represents length of interval of features of symbolic knowledge base. And L_s corresponds to centered span length of both TFV and SKB$_i$.

The similarity component due to 'content' is a measure of the common parts between two feature values. Further, to compute the similarity component between the TFV and the ith user in the SKBi due to content, the following values are calculated,

$$CF(TFV, SKB_i) = \begin{cases} 1 & if \ |TFV - SKB_i| \leq 0.05 \quad \forall 1 \leq i \leq N \\ 0 & Otherwise \end{cases}$$
(26)

where CF(TFV, SKB$_i$) is the common features(CF) between TFV and the ith user in the SKB$_i$. Empirically the level of significance is assumed to be 0.05 in this work.

$$inters = \sum_{i=1}^{N} CF(TFV, SKB_i) \quad \forall \ 1 \leq i \leq N$$
(27)

$$Ls = (L_a + L_b - inters)$$
(28)

Where, N = Number of users i.e.100, L_a and L_b are computed using Eq. (24) and *inters* represents the number of common features between TFV of the input voiceprint and SKB. The content similarity measure is computed by Eq. (29)

$$S_{Content}(TFV, SKB_i) = inters/L_s$$
(29)

The similarity components due to "position" indicates the distance of one object to the initial position of other object. Further, to compute the similarity component between the TFV of the input voiceprint and the ith user in the SKB$_i$ due to position is computed using Eq. (30)

$$S_{Position}(TFV, SKB_i) = 1 - |(A_l - B_l)|/U_k \qquad (30)$$

where U_k is the length of the maximum interval of kth feature of SKB$_i$ computed using Eq. (31)

$$U_k = |A_l - A_u| \qquad (31)$$

The A_l and A_u are computed using Eq. (24). The symbolic similarity between the TFV and the ith user in the SKB$_i$ is computed by using the Eq. (23). As the identification is one-to-many matching, the combined similarity between the TFV and the SKB of all the N (i.e. 100) users is computed by adding the span, content and position similarity values calculated using the Eqs. (25), (29) and (30) respectively and represented as a $1 \times N$ vector. Finally the best matching utterance is identified using the maximum similarity value. The evaluation is carried out for 100 users and performance is brought out, and is discussed in next section.

4   Experimentation

The performance of the proposed speaker recognition method is evaluated on voice samples from VTU-BEC-DB Voice Corpus of 100 users. For the experimentation, the English number utterance from "21" to "29" of 100 persons are considered as they contain the inter-lexical pauses. Every utterance is modelled as assertion symbolic object by extracting the voice features. From each user five utterances of first session are used for constructing the symbolic knowledgebase and the five utterances of second session are employed for testing. Hence, for 100 users the symbolic knowledgebase with a collection of assertion symbolic objects is constructed. During testing, the symbolic data representation of the user is compared with the symbolic knowledgebase using span, content and position symbolic similarity measure. Finally the best matching utterance in symbolic knowledge base is identified using the maximum symbolic similarity. The system performance is given in Table 3.

Table 3. Recognition performance with symbolic similarity

Sl. no.	English number utterance	Speaker identification rate (%)	Sl. no.	English number utterance	Speaker identification rate (%)	Sl. no.	English number utterance	Speaker identification rate (%)
1	21	91.60	4	24	90.60	7	27	90.40
2	22	92.20	5	25	89.80	8	28	90.00
3	23	90.60	6	26	89.60	9	29	90.20
Average of speaker identification rate = 90.56%								

The experimentation for the identification of user with symbolic similarity is conducted by varying the rank 'k' value in the range of $1 \leq k \leq 100$. The overall speaker identification rate of 90.56% achieved for the rank k = 1. The performance of speaker identification for 100 users with English number utterance from Twentyone (21) to Twentynine (29) is depicted pictorially in Fig. 4. By using the 4500 voice samples of 100 users by considering English number utterances of VTU-BEC-DB database, the overall correct identification rate of 90.56% is achieved with the help of symbolic similarity measure.
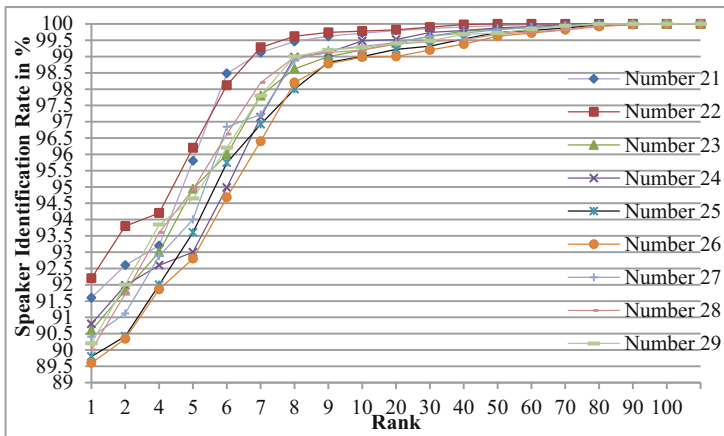
Fig. 4. Identification performance of the proposed speaker recognition system

5   Conclusion

In this paper, a robust text-dependent speaker recognition system using symbolic modeling of voiceprint is presented. The symbolic data object is constructed using the voice features namely the inter-lexical pause position, complementary spectral features i.e. spectral entropy, spectral centroid and spectral flatness, pitch, loudness and formants. The user recognition is performed using a new symbolic similarity metric i.e. span, content and position. The symbolic data representation of the characteristics of user voiceprint is the novel feature of this work. The proposed method is experimented with the voice samples of the English number utterance separately by considering the phrase "Twenty One" to "Twenty Nine" of VTU-BEC-DB voice corpus. The overall speaker identification rate of 90.56% is achieved at rank-1 with symbolic modelling and symbolic similarity analysis technique. In future, extensive experiments of the proposed method can be performed to make it more robust.

References

1. Faundez-Zanuy, M., Monte-Moreno, E.: State-of-the-art in speaker recognition. IEEE Aerosp. Electron. Syst. Mag. **20**(5), 7–12 (2005)
2. Angadi, S.A., Kagawade, V.C.: A robust face recognition approach through symbolic modeling of polar FFT features. Pattern Recogn. **71**, 235–248 (2017)
3. Gowda, C.K.: Symbolic objects and symbolic classification. In: International Conference on Symbolic and Spatial Data Analysis: Mining Complex Data Structures, pp. l–18 (2004)
4. Nagabhushan, P., Angadi, S.A., Anami, B.S.: Symbolic data structure for postal address representation and address validation through symbolic knowledge base. In: Pal, Sankar K., Bandyopadhyay, S., Biswas, S. (eds.) PReMI 2005. LNCS, vol. 3776, pp. 388–394. Springer, Heidelberg (2005). https://doi.org/10.1007/11590316_59
5. Jourlin, P., Luettin, J., Genoud, D., Wassner, H.: Integrating acoustic and labial information for speaker identification and verification. In: Fifth European Conference on Speech Communication Technology, pp. 1603–1606 (1997)
6. Rabiul, I.M., Fayzur, R.M.: Improvement of text dependent speaker identification system using neuro-genetic hybrid algorithm in office environmental conditions. Int. J. Comput. Sci. Issues **1**, 42–47 (2009)
7. Dash, K., Padhi, D., Panda, B., Mohanty, S.: Speaker identification using mel frequency cepstral coefficient and BPNN. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **2**(4), 326–332 (2012)
8. Erzin, E., Yemez, Y., Tekalp, A.M.: Multimodal speaker identification using an adaptive classifier cascade based on modality reliability. IEEE Trans. Multimed. **7**(5), 840–852 (2005)
9. Aladwan, A.A., Shamroukh, R.M., Aladwan, A.: A novel study of biometric speaker identification using neural networks and multi-level wavelet decomposition. World Comput. Sci. Inf. Technol. J. **2**(2), 68–73 (2012)
10. Li, Z., Gao, Y.: Acoustic feature extraction method for robust speaker identification. Multimed. tools Appl. **75**(12), 7391–7406 (2016)
11. Stafylakis, T., Jahangir, A.M., Kenny, P.: Text-dependent speaker recognition with random digit strings. IEEE/ACM Trans. Audio Speech Lang. Process. **24**(7), 1194–1203 (2016)
12. Zhang, S-X., Chen, Z., Zhao, Y., Li, J., Gong, Y.: End-to-end attention based text-dependent speaker verification. In: IEEE Workshop on Spoken Language Technology, pp. 171–178 (2017)
13. Büyük, O.: Sentence-HMM state-based i-vector/PLDA modelling for improved performance in text dependent single utterance speaker verification. IET Sig. Process. **10**(8), 918–923 (2016)
14. You, H., Li, W., Li, L., Zhu, J.: Lexicon-based local representation for text-dependent speaker verification. IEICE Trans. Inf. Syst. **E100–D**(3), 587–589 (2017)
15. Sun, H., Kong, A.L., Ma, B.: A new study of GMM-SVM system for text-dependent speaker recognition. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 4195–4199 (2015)
16. Jang, R., Shing, J.: End point detection, pp. 1–23. MIR lab, CSIE department national Taiwan university, Taiwan (2015)
17. Zellner, B.: Pauses and the temporal structure of speech. In: Fundamentals of Speech Synthesis and Speech Recognition, pp. 41–62 (1994)
18. Misra, H., Shajith, I., Bourlard, H., Hermansky, H.: Spectral entropy based feature for robust ASR. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 1–6 (2004)

19. Paliwal, K.K.: Spectral subband centroid features for speech recognition. In: IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 2, pp. 617–620 (1998)
20. Tzanetakis, G., Cook, P.R.: Musical genre classification of audio signals. IEEE Trans. Speech Audio Process. **10**, 293–302 (2002)
21. Rouat, J., Liu, Y.C., Morissette, D.: A pitch determination and voiced/unvoiced decision algorithm for noisy speech. Speech Commun. **21**, 191–207 (1997)
22. Esben, S., Soren, H.N.: Evaluation of different loudness models with music and speech material. In: Audio Engineering Society 117 Convention, pp. 1–34 (2004)
23. Zwicker, E., Fastl, H., Widmann, U., Kurakata, K., Kuwano, S.N.: Program for calculating loudness according to DIN 45631 (ISO 532B). J. Acoust. Soc. Jpn. **12**, 39–42 (1991)
24. Roy, S.C., Fausto, M.: Formant location from LPC analysis data. IEEE Trans. Speech Audio Process. **1**(2), 129–134 (1993)
25. Rabiner, L.R., Schafer, R.W.: Introduction to digital speech processing. Found. Trends Sig. Process. **1**(1–2), 1–194 (2007)

An IOT Based Smart Shopping Cart for Smart Shopping

Srinidhi Karjol^(✉) , Anusha K. Holla^(✉) , C. B. Abhilash^(✉) ,
P. V. Amrutha, and Y. V. Manohar

JSS Academy of Technical Education, Bengaluru, Bengaluru, India
srinidhikarjol@gmail.com, anusha.k.holla@gmail.com,
abhilashcb@jssateb.ac.in

Abstract. Today's world have a fast growing population with a wide range of demand from a variety of domains. Customers who need to purchase different products in Walmart or supermarkets needs lots of time and patience in coordinating among them self for successful shopping. We need to address this problem by efficiently using our technologies. In the advancement of technologies, the world is getting automated in many aspects. In this Paper, we depict reasonable and cost-effective Smart Shopping Cart utilizing IoT (Internet of Things) innovations. Such a framework is appropriate for use in spots such as Walmart & supermarkets, where it can help in lessening work and in making a superior shopping knowledge for the clients. Rather than influencing the clients to sit tight in a long line for looking at their shopped things, this framework helps in mechanizing the easy and comfortable billing process. The shopping is processed with two aspects, with a predefined list and random shopping. Our proposed system provides the nearest route to pick-up the listed items present in different racks of the Walmart. Also, with the added feature we have an approach where Cart-to-Cart communication is enabled that allows a customer to share their shopping list with co-shopper to enable parallel shopping using two or more carts. These features save time and make shopping easy. Along with these abilities, this system design is also capable of detecting theft by shoplifters. In addition, the Walmart or supermarket management will be able to analyze the shopping behaviors of various customers to arrive at valuable business insights. These will be very beneficial for the retail stores. Accordingly, the management team will have the ability to predict the rate of sales of all individual products and make the stock available is based on the ongoing customer requirements. Overall, this system will ensure that the customers will have the best shopping experience and very often, they visit the Walmart for the shop.

Keywords: Shopping cart · IOT · Cart communication · Automation
Business analysis

1 Introduction

With the increasing reliability and cost effectiveness of—Internet of Things (IoT) based connected smart things in the field of consumer applications, it makes better sense to ensure such technologies are put to use in addressing the day-to-day concerns of the

common man. In this framework, we portray the execution of a dependable, reasonable and cost effective Smart Shopping Cart. Such a framework is reasonable for use in any Brick and Mortar shopping spots, for example, general stores, where it can help in diminishing work and in making a superior shopping background for its clients. Rather than influencing the clients to sit tight in a long line for looking at their shopped things, the framework helps in mechanizing the charging procedure. Alongside this capacity, the framework configuration additionally guarantees identification of instances of duplicity conjured by deceptive clients, which influences the savvy framework to reasonable and alluring to both the purchasers and merchants. The framework outline alongside the execution is exhibited here. The outcomes are empowering and make shopping less demanding and helpful to the clients. The fundamental target of the proposed framework is to give an innovation arranged, keen, ease, adaptable and rough framework for a superior in-shop involvement for the cutting edge world client.

Walmart could at long last end the disappointment of not having the capacity to locate the last sustenance thing on your shopping list - and ending up totally dumbfounded the second you set foot in one of their stores. It is working with a mechanical autonomy firm to create "driverless" shopping baskets that enables clients to examine in their shopping list - and after those aides them to the correct path and rack.

This new advancement is a route for Wal-Mart to contend with the comfort that Amazon and other online stores offer buyers.

Wal-Mart says utilizing the automatons empowers it to check stock in about a day or less, rather than a month that it takes physically. At the present time, representatives remain on lifts that go here and there the stacks, and output things to ensure that containers are in the perfect place.

As indicated by Walmart's patent demand, clients will have the capacity to summon one of these truck pullers each furnished with cameras and sensors with their user interface gadget, maybe a Smartphone application and a mechanized unit will append to a truck stopped in a docking station and force it to the client. When client and truck meet, the vehicle unit will fill in as individuals.

Some of these issues would be understood if Walmart chooses to proceed with the Dash. It could incorporate with a Walmart shopping-list application on your telephone, for instance, so you could be taken appropriate to the things you have included since your last visit. Now, however, you begin to ask why you don't simply arrange those basic needs on the web and be finished with it. Which is amusing, as endeavors prefer the Dash shopping cart seem to be, says Bloomberg, intended to enable Walmart to finish with online retailers like Amazon [6].

The brilliant shopping cart will be an across the board shopping cart. It will enable the client to monitor the aggregate cost also, when things are added to the shopping cart. The client will be aware of his budget and the offers that are available at the market. It will likewise discuss remotely with an in store segment to make simple installments in a hurry. The client has choice to make easy online payments through the application. Because of any vagueness, the customer will likewise have the alternative of going up to the checkout counters. This new framework would diminish the long hold up times at the checkout counters, increment the productivity of the checkout technique, and would furnish the customer with a la mode cost and aggregate data, which makes the entire experience more helpful. This way it minimizes the labor

required at billing counters hence, reducing the amount spent on the labor. The application can promote live offers, and can be updated as the season changes.

This framework addresses one of the common issues that clients face in the existing system such as unable to find the items in the inventory or employee for any help. The application will help the clients to find items at the right inventory by providing the information about the items in the list along with a route map of the super market thereby providing new experience to the clients.

This will also have the history of the items bought by the client. This will enable client to use the data for next purchase. The supermarket can understand the trend and hence stock the inventory or promote offers accordingly.

This paper aims to outline a framework which peruses the standardized tag on everything that is put in the shopping cart and updates the item data which is accessible to the customer. Weight/Weight sensors will be utilized to distinguish the nearness of new things in the shopping cart. The standardized identification scanner separates the standardized tag which is transmitted to the microcontroller through an USB association. The microcontroller peruses data from a SD card embedded into the microcontroller. This SD card has all the data about the item. This information is at that point organized and exhibited to the client for survey and affirmation on a LCD screen. New things in the shopping cart will be recognized by following the adjustment in the yield of weight sensors.

Similar sensors will be utilized to recognize when things are expelled from the shopping cart. A program will be executed to affirm the expulsion from the customer's shopping basket. Another program will be executed to function as an against burglary system to keep the customer from leaving without an effective installment.

The carts inbuilt programmed charging framework makes shopping a breeze and has other positive turn offs, for example, liberating staff from tedious checkout filtering, lessening an aggregate number of staffs required and expanding operational effectiveness of the framework. In conclusions, we likewise talk about open doors for enhancing the proposed framework to influence it into a monetarily suitable item as a phenomenal approach to help clients to lessen the time spent in shopping by showing the rundown of items, their cost, the best arrangements/rates on the items and programmed charging. The framework assists the store administration with a programmed refresh of the stock on each buy of a thing Intelligent shopping basket (proposed framework) can possibly make shopping more pleasurable and effective for the customer and the stock control less demanding for the store administration. Clever shopping basket (proposed framework) can possibly make shopping more pleasurable and effective for the customer and the stock control less demanding for the store administration.

2 Literature Survey

The presence of Wireless Sensor Network (WSN) advances offers the capacity to gauge and comprehend ecological pointers from sensitive ecologies and normal assets to urban conditions. With the quick increment of these gadgets in a discussing impelling system makes the Internet of Things (IoT), where sensors and actuators mix well with

the earth around us, and the data is shared crosswise over stages keeping in mind the end goal to build up a typical working picture (COP). With the current adaption of an assortment of empowering gadget advances, for example, RFID labels and per users, close field correspondence (NFC) gadgets and installed sensor and actuator hubs, the IoT and is the following progressive innovation in changing the Internet into a completely incorporated Future Internet [1].

With the expanding utilization of wide region Wireless Sensor Networks (WSN) in the field of purchaser applications, it ends up noticeably essential to address the worries, for example, dependability, vitality utilization and cost-viability. This framework is utilized as a part of spots, for example, general stores. It can help in diminishing labor and in making a superior shopping background for the clients. Rather than influencing the clients to hold up in a long line while looking at, the framework robotizes the charging procedure. The client can likewise track the subtle elements of the acquired things and additionally the present bill sum on the screen, which is joined to the truck [2].

Automation of shopping cart using RFID module and ZIGBEE module, in this system, RFID tags are used instead of barcodes. These RFID tags will be on the product. When the customer takes a product and places it in the trolley, the trolley will contain an RFID reader which will sense the RFID tag which is present on the product. Thus displays the product price on the LCD display. Like this, the process continues. Along with it, comes a ZIGBEE transmitter in the trolley, which transfers data to the main computer. The ZIGBEE receiver is placed near the main computer which receives the data from transmitter [3].

This paper also, explains about an automated and a time saving system for retail to improve shopping experience. It is designed to make it customer friendly and secure. In this paper, the proposed smart cart is capable of generating a bill from the cart itself. The customer will make the payment through a credit card which will help to maintain database. This data can be used to introduce schemes and offers in the stores accordingly and also provide additional offers for specific customers. The smart cart uses RFID technology for shopping and payment, AVR microcontroller for peripheral interfacing and inventory management [4–6].

3 Existing System

In this work, our configuration in view of IoT is utilized to address distinctive instances of Super-Market. A Supermarket is a self-benefit shop offering a wide assortment of nourishment and family unit items sorted out into numerous passageways. Clients invest a considerable measure of energy to discover the items they require. In the event that the Customers don't discover the thing or the staff to help them out it is more probable that they leave the general store with no buy, which is thought to be an extraordinary misfortune to the sellers. Disappointment communicated by clients as a result of long holding up time amid the Checkout procedure is another real concern. Controlling the operational expenses is absolutely one of the greatest difficulties that any retailer faces. On the off chance that work cost lessening isn't overseen appropriately, client administration and store conditions may endure. This, obviously, brings

about lost clients and deals. These reason a considerable measure of time wastages, gear wear, and strain on the general population who take the necessary steps. Frequently, wastefulness trouble markets that work on a manual premise. Wasteful aspects are all over and are regularly discovered when items are touched different circumstances or conveyance courses are vague.

4   Proposed System

The basic purpose of innovation in technology, irrespective of the domain, has been in simplifying everyday chores and making it easier, faster and efficient. One trivial task, that human beings spend considerable amount of time, is in shopping. However, the shopping carts are the same, serving simple purposes since they were first manufactured, not undergoing any changes either to design or their purpose. This motivated we to build an innovative shopping cart that not only make the shopping efficient but also enjoyable and convenient to the customers (Fig. 1).
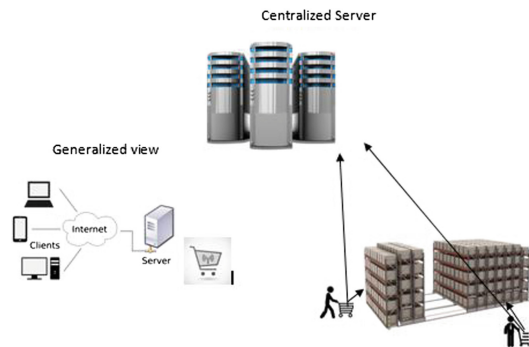
Fig. 1. Main architecture of proposed system

Every Shopping Cart is equipped with a barcode scanner, a camera, a weight sensor, a small computer for local processing and a display device (to save cost, customer's Smartphone will be used as display). The Base Station at a centralized location consists of a database that stores information of all the products, and can communicate with all the Smart Carts via the Wi-Fi network. When a customer starts shopping, she/he has to login with a Customer ID and link the Cart ID with the Customer ID, once registered; they can scan the products one by one with the barcode scanner present at the cart and keep adding them to the cart. In order to handle all the cases of mistake/dishonesty, the design includes the use of weight sensor at the cart.

In addition, cart-to-cart communication is enabled that allows the customers to share their shopping list with their partner thereby enabling them to shop in parallel without duplicating their purchases. Once the customer finishes shopping, she/he then proceeds to the payment counter to pay the bill amount. In addition, the supermarket management will be able to analyze the shopping behaviors of various customers to

arrive at valuable business insights. These will be very beneficial for the retail stores. Overall, this system will ensure that the customers will enjoy the shopping experience and come more often to shop [7] (Fig. 2).
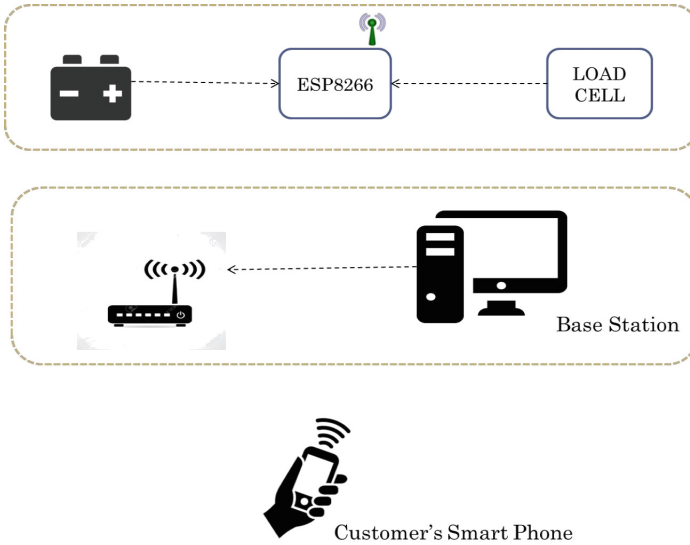
Fig. 2. An overview of the system

Benefits

- Improve the shopping experience for all the customers of the store
- Increase efficiency of the exit process
- Eliminates a long waiting queues at the exit counter

Features

- User interface with LCD monitor for user inputs
- Automated shopping items detection system
- Automated communication system to make payments at counter or via app
- Automated data formatting in case of item deletion or additions and to organization the shopping in a systematic way.

5   Implementation

5.1   Hardware Implementation

Wheatstone Bridge Principle

A Wheatstone bridge is an electrical circuit used to measure an unknown electrical resistance by balancing two legs of a bridge circuit, one leg of which includes the

unknown component. The primary benefit of a Wheatstone bridge is its ability to provide extremely accurate measurements (in contrast with something like a simple voltage divider). (Figs. 3 and 4).
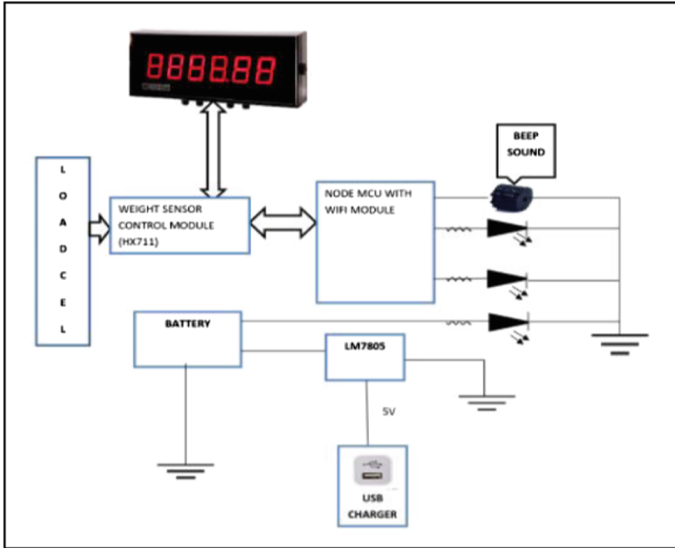
Fig. 3. Circuit diagram
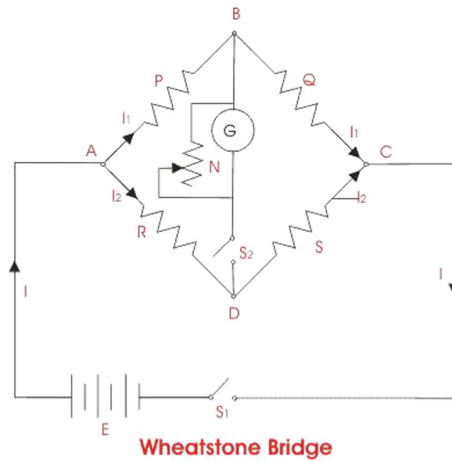
Fig. 4. Wheatstone bridge circuit diagram

NodeMCU Programming

The Arduino Integrated Development Environment - or Arduino Software (IDE) – to program the NodeMCU.
This IDE contains

1. An Editor
2. Text area
3. Text Console
4. A toolbar with common facilities
5. Menus

It interfaces with the Node-MCU equipment to transfer programs and speak with them. Projects composed utilizing Arduino Software (IDE) are called draws. These portrayals are composed in the content tool and are spared with the record extension—. ino. The editorial manager has highlights for cutting/gluing and for looking/supplanting content. The message range gives input while sparing and sending out and shows mistakes. The comfort shows content yield by the Arduino Software (IDE), including complete blunder messages and other data. The base right-hand corner of the window shows the designed board and serial port. The toolbar catches enable you to confirm and transfer programs, make, open, and spare portrays, and open the serial screen.

Pseudocode for the NodeMCU Firmware

Include the Wi-Fi and Serial Port Header File
Define SSID and Password values for Wifi Connection
Setting up the static IP Address for the Node-MCU, so that every time we get the same IP address Setup a Software Serial Port to communicate to the Weight Sensor
Define the GPIO pins used for Led and Beeper

```
In setup ()
    Set Pin Mode for Led and Beeper
    Connect to the Wifi network
    Start the Server
 In loop ()
    Check if client is connected
    If not connected return.
    Otherwise
       ReadWeight()
       Output (weight) //As HTTP response
       Read request if any
       If request is _BEEP=ON'
       Set Beeper to HIGH
       If request is _BEEP=OFF'
       Set Beeper to LOW
  In ReadWeight ()
   //to read weight from serial input
   While (1)
   Do
       For i=0 to 22
            If Serial data available then
                  Read serial data into an array
       End for
       Extract weight value from the array of data
   End while
```

At the point when the client enters the shopping center, he/she will get a trolley on which there will be an RFID.

Peruse, LCD Display, and Android gadget. The trolley will begin its preparing once the merchant presses the `Begin'' catch on the trolley which will be just observed by the merchant. In this way, now the client can utilize the keen trolley for shopping. At whatever point the client places an item into the trolley, the RFID Reader will read the label data and show the related outcomes on the LCD Display. These means are rehashed till the client completes his shopping and the "STOP" catch is squeezed by the seller. Once the "STOP" catch is squeezed there is an alternative gave to end the shopping with the same obtained items or to erase or expel a portion of the items from the trolley. This absolutely relies upon the client decision. Toward the finish of shopping, the client would straight be able to away pay the bill, what's more, leave the mall. Stock status of the items is additionally refreshed toward the finish of shopping. Also, the cart to cart communication enables the best shopping experience to customers and save time.

6   Results and Discussion

The Application user has to register into the application for the first time. This screen asks for his/her name, Email ID and password. Later the user can login using the Email ID and password (Fig. 5).
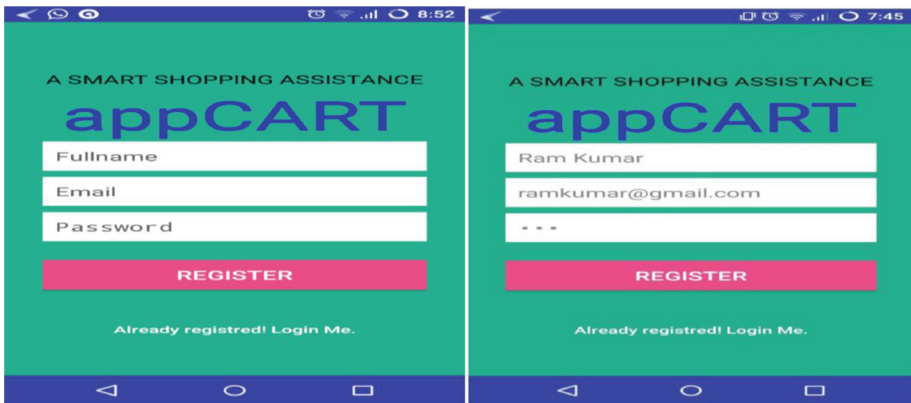
Fig. 5. Registration screen

If user has already registered then he/she can login with the Email ID and password otherwise has to register first and sign in (Fig. 6).

User can choose any mode from this screen. Home mode option allows user to create a list of items. Shop mode allows user to shop using the app. User can also logout from the session (Fig. 7).
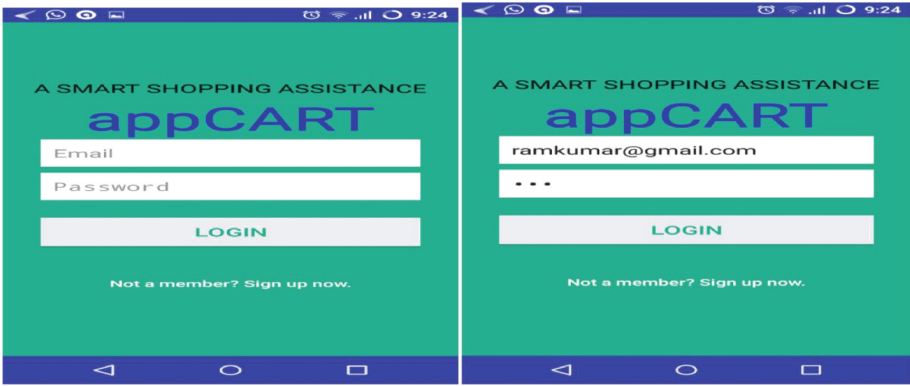
Fig. 6. Login screen
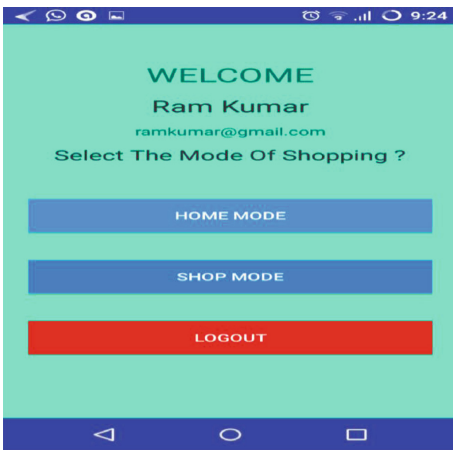
Fig. 7. Welcome screen after login   **Fig. 8.** Home mode screen
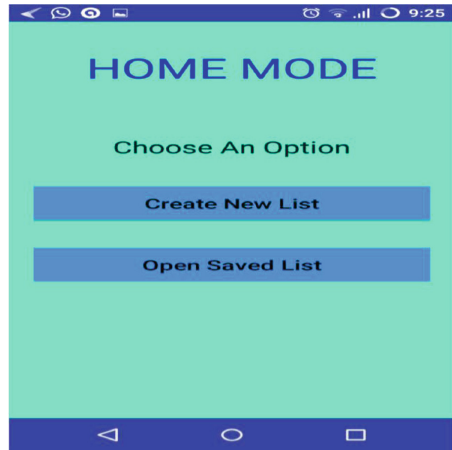
Home mode screen appears when user clicks on home mode button. This provides the user options to create a new list and to open previously saved list (Fig. 8).

The products that are available in the market are available for the user to select and save it as a list (Fig. 9).

The List of items saved by the user is displayed when open saved list button is pressed (Fig. 10).

User can shop by picking random items hence the random shopping option. He/she can shop based on the already saved list (Fig. 11).

The user can shop with his partner in order to save time (Fig. 12).

User can share his/her shopping list with a partner in order to save time (Fig. 13).

If the user wishes to shop from the partner's list, then he/she can provide that partner's (owner) Email ID (Fig. 14).

Fig. 9. Product list from the database
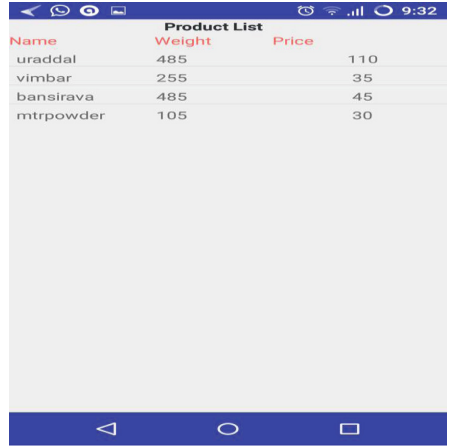
Fig. 10. Saved product list by the user
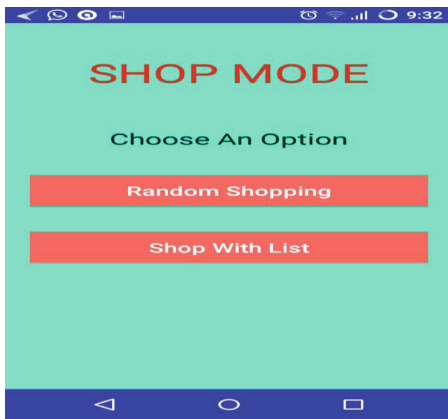
Fig. 11. Shop mode screen
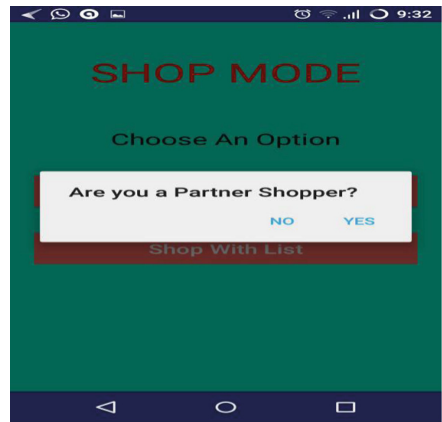
Fig. 12. Asking for partner shopper

If the user shares his/her list with a partner, then he has to provide partner's Email ID (Fig. 15).

The user can start shopping with his saved list sorted according to the bay. Scan button is used to scan the barcode. Done button is clicked to view the total amount (Fig. 16).
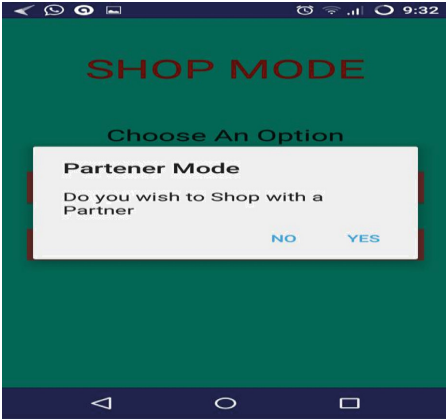
Fig. 13. Asking to shop as owner
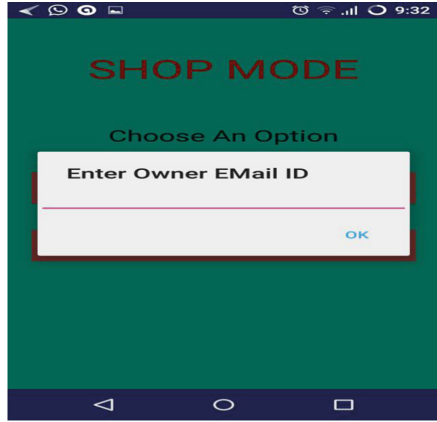
Fig. 14. Pop up asking Email ID
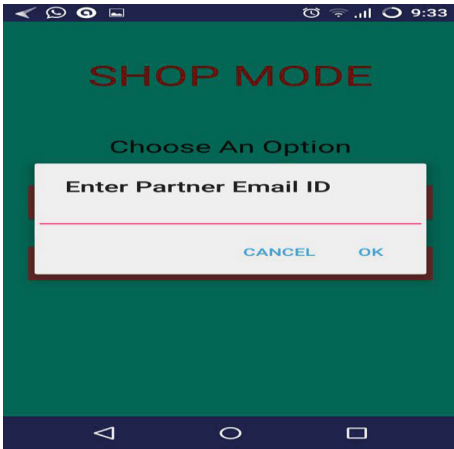
Fig. 15. Pop up asking Email ID of partner
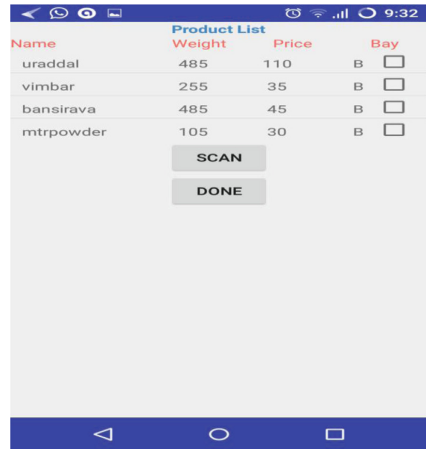
Fig. 16. Shopping with the saved list

7   Conclusion

The Internet of Things is one such technology that connects various objects in a network and is a milestone in the era of the smart world. The smart shopping cart features these technology enabling users to shop efficiently. Internet of things is the leading technology that makes the world experiences a seventh sense. By the year 2020, around 1 billion objects will be connected thus making the world smart. This smart shopping cart is implemented in such a way that it allows the customer to scan the item that he/she wants to purchase and automatically updates the bill thus preventing long queues at the checkout. Also, another interesting feature of this smart shopping cart is the cart-to-cart communication that helps the customers to shop parallel with friends and family.

References

1. Gubbi, J., Buyya, R., Marusic, S., Palaniswami, S.: Internet of Things (IoT): a vision, architectural elements, and future directions. IEEE (2011). https://doi.org/10.1109/i-smac.2017.8058399
2. Gangwal, U., Roy, S., Bapat, J.: Smart shopping cart for automated billing purpose using wireless sensor networks. IEEE (2013). https://doi.org/10.1109/icices.2014.703399
3. Yathisha, L., Abhishek, A., Harshith, R., Darshan Koundinya, S.R., Srinidhi, K.: Automation of shopping cart to ease queue in malls by using RFID (2015). https://doi.org/10.1109/icices.2014.7033996
4. Kaur, A., Garg, A., Verma, A., Bansal, A., Singh, A.: Arduino based smart cart. Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET) **2**(12) (2013)
5. Dash Robotic Shopping Cart. https://www.fastcompany.com/3061405/walmart-is-testing-a-robot-shopping-cart-so-you-can-do-the-job-of-low-wage-workers
6. Sanghi, K., Singh, R., Raman, N.: The Smart Cart – An Enhanced Shopping Experience. TA: Justine Fortier Team 41 (2012)
7. Dubey, V., Sangeeth Sagar, V.R., Sumalya, S., Abhilash, C.B.: An Android approach for wireless power harvesting from radio waves. In: Contemporary Computing and Informatics (IC3I), pp. 1235–1239. IEEE (2014). https://doi.org/10.1109/ic3i.2014.7019670

Pathnet: A Neuronal Model for Robotic Motion Planning

V. M. Aparanji[1](✉) ⓘ, Uday V. Wali[2] ⓘ, and R. Aparna[2] ⓘ

[1] Siddaganga Institute of Technology, Tumakuru, Karnataka, India
vmal508@gmail.com
[2] KLE Dr. MSS CET, Belagavi, Karnataka, India
udaywali@gmail.com, raparna27@gmail.com

Abstract. This paper proposes a new type of Multi-layered Artificial Neural Network (ANN) suitable for motion control of multi-joint robotic mechanisms with arbitrary Degrees of Freedom (DoF). Input layer classifies the incoming data using Auto Resonance Network (ARN) while higher levels implement Pathnet, a connection oriented neural network with Hebbian reinforcement learning capability. ARN networks grow with training input. Perturbation of ARN nodes allows the network to classify and recognize events with no previous history, Multilayer pathnets can recognize and recall temporal sequences. The network can memorize low cost paths and use parts of such segments in establishing new paths. We have used the system to control a multi segmented robotic system in R. Results of simulation presented in this paper encourage further explorations.

Keywords: Artificial Neural Networks (ANNs)
Auto Resonance Network (ARN) · Pathnet

1 Introduction

Moving a robotic end effector from current location to target location involves stepping through known nodes avoiding obstacles in the path. Such path is rarely a straight line. The problem is more complex in case of humanoid motion as the robot is mobile, possibly on an uneven terrain. Path of end effector would contain several steps and bends except in trivial cases. Inverse kinematic equations that yield torque and angles at joints can be solved iteratively using Newton-Euler force equations or Euler-Lagrange energy conservation equations [1] if the path is known apriori. Moving a convex volume attached to the end effector in constrained channels has been described as PSpace-Hard [2]. Inverse kinematic problems are categorized as structural analysis of mechanism, conventional approaches, intelligence or soft computing approaches and optimization based approaches [3]. Modern robotic systems with large DoF can benefit from use of soft control algorithms. For example Tai and Ming [4] have discussed the use of deep learning methods for robotic motion control. Veslin et al. [5] have used a genetic algorithm to estimate control parameters in a 7-DoF system.

Path planning for a joints system involves searching for a path between current location of the end effector and a desired location in the given 3-D work space. Path

planning has to precede actual motion of the end effector because it has to incrementally move towards the target possibly minimizing the energy loss, avoiding obstacles, etc. There may be other physical constraints to move the end effector in a particular way, possibly avoiding obstacles or working around a part of the work piece. Further, contents of the work space can change dynamically. Therefore, we may not be able to use a predefined path to move between two given points. Smooth motion of the joints system ensures efficient utilization of energy. Identification of sets of angles that can minimize energy utilization is also a complex problem in itself.

Any path traversed by the end effector can be seen as a series of angles to be set at various joints. Motion in a joints system occurs in a three dimensional space. With increasing number of DoF, inverse kinematic equations get harder to solve. As the number of steps increases, the number of possibilities grows exponentially. Therefore, a generic search in 3D space has a high complexity. Starting from the early works of Reif [2], which recognized the complexity of robotic motion, research has progressed a long way and is continuing. Zhou and Cook [6] in 1991 suggested incremental computation of sub-optimal path avoiding obstacles, for redundant joint systems. They used a deterministic approach based on geometric relations and Newton's equations for non-linear equations. Kawasaki et al. [7] used adaptive control for industrial robotic manipulators, based on a earlier work of Slotine and Li [8]. They used Newton-Euler equations for systems with both translational and rotational joints.

Many of the published works use a combination of neural and computational approach to the joints motion problem. Duka [9] has described use of a multi layer feed forward neural network for simulating a planar 3 segment joint. The network is trained with forward calculations and Levenberg-Marquardt training algorithm. During run time, the neural network will estimate the joint angles, which are then used to plan the trajectory. Faigl [10] has considered motion planning like a Travelling Salesman Problem (TSP) and used Self Organizing Maps (SOM) to generate motion schedules. By the turn of the century, feasibility of biped motion started attracting many researchers, especially in Japan and China. Huang, Yokoi et al. [11] have reported use of neural network for biped motion using a 18 DoF mechanism with trunk. They have analyzed both single and double support mechanism for stable walking using a smooth gait. While most of the robotic systems use a rotary actuator, pull relax mechanisms have also been suggested, e.g., biologically inspired Eccerobot by Potkonjak et al. [12]. Primary motion in such joints is caused by the pull tendon while the other tendon simply follows to avoid slackness.

Recently, He et al. have reported use of a machine learning technique for route planning [13]. This review paper covers many new machine learning techniques for robotic applications including Support Vector Machines (SVM), Auto regression models, Fuzzy Adaptive Resonance Theory (Fuzzy ART), Radial Basis Functions (RBF), Self Organizing Maps (SOM) etc. A low resolution Fuzzy decision system called Linguistic Decision Trees (LDT) is discussed in this paper. LDT classifies the actuation in to fuzzy classes like full-right, small-right middle right etc. These language descriptions along with mass assignments can be used to approximate the actuation forces. Convolutional neural networks have been used in vision related robotic control [4]. Zhang proposed a deep reinforcement learning algorithm based on successor feature [14]. Q learning algorithm is proposed for path planning that makes use of

reinforcement learning [15]. Deep learning methods for different applications are reported in [16]. It is possible to use several of the existing types of neural networks discussed in [13, 17] for path planning of robotic motion. We started with a belief that a network based on pull-relax model of biological systems is better suited for such applications. We have reported this type of network called Auto Resonance Network (ARN) in our earlier work [18]. A hierarchical network for path planning built using ARN has been reported in [19]. Pathnet is outcome of continued research in this area. PathNet nodes use a kind of Hebbian learning where two nodes that get activated in a time sequence will strengthen their connections. Activities on the path either enhance or weaken the network connections, adapting to the environment.

2   Auto Resonance Network (ARN)

In ARN each node is tuned to recognize a specific pattern of input vector. Each node will record the input at which it resonates and the corresponding outputs and other associated states. All nodes in a network will receive the same input but only the matching node will be driven to resonance. Resonance conditions will provide certain selectivity and noise tolerance [17]. An active node can be at resonance, near resonance or not in resonance. For a given input, one or none of the nodes are in resonance. If there is a resonating node and its output is above threshold value, we have a recognition. Otherwise, a new node is inserted into the network and is set to resonate at the values of present input. Under specific circumstances, such addition of new nodes may be suppressed.

The resonance equation is given by

$$y_k = \frac{4}{N} \sum (1 - w_{ki} x_i)(w_{ki} x_i), \quad i = 1 \ldots N, k = 1 \ldots K \tag{1}$$

where w_{ki} is the *resonant weight* for i-th input of k-th node as shown in Fig. 1(a). Note that w_{ki} and $(1 - w_{ki} x_i)$ reflect the pull-relax model we are expecting to implement. For an input limited range of $\{0-1\}$, this equation will produce maximum value of 1 when $x_i = 1/2w_{ki}, \forall i$

(a) Single input matching

(b) One node with N-inputs

(c) ARN Network K- patterns
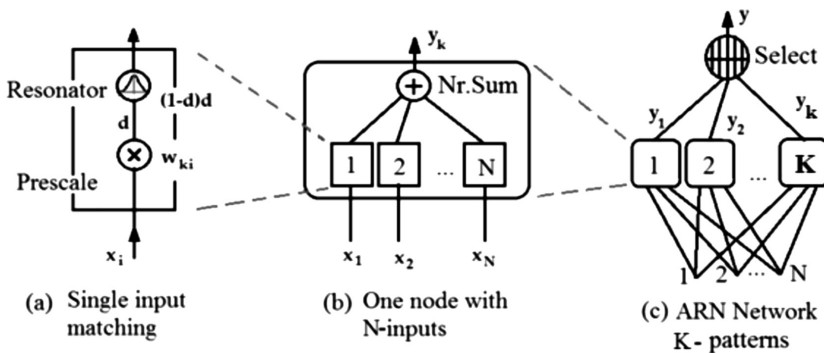
Fig. 1. Structure of ARN

Coverage of a node can be defined as the range of input values within which the node produces maximal output among all nodes and the output is above a set threshold. Therefore, coverage of k-th node can be expressed as

$$C_k = \{X | (y_k > \rho) \, and \, (y_k > y_i, \forall i \neq k)\} \tag{2}$$

where ρ is the threshold value. Coverage of a node is an import characteristic of the node. It provides noise tolerance by ensuring maximum output even in presence of noise. It also provides a condition to check if a new node has to be inserted into the network. Coverage of the node can vary dynamically.

We have used w_{ki} as a scaling constant that maps the input to ½ to make the node resonate. Input range for (1) is limited to {0–1}. While it is not too restrictive, it would be easier to have a function that takes care of input mapping. The quality of resonance, measured by the coverage of node, is not uniform across the input range. These limitations can be overcome by choosing complementary functions that can implement $(wx)(1 - wx)$ type of resonance, with wider input range and uniform coverage. One such excellent function is the sigmoid function given by $wx = 1/(1 + e^{-x})$ Eq. (1) can now be written as

$$y = 4 \left[\frac{1}{(1 + e^{-\sigma(x-M)})} \right] \left[\frac{1}{(1 + e^{\sigma(x-M)})} \right] \tag{3}$$

This function maps real valued input to 0–1 range and has a controllable resonance. Figure 2 shows the coverage, range and controllability of (3).
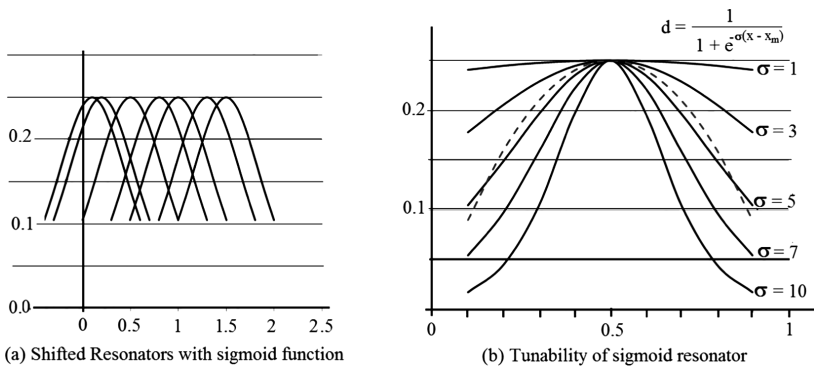
Fig. 2. Resonance curves for various x_r

Organization of a neural network based on (1) is shown in Fig. 1(b) and (c). We have indicated the resonator function as $d(1 - d)$ where $d = wx$ Each resonator corresponds to one input of one cc will be $K \times N$. Note that K changes dynamically but N is fixed.

It is easy to see that ARN like system could be possible in a biological system: biological joints work like a U-control, pulling and relaxing the segment to move the tip to a desired location. The original ARN model therefore was based on a very similar model as expressed in Eq. (1). We started with a minimal system and improved the equations to provide better control.

3   Pathnet

ARN can perform real-world input classification. However, it cannot identify temporal or spatial relations among the input set. One way of addressing this is to add associative data like labeling of nodes. However, it is more interesting to add a connection layer on top of ARN. This layer consists of cells that connect two or more cells with a causal or temporal order. Strength of the connection is akin to conductivity of a channel: Higher the conductivity, better is the connection.

Assume that during training, the robotic arm will move across the workspace on a random path. This will cause several nodes to recognize the position of the arm in a temporal series. This activity may be used to add or improve the connectivity between such nodes. Such network will essentially use Hebbian learning, reinforcing the connections with every learning opportunity. The spatial coordinates input to ARN will give a sense of location but the connection network will give direction to traversable paths. Using a connection network also allows us to see the ARN Connection network as implementation of a graph G(E,V), which allows us to convert many real life identification and classification problems into graph theoretic problems. Additional networks layers can be implemented on top of such a network to implement rich functionality. One such possible structure for joints motion control is given in [17]. An overview is shown in Fig. 3.
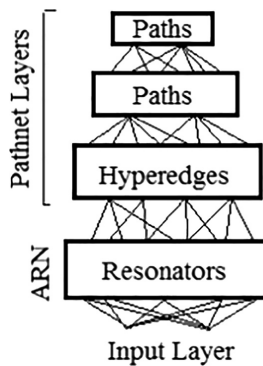
Fig. 3. ARN-pathnet hierarchical network

When an input is applied, ARN tries to classify the input among the existing nodes. If it cannot, a new node is added to the network. Therefore, the ARN contains nodes

that memorize previous input values. ARN nodes may be labeled to address disjoint or concave regions as belonging to the same class. Nodes that are activated in a temporal or spatial sequence are connected in a layer above ARN layer. Now, the path planning problem can be seen as a graph search problem and solved heuristically. Alternately, we could add layers of pathnet nodes above ARN layer that aggregate sequentially activated connections in lower layers. Therefore, a transit on the combined network can be seen as reflection of transit in the real work space. The advantage of ARN is that it provides a certain granularity in the otherwise contiguous work space. Adding granularity increases search efficiency. It is easy to observe that the nodes in pathnet behave like a Hebbian network: Connections between nodes on a path are strengthened and unused ones slowly degenerate. Stronger connection on a pathnet will have a lower cost and hence, the same path will have a priority. Degeneration is an important aspect of pathnet as it allows exploration of alternate paths: A path that exists but leads nowhere is not very useful. Direction of information flow in a Hebbian system is established when there is spatio-temporal sequence of firing of nodes. This is generally called the causality condition of Hebbian learning. The motion of the end effector from the current location to the target location provides such a sequence. For a joints control system, we need to store possible end locations Connections in a pathnet layer can store the coordinates and faster access later. A stored path is like a stored procedure for a joints system.

From the foregoing discussions, it appears plausible that a hierarchical network consisting of ARN and pathnet layers can implement a reinforcement learning system suitable for motion planning of a robotic system.

We call our model as *Pathnet*, because it is a network of several paths. Pathnet is best suited for middle layers of a multi-layer neural network because of following properties: It has no input processing capability like ARN. At the output, some decision layer has to be present which the pathnet can not assure. On the other hand, pathnet does some very important tasks of a neural network. It can memorize and identify sequences of events separated in space and time. It can implement Hebbian learning. Therefore pathnet can be used as a basic structure to solve problems which can benefit from it. Some of the areas we foresee are motion control, natural language processing, time series prediction, etc. Applications in these areas provide ample opportunity for research. We will discuss some details of pathnet in the following paragraphs.

Input to a node in pathnet is a set of nodes firing in a temporal order. These input nodes may be nodes of other structures like ARN or other pathnet nodes. These pathnet nodes model the synaptic connection between two nodes, called the pre-synaptic and post-synaptic nodes.

Assume that a cell has a temporal behavior indicated in Fig. 4. Action potential graphs of typical nerve cell are much sharper [20] and oscillatory than what we have represented here. Similar flattening of the curve, as in Fig. 3, has been reported [21] in optical nerves. Firing occurs only if the neuron accumulates charges beyond a threshold. Depolarization does not hold if the charge build-up is not enough. Dynamics of such nodes can lead to formation of stronger chord connection with repeated use. Let us consider the connection between three cells as shown in Fig. 5(a).

As N1 fires, N2 and N3 both receive charges at different rates as shown in Fig. 5(b). A weaker connection to N3 implies a slower build-up. However, when N2 receives
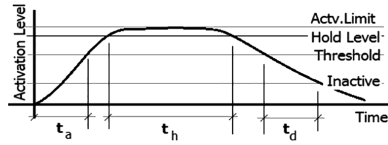
Fig. 4. Temporal behavior of a edge cell

(a) Three neurons

(b) Dynamics of three neurons emphasising N1-N3 connection
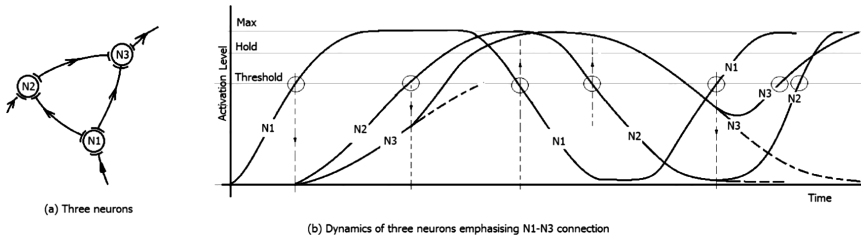
Fig. 5. Connections between neurons modeled by edge cells

enough charge, it fires and feeds N3. This will add to the incoming charge from N1 and starts accumulating faster. After N1 discharges N2 holds for some time and then starts discharging. It will take some time for N3 to start discharging. Pathnet connecting N1 to N3 now gets stronger because of the catalyst like action of N2, leading to formation of chord between N1 and N3, bypassing N2. Every time this occurs, N1–N3 bond grows stronger. Formation of chordal edge between nodes is important for optimization of pathnet. Delayed firing of neighbors gives rise to many interesting scenarios similar to one explained above. It is possible that a pathnet node or a set of nodes are activated by several paths. Such nodes or sets of nodes improve their charge accumulation rate by increasing the number of receptors on the post-synaptic neuron and number of charge gates on the pre-synaptic neuron.

4   Motion Planning with Pathnet

We have used the ARN and Pathnet neural networks in a hierarchical structure to address the path planning problem, as shown in Fig. 3. The system works independent of the number of joints in the mechanical system, as long as the individual joint angles and the displacement can be input to the network. Input to the network consists of location of the end effector. Outputs are the joint angles. The ARN layer recognizes the location of end effector and triggers nodes in series as the end effector moves in the work space. The output of ARN is tied to a pathnet layer, which learns the motion. The ARN layer spawns new nodes in the neighborhood of existing nodes to create approximations to unexplored areas in the work space. This reduces the training effort and increases the ability of the network to identify possible paths in untrained areas. First pathnet layer forms edges between immediate spatio-temporal neighbors, i.e., nodes which are close to each other or triggered one after the other form nodes in this

layer. Traverses of robots end effector are an aggregation on such nodes, again, triggered in sequence. Therefore, the second pathnet layer consists of possible paths. The third layer identifies preferred paths, which are used frequently. Obstacles are avoided by gating the input to such nodes or by negating the resonance condition.

Figure 6 shows results of a simulation using ARN-Pathnet hierarchical network. The simulation has been done using R language. Multiple paths between a given set of start and end points are shown in Fig. 6(a) for a 3-segment joint system. Each joint has 2DoF and support motion in constrained 3D space. Second layer of pathnet identifies multiple paths and a computational third layer evaluates the cost of each path and selects the one with lowest cost. Sweep of the robotic arm for a 5-segment robotic arm is shown in Fig. 6(b).

(a) 3-Segment Joint System showing paths in a work space   (b) 5-Segment Joint System, showing segments
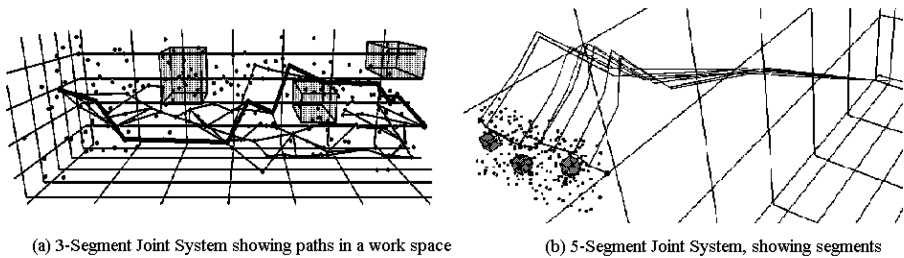
Fig. 6. Path planning for 3 and 5 segment Joint systems with 2DoF per joint.

5   Conclusions

A Pathnet is a connection oriented neural network that implements Hebbian reinforcement learning. Pathnet essentially implements dynamic path search and identification. Multi-layered pathnet structures can identify spatio-temporal events and hence can find applications in several areas of research. A single layer of pathnet is similar to a directed graph. Each layer of pathnetsupport one level of abstraction. We have implemented and used the ARN-pathnet combination to address path planning problem in robotic motion control. Efforts are on the way to use the structure in various areas of current research.

References

1. Saha, S.K.: Introduction to Robotics. McGrawHill, New Delhi (2008). ISBN 978-0070140011
2. Reif, J.H.: Complexity of the mover's problem and generalizations. In: Department of Computer Science, University of Rochester, Research report - TR58, pp. 421–427 (1979)
3. Jha, P.: Inverse kinematic analysis of robot manipulators. Dissertation, National Institute of Technology, Rourkela (2015)
4. Tai, L., Ming, L.: Deep-learning in mobile robotics - from perception to control systems: a survey on why and why not. Cornell University Archives. arXiv:1612.07139v3 (2017)

5. Veslin, E.Y., Dutra, M.S., Lengerke, O., Carreno, E.A., Tavera, M.J.M.: A hybrid solution for the inverse kinematic on a seven DOF robotic manipulator. IEEE Lat. Am. Trans. **12**(2), 212–218 (2014)

6. Zhou, L., Cook, G.: Path planning for robotic manipulators with redundant degrees of freedom. IEEE Trans. Ind. Electron. **38**(6), 413–420 (1991)

7. Kawasaki, H., Bito, T., Kanzaki, K.: An efficient algorithm for the model based adaptive control of robotic manipulators. IEEE Trans. Robot. Autom. **12**(3), 496–501 (1996)

8. Slotine, J.J.E., Li, W.: On adaptive control of robot manipulators. Int. J. Robot. Res. **6**(3), 50–59 (1987)

9. Duka, A.: Neural network based inverse kinematics solution for trajectory tracking of a robotic arm. Procedia Technol. **12**, 20–27 (2014)

10. Faigl, J.: An application of self-organizing map for multirobot multigoal path planning with minmax objective. Comput. Intell. Neurosci. **2016**, Article ID 2720630 (2016). https://doi.org/10.1155/2016/2720630

11. Martin, A.E., Gregg, R.D.: Incorporating human-like walking variability in HZD-based bipedal model. IEEE Trans. Robot. **32**(4), 943–949 (2016)

12. Potkonjak, V., Svetozarevic, B., Jovanovic, K., Holland, O.: The puller-follower control of compliant and noncompliant antagonistic tendon drives in robotic systems. Int. J. Adv. Robot. Syst. **8**(5), 69 (2011)

13. He, H., McGinnity, T.M., Coleman, S., Gardiner, B.: Linguistic decision making for robot route learning. IEEE Trans. Neural Netw. Learn. Syst. **25**(1), 203–215 (2014)

14. Zhang, J., Springenberg, J.T., Boedecker, J., Burgard, W.: Deep reinforcement learning with successor features for navigation across similar environments. arXiv:1612.05533 (2017)

15. Zhang, Q., Li, M., Wang, X., Zhang, Y.: Reinforcement learning in robot path optimization. J. Softw. **7**(3), 657–662 (2012)

16. Carrio, A., Sampedro, C., Rodriguez-Ramos, A., Campoy, P.: A review of deep learning methods and applications for unmanned aerial vehicles. J. Sens. **2017**(2), 1–13 (2017). https://doi.org/10.1155/2017/3296874. Article ID 3296874

17. Aparanji, V.M., Wali, U.V., Aparna, R.: Robotic motion control using machine learning techniques. In: International Conference on 6th IEEE Communication and Signal Processing (ICCSP) Melmaravattur, April 2017. IEEE Xplore (2017, in press)

18. Aparanji, V.M., Wali. U.V., Aparna, R.: A novel neural network structure for motion control in joints. In: ICEECCOT, Mysore, pp 227–232 (2016). Also available from IEEE Xplore

19. Aparanji, V.M., Wali, U.V., Aparna, R.: Automated path search and optimization of robotic motion using hybrid ART-SOM neural networks. In: International Conference on Recent Advancement in Computer and Communication, Bhopal, May 2017, ICRAC-2017. LNNS. Springer, Heidelberg (2017, in press)

20. Stufflebeam, R.: Neurons, Synapses, Action Potentials, and Neurotransmission, Consortium on Congnitive Science Instruction (2008). http://www.mind.ilstu.edu/curriculum/neurons_intro/neurons_intro.php

21. Verbny, Y., Zhang, C.L., Chiu, S.Y.: Coupling of calcium homeostasis to axonal sodium in axons of mouse optic nerve. J. Neurophysiol. **88**(2), 802–816 (2002)

Impact of Named Entity Recognition on Kannada Documents Classification

R. Jayashree[✉] [iD], Basavaraj S. Anami, and S. Teju

PES Institute of Techology, Bangalore, India
jayashree@pes.edu

Abstract. Natural language processing in Kannada language is promising research field due to the unavailability of tools and challenges in various aspects such as lack of annotated Kannada corpus. The important aim objective of this paper is to study and understand the impact of named entity recognition (NER) on Kannada text documents classification. Rule based Kannada named entity recognition system is implemented and integrated with Naïve Bayes classifier using a tool for this purpose. Rule based approach is considered for the purpose of experimentation. Another important aspect of this work is the attempt made to improving the classifier performance for Kannada Documents through NER. Comprehensive study is conducted to investigate the impact of Kannada named entity recognition on Kannada text document classification using Naïve Bayes classifier. Experimental results shows classification algorithm produces better results for Kannada text documents with previously recognized named entities.

Keywords: Natural language processing · Kannada named entity recognition
Rule based approach · Text classification

1 Introduction

Natural Language Processing is a field of computer science, which is intended to create interactions between human and machine. It is widely being used in variety of applications such as emails, news articles, web portals, social media, to name a few. Since many decades, natural language processing has been a part of our life in many aspects such as spell check, detecting mails as spam or not spam, language translation, grammar correction and question answering system etc.

Information Retrieval (IR) is an application of NLP and Named Entity Recognition is another subarea of information retrieval. Named Entity Recognition is high end application of Natural language processing which includes identifying proper names in text and classifying these identified proper names into previously specified categories such as nouns, names of organizations, names of location, number expressions (date and time).

Kannada, is predominantly being spoken in the state of Karnataka, India. For the language Kannada, preprocessing tools such as annotated corpora, name dictionaries, good morphological analyzers, Parts of Speech (POS) taggers etc. are not yet available

in the required measure and not much work has been done in NLP, with respect to the Kannada language.

Here, we have made an attempt to develop rule based system for Kannada NER. A rule based systems needs more grammatical and linguistic analysis to make rules. The Rule Based approaches may give good result with sufficient gazetteers lists, language dependent features and rules for purely particular language. Named entities are open class words, every day new words added to languages and gazetteers list is infinite to store all words is not possible, hence gazetteers are needed to divide into finite tests like suffix, prefix, context words etc. All rule based approaches are language dependent.

2   Literature Survey

Approaches to named entities identification in south Asian languages and study [1] is the work which highlights the various approaches for finding Names Entities (NE) in Indian Languages, challenges faced for Indian languages and results for identifying named entities in text document using these approaches.

Rule based method to find named entities in Kannada [2] proposed a rule based approach to recognize named entities (Names of person, location, time, organization, number and measurement) in Kannada language. Manually collected/created suffix, prefix list and proper noun list of 5000 words are used. English have NER features such as capitalization which helps in finding named entities. Lack of capitalization feature in Kannada makes NER task more complex than in English. The proposed system gives result 86% of precision and 90% of recall over 20 files.

Kannada named entity recognition and classification [3] provides comprehensive study in implementing NLP Models based on Noun Taggers. Authors proposed and implemented HMM (Hidden Markov Model), supervised learning techniques. Unannotated Kannada text file are inputted to proposed system which recognizes the Named Entities and annotated text document file is obtained as output. Suitable cryptographic algorithm is applied on output of classification system in order to secure corpus.

Rule based approach Named entity recognition for Malay language [4] gives detailed description of proposed system for recognizing named entities in Malay language which contains rules for identifying person entities, location rules, organization rules etc. This algorithm implements both rule based method and machine learning methods to obtain better results. The performance of the system can be made better by updating dictionaries and lists.

The detailed challenges for recognizing named entities in Indian languages are discussed in rule based approach to identify named entities in Urdu language [5]. The algorithm for identifying named entities in Urdu using rule based method is proposed. Concept of n-grams are used in this proposed system and get promising results. 6-g are used to implement this model and authority files are created which consist of list of common person names, location names, organization names, person name prefixes, suffixes etc. considering 3-g would be ideal for any named entity recognition system.

For recognizing named entities in Telugu [6], authors developed a CRF approach with rule based approach to identify named entities in Telugu. Not much work has been

done in Telugu language because of lack of annotated corpus they manually created 13,425 words with are manually tagged as noun or not-noun and also extracted few features which helps in identifying nouns in a given Telugu text. The work indicates that, the performance of the system can be increased by increasing the gazetteer list and better accuracy can be achieved by hybrid approach than any other approach.

A brief review for basic approaches to named entity recognition is discussed in named entity recognition methods [7]. The authors tried to improve precision and portability for the developed system as the most and difficult problem in NER is portability. Rule based approach improves the precision by adding more rules but it will automatically decreased the portability because of its fix rules. In order to overcome this problem this paper proposed fuzzy NER and produced better results.

An algorithm is proposed to classify Punjabi Documents text documents [8] by creating domain based ontology in Punjabi that consist of domain related terms. The main advantage of classification with help on ontology is that it does not need training dataset.

The sentence level text classification [9], is a work, which highlights the importance of sentence level text classification in the Kannada language. The performance of Naive Bayesian classifier is looked at in this work. The exhaustive literature survey made highlights the importance of Named Entities and their impact on text classification. No considerable work has been recorded in the case of NER in the Kannada language. This motivated us to take up the task of NER and its impact on text classification.

3   Proposed System

There are mainly two parts in this paper, first identification of named entities in the given text document and secondly, classification of the text document in Kannada Language. For named entity recognition, rule based approach is used, because it is expected to produce better results for Indian languages (Fig. 1).
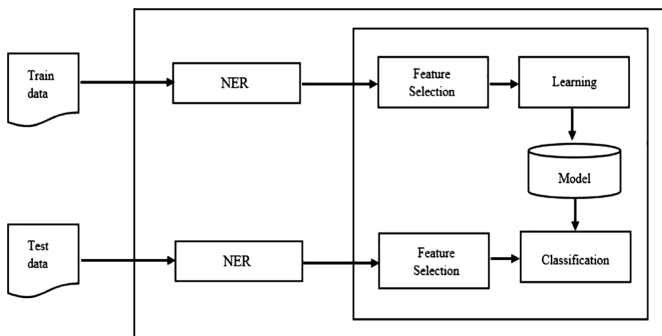
Fig. 1. Architecture of proposed system

3.1   Named Entity Recognition

Rules are handcrafted regular expressions which are designed by language dependent features and context features. Authority file is created which consist of names of people, organization names and location names. Another authority file is created which consist of list of common person name prefixes, suffixes, location suffixes, organization suffixes. 3-g are considered to get context features and identify named entities.

3.1.1   Person names: This file consist of 500 person names such as Kuvempu, Chandrasheker Kambar, Goruru Ramaswamy Iyenger etc.

3.1.2   Person prefix: This file consist of 51 person name prefixes such as Pradhani, Sri, Pro, Rastrapathi etc.

3.1.3   Person suffix: This file consist of 60 person name suffixes such as Singh, Kapoor, Gowda etc.

3.1.4   Location names: This file consist of 500 names of locations such as Karnataka, Andra Pradesh, America etc.

3.1.5   Location suffix: This file consist of 15 location name suffixes such as Uru, Halli etc.

3.1.6   Organization names: This file consist of 200 names of organizations such as SCC, TATA Steel, Google, Yahoo etc.

3.1.7   Organization suffix: This file consist of 34 organization name suffixes such as limited, Nigama, Samsthe etc.

Table 1 represents the description of tags used in this system with examples.

Table 1. Named entity tags description

Named Entities	Tag	Description	Example
Person	PER	Person name	ಕುವೆಂಪು, ಕರುಣಾನಿಧಿ
Person prefix	PER	Title before the person name	ಶ್ರೀ, ಪ್ರಧಾನಿ
Person suffix	PER	Word after/within the person name	ಸಿಂಗ್, ಗೌಡ, ವ್ಯ, ರಾವ್
Location	LOC	Location name	ಕೊಲಂಬಿಯಾ, ಕರ್ನಾಟಕ
Location suffix	LOC	Word after/within the name of location	ಪುರ, ಪ್ರದೇಶ
Organization	ORG	Organization name	ಬಿಜೆಪಿ, ಗೂಗಲ್
Organization suffix	ORG	Word after name of organization	ಲಿಮಿಟೆಡ್, ಇಂಡಸ್ಟ್ರೀಸ್

Figure 2, shown below gives the detailed explanation for the module to identify Kannada named entities. Input given to this module are Kannada text documents, a sample document is shown in Fig. 3. The output obtained after execution are Kannada text documents with named entities identified with appropriate tags, which is shown in a sample document in Fig. 4. The given Kannada text document is divided into sequence of sentences, again each sentence is divided into sequence of tokens. The program takes each token and match it with the dictionary available. In this work, dictionary means created files with list of person names, location names and

organization names. If token finds direct match with the dictionary created, then that token is output as named entity with appropriate tags. For token which does not find match in the provided dictionary, NER features are checked upon that token. NER features such as person name prefixes, person name suffixes, organization suffixes, location suffixes are used here. The previous token and the next token to the current token are considered at this part to obtain the named entities. If any of the NER features are matched, then the token is recognized as named entity and attached with appropriate tag. To achieve the named entities attached with tags out of plain Kannada text rule based approach is utilized. Each and every step in this module is depicted as rule.
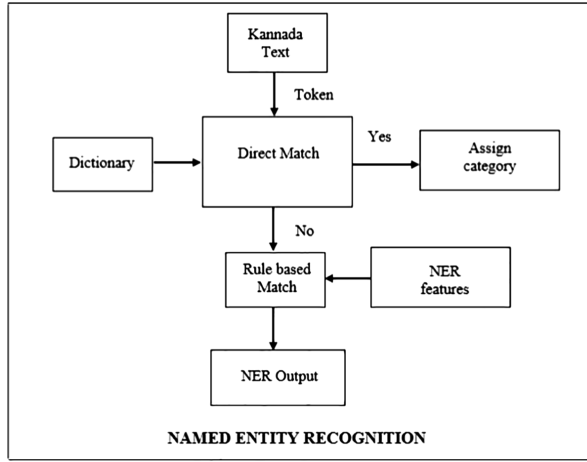
Fig. 2. Kannada named entity recognition module

Fig. 3. Sample input Kannada text

೧೯೪೪ರಲ್ಲಿ\NUM ಗುತ್ತಿಗೆಯನ್ನು ತಾನು ಪಡೆದುಕೊಂಡಿರುವುದಾಗಿ ವಿಪ್ರೋ\ORG ತಿಳಿಸಿದೆ . ಅಂಕಿ ಅಂಶಗಳ ಪ್ರಕಾರ ಏಪ್ರಿಲ್\TIME ೨೦೦೭ರ\NUM ಅವಧಿಯಲ್ಲಿ ಇಸಿಬಿಗಳ ಮೂಲಕ ಸುಮಾರು ಬಿಲಿಯನ್ ಡಾಲರ್ ಗಳನ್ನು ಏರಿಸಿದೆ . ಸೆಲೆಬ್ರಿಟಿ ಎಂಬ ಪದವಿ ತನಗೆ ಇಷ್ಟವಿಲ್ಲ ಎಂದು ಕುವೆಂಪು\PER ಅಭಿಪ್ರಾಯಪಟ್ಟಿದ್ದಾರೆ . ಸೆಲೆಬ್ರಿಟಿ ಎಂಬ ಪದವಿ ತನಗೆ ಇಷ್ಟವಿಲ್ಲ ಎಂದು ಚಂದ್ರಶೇಖರ ಕಂಬಾರ\PER ಅಭಿಪ್ರಾಯಪಟ್ಟಿದ್ದಾರೆ . ಸೆಲೆಬ್ರಿಟಿ ಎಂಬ ಪದವಿ ತನಗೆ ಇಷ್ಟವಿಲ್ಲ ಎಂದು ಗೊರೂರು ರಾಮಸ್ವಾಮಿ ಅಯ್ಯಂಗಾರ್\PER ಅಭಿಪ್ರಾಯಪಟ್ಟಿದ್ದಾರೆ . ಭಾರತದ\LOC ಪ್ರಧಾನಿ ಮನಮೋಹನ್ ಸಿಂಗ್\PER ಹಣಕಾಸು ಸಚಿವ ಪಿ ಚಿದಂಬರಂ\PER ಭಾರತವನ್ನು ಉತ್ತಮ ಪಥದಲ್ಲಿ ಮುನ್ನಡೆಸುತ್ತಾರೆ. ಬೆಂಗಳೂರು\LOC ಉತ್ತರ ಜಿಲ್ಲೆಗಳಲ್ಲಿರುವ ದೇವಸ್ಥಾನಗಳು ಹೊರನಾಡಿಗರು ವಿದೇಶಿಪ್ರವಾಸಿಗರನ್ನು ಹೆಚ್ಚು ಆಕರ್ಷಿಸುತ್ತವೆ ಕಳೆದ ವರ್ಷದ ಅವಧಿಯಲ್ಲಿ ಭಾರತ್ ಸಂಚಾರ್ ನಿಗಮ್ ಲಿಮಿಟೆಡ್\ORG ರೇವಣಿ ಬಡ್ಡಿ ದರ ಶೇ೭\NUM ರಷ್ಟಾಗಿತ್ತು . ಕೃಷ್ಣಶಾಸ್ತ್ರಿ\PER . ಮನಮೋಹನ್ ಸಿಂಗ್\PER . ಸಚಿವ ಚಿದಂಬರಂ\PER . ಪ್ರಧಾನಿ ಮನಮೋಹನ್ ಕಂಬಾರ ಸಿಂಗ್\PER .

Fig. 4. Sample output Kannada text

3.2 Text Classification

See Fig 5.
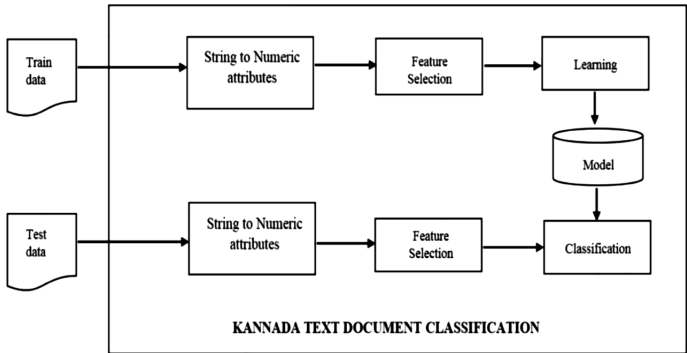
Fig. 5. Kannada text documents classification module for classifying Kannada text documents

4 Tests and Results

Train dataset considered for this system includes 155 Kannada text documents under 3 categories: Entertainment, sports and Politics.

It is observed that, a slight increase in precision, recall and F-measure of classifier results are recorded in Tables 2 and 3. Classifier results show increased accuracy with named entities. Noticeable decrease in mean absolute error and root mean squared error is also recorded. This is indicative of the fact that, classifying documents after recognizing named entities increase the accuracy of the classifier.

Table 4 precisely indicates the impact of named entities on Kannada documents with increasing precision, recall and F-measure.

Table 2. Classification results of Kannada text documents (without using named entities)

No. of Kannada documents	Precision	Recall	f-measure
10	0.75	0.7	0.711
20	0.846	0.75	0.711
33	0.795	0.727	0.714

Table 3. Classification results of named entity recognized Kannada text documents (using named entities)

No. of Kannada documents	Precision	Recall	f-measure
10	0.75	0.7	0.711
20	0.867	0.8	0.786
33	0.808	0.758	0.752

Table 4. Comparison of Precision, Recall and F-score with and without named entities.

Results	Kannada text documents	Kannada text documents with named entities identified
Precision	0.795	0.808
Recall	0.727	0.758
F-measure	0.714	0.752

5 Conclusion

Named entity recognition in Kannada language is promising research field due to lack of work in the language and challenges in various aspects such as lack of annotated Kannada corpus.

In this work, Rule based named entity recognition system is successfully implemented with manually created dictionaries (list of proper names). NER system is integrated with machine learning and data mining tool (WEKA). Using the collected corpus and the proposed system, a comprehensive study is conducted to investigate the impact of Kannada named entity recognition on Kannada text document classification using Naïve Bayes classifier. Named entities have direct impact on document classification and Naïve Bayesian classifier gives better classification accuracy.

6 Future Work

Future enhancement to this work includes adding more proper names to the dictionary, identifying more Kannada NER features such as person name suffixes, person name prefixes, location suffixes, organization suffixes etc., identifying more context features and creating rules for corresponding features.

References

1. Hiremath, P., Shambhavi, B.R.: Approaches to named entity recognition in Indian languages: a study. Int. J. Eng. Adv. Technol. (IJEAT) **3**(6), 191–194 (2014)
2. Bhuvaneshwari, C.M.: Rule based methodology for recognition of Kannada named entities. IJLTET **3**, 50–59 (2014)
3. Amarappa, S., Sathyanarayana, S.V.: Named entity recognition and classification in Kannada language. Int. J. Electron. Comput. Sci. Eng. Trans. Mach. Learn. Artif. Intell. **2**, 281–289 (2012)
4. Alfred, R., Leong, L.C., On, C.K., Anthony, P.: Malay named entity recognition based on rule-based approach. Int. J. Mach. Learn. Comput. **4**(3), 300–306 (2014)
5. Riaz, K.: Rule-based named entity recognition in Urdu. In: Proceedings of the Named Entities Workshop, pp. 126–135 (2010)
6. Srikanth, P., Murthy, K.N.: Named entity recognition for Telugu. In: Proceedings of the IJCNLP-2008 Workshop on NER for South and South East Asian Languages Hyderabad (2008)
7. Mansouri, A., Affendey, L.S., Mamat, A.: Named entity recognition approaches. IJCSNS Int. J. Comput. Sci. Netw. Secur. **8**(2), 339–344 (2008)
8. Kaur, K., Gupta, V.: Named entity recognition for Punjabi language. Int. J. Comput. Sci. Inf. Technol. Secur. (IJCSITS) **2**(3) (2012)
9. Jayashree, R., Srikanta, M.K., Anami, B.S.: An analysis of sentence level text classification in the Kannada language. In: International Conference of Soft Computing and Pattern Recognition (SoCPaR), pp. 147–151. IEEE (2011)

Big Data Analysis - An Approach to Improve Power System Data Analysis and Load Research

Sandhya S. Shankarlinga[1]([✉]) [iD], K. T. Veeramanju[2] [iD],
and R. Nagaraja[3] [iD]

[1] Karnataka Power Transmission Corporation Limited, Bengaluru, India
sandhyakrish2002@gmail.com
[2] Sri Jayachamarajendra College of Engineering, Mysuru, India
veeramanju.kalyan@gmail.com
[3] Power Research & Development Consultants Pvt. Ltd., Bengaluru, India
nagaraja@prdcinfotech.com

Abstract. The introduction of various intelligent electronic devices (IEDs), sensors and other network controls for smarter operation of the electric grid has resulted in massive data explosion. With an exponential growth in volume and diversity of data sources, developing an effective data management system is challenging and also imperative. This paper explores the current state of data analysis and associated problems in power sector with reference to Indian scenario and narrates how Big Data Analysis and statistics analysis tools could be adapted to improve power system data analysis and load research by acting on the deluge of big data and leveraging various statistical algorithms. We leverage Apache Hadoop BigData ecosystem for large volume of load research data and 'R' for pattern recognition and load forecasting. The paper also suggests an approachable roadmap to the power utility for sub-station data analysis embracing identification of peaks and valleys in sub-station demand, detection of anomalies in the data and also conducting short term load forecasting of the sub-station peak demand using these technologies.

Keywords: Big data · Power system data analysis · Load research
Short term load forecasting · Seasonal ARIMA

1 Introduction

The electric grid has evolved over the past century from a small independent community based systems to one of the largest and most complex systems today. There has been drastic change in the power industry in terms of its professional as well as operational models in order to plug the gap between demand and supply. The resulting transformation and ongoing rapid growth of computational technology are generating tremendous volumes of data each day and is sure to become even more gigantic in the near future. This compounding increase in the **volume, velocity** and **variety of data** is referred to as ***big data***. Big Data is a term used to data sets whose size is beyond the capacity of commonly used software tools to capture, manage and process within a allowable elapsed time.

It is increasingly becoming imperative for the power utilities to mine this data or deploy big data analytics and statistical analysis tools in order to enhance the operational efficiency. Big Data analytics can pave way for better analysis of power system data and power system load research.

2   Current State of Data Analysis and Problems

The various power system parameters like voltage, current, active power, reactive power, active energy import and active energy export pertaining to a sub-station transformers and connected feeders are pumped into the server at the central control room through Remote Terminal Units at an interval of every one minute. This tremendous volume of data termed as 'Big Data' has posed various problems; to list a few,

1. **Difficulty in Analysing**
   Massive volumes of data are difficult to analyze with standard database software.
2. **Time Consuming**
   It takes lot of time to process and produce analytics information, as data is voluminous.
3. **No comprehensive approach**
   The variety of information like active/reactive power, voltage current power factor etc. are collected but there are no comprehensive approaches for centralizing the information.
4. **Delay in furnishing the information**
   There has been delay in furnishing the information due to huge volume of data which would result in taking right decisions at right time.
5. **Probability of furnishing erroneous information**
   As the data is voluminous, the manual analysis may end up in erroneous results.

3   Handling Big Data

Traditional data management solutions or tradition Relational Data base management systems (RDBMS) cannot achieve the required performance and also are not efficient for storing Big data due to the volume and data variety or data structure. As such, 'Hadoop' which is an open source software framework is employed for handling Big Data.

Distributed computing and processing frameworks are getting built from multiple vendors. On the same lines Hadoop is an Apache open source project in this area of distributed computing. Hadoop is based on MapReduce concept and provides parallel storage and processing. MapReduce is modeled with 2 phases, namely Map; which maps the data which is data extraction phase and second phase is for transforming the data which is called reduce phase. Foundation of Hadoop is based on HDFS which provided distributed file storage capability enabling spreading across the huge amount of data across servers.

For enabling MapReduce operations, Hadoop provides 3 types of techniques.

1. Java based MapReduce program which is targeted for structured data, semi structured data and unstructured data.
2. Pig is another way, which is scripting based (similar to SQL scripting) and is primarily targeted for semi structured data.
3. Third method is Hive Query Language (also called as HQL) which is another options for processing the structured data.

4   Methodology

The electrical Sub-station data is processed and summarized using Hadoop and then processed data is used in statistical analysis platform 'R' to arrive with the following insights:

1. Identification of Peaks and Valleys in Sub-station Demand
2. Detection of anomalies in the data which would indicate an opportunity to make actionable decisions
3. Short term load prediction for sub-station demand.

In order to achieve the objective, following methodology has been adopted as depicted in Fig. 1.
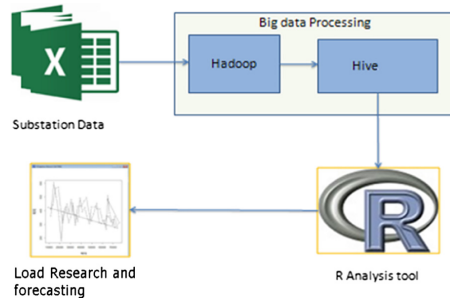
Fig. 1. Diagram depicting methodology

4.1   Enormity of Data

In a micro level, one 220 kV station and connected eight 66 kV sub-stations data are collected for a period of 1 year (Nov 2013 to Oct 2014). The various power system parameters like voltage, current, active power, reactive power, active energy import and active energy export pertaining to transformers and connected feeders of nine substations are collected. Each sub-station contain about 16–32 feeders and in overall, data is collected from 240 number of feeders. The total volume of data thus collected amounts to 33.24 GB and selected subset of data is considered for illustration purpose.

Figure 2 shows the data inflow and outflow to achieve the objectives cited and entire process is explained in below steps:
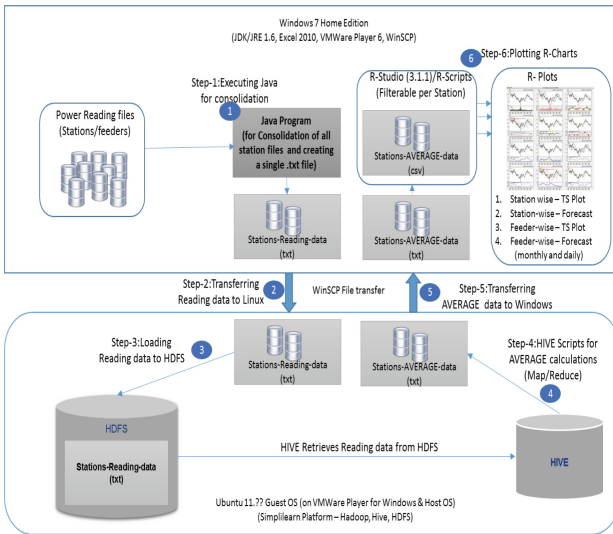
Fig. 2. Detailed architectural diagram

Step 1: A Java program reads all the files from configured Station data folder and creates a consolidated '.txt' file.

Step 2: Created consolidated '.txt' file is transferred to Linux using WinSCP utility.

Step 3: Consolidated '.txt' file (Stations Reading consolidated data) is loaded to HDFS using HDFS commands.

Step 4: AVERAGE Power calculations for each day for each stations is calculated and fetched using HIVE queries using HIVE commands.

Step 5: Created AVERAGE calculations file '.txt' file is transferred to Windows using WinSCP utility.

Step 6: Various R-Charts plotted using R-Studio and R-Scripts.

Station-wise – TS Plot

Station-wise – Forecast

Station-wise – Anomaly Detection

Feeder-wise – TS Plot

'R' is a **statistical tool/programming language** which is free and open source is proposed to be used for data analysis and forecasting. It is a comprehensive statistical platform, offering all manner of data analytic techniques. It can easily import data from a wide variety of sources.

5   Results

5.1   Identification of Peaks and Valleys of Sub-station Peak Load Time Series Using R

The total station peak load of 220 kV Station for a period of one year from November 2013 to October 2014 is plotted using R. The plot is shown in Fig. 3:
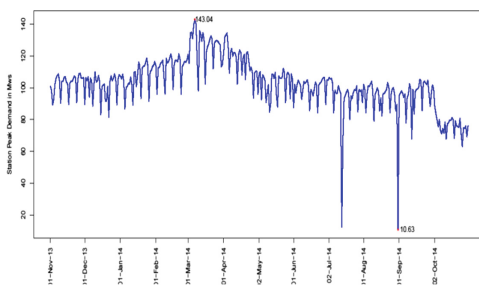
Fig. 3. Graph showing variation of station demand over a period of one year

Above graph is plotted starting from November 2013 till October 2014. It is seen from the above plot that the peak load of the station is highest during March-2014. This is due to summer where power consumption is high due to cooling appliances and various industrial and home appliances as well. Table 1 shows the percentage increase in station peak load, month on month.

Table 1. Percentage increase in station peak load

Sl. no.	Month	Highest peak demand in MW	Average peak demand in MWs	% increase in the peak demand compared to previous month
1	Nov 2013	109.09	102.66	
2	Dec 2013	109.89	101.12	−1.52
3	Jan 2014	118.53	106.84	5.35
4	Feb 2014	121.41	113.46	5.84
5	March 2014	143.04	126.68	10.43
6	April 2014	134.40	117.63	−7.69
7	May 2014	112.23	102.78	−14.44
8	June 2014	106.47	99.70	−3.09
9	July 2014	106.51	92.92	−7.31
10	Aug 2014	103.90	92.36	−0.60
11	Sep 2014	105.39	96.20	3.99
12	Oct 2014	87.81	75.93	−26.69

After the month of April, May, it is observed that the trend is slight downward till the month of Sep, October.

The decomposition of the time series into trend, seasonal and random components are plotted using 'R' statistical platform as shown in Fig. 4:
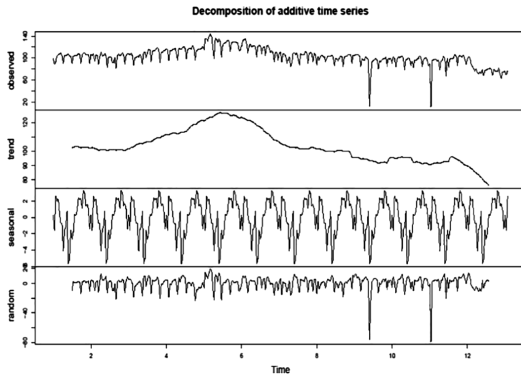
Fig. 4. Decomposition of station demand time series

Weekly Analysis

The plot of variation in the peak load in 4 different weeks for the months of November 2013 and December-2013 are shown in Figs. 5(a) and (b) respectively. (Week is taken from Monday to Sunday). Figures 6(a), (b), (c) and (d) shows the variation of the station demand during each week in a month.
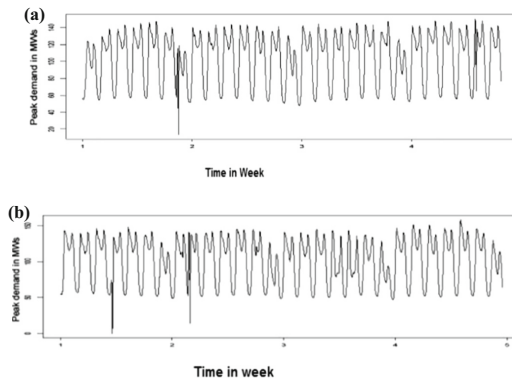
Fig. 5. (a) Weekly station demand variation in November 2013, (b) weekly station demand variation in December 2013

Daily Analysis

The plot of variation in the peak load for 1.11.13 & 2.11.13 is shown Fig. 7.
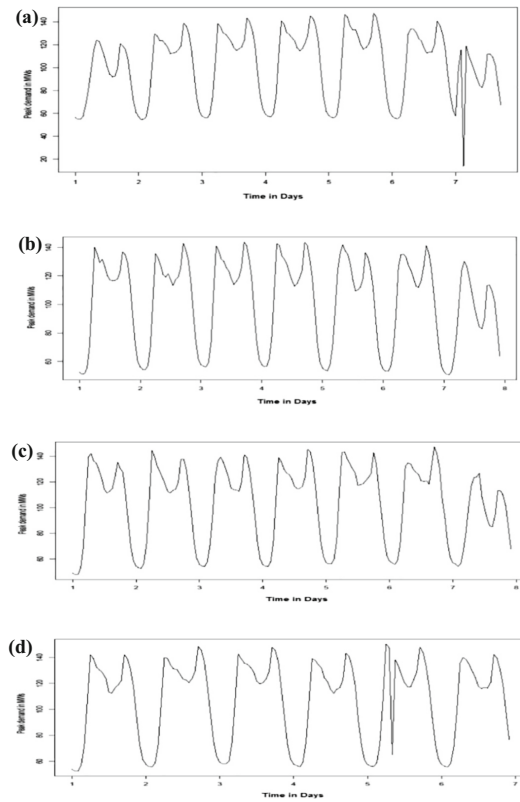
Fig. 6. (a) Variation of station demand during 1st week of November 2013: (4.11.13 to 10.11.13), (b) variation of station demand during 2nd week of November 2013: (11.11.13 to 17.11.13), (c) variation of station demand during 3rd week of November 2013: (18.11.13 to 24.11.13), (d) variation of station demand during 4th Week of November 2013: (25.11.13 to 30.11.14)
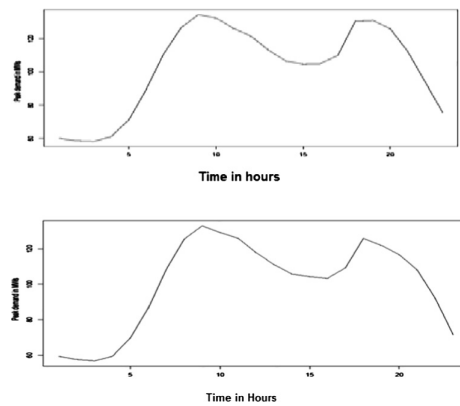
Fig. 7. Variation of station demand during each hour in a day

5.2 Anomaly Detection by 6 Sigma Method

Anomaly detection which is also called as outlier detection is the way of identifying obser-
vations, events or data items which do not confirm expected patterns or boundaries of the other
items. These are also called as noise, novelties, outliers, exceptions and deviations etc.

'Three-Sigma Limits' is a statistical calculation that refers to data within three
standard deviations from a mean. It is used as a measure of how much some series of
events vary around the average. 3-sigma limits are used to set the upper and lower
control limits (3 limits in upper side and 3 to the lower side). Figure 8 shows Detection
of anomaly in station demand using 6 sigma method.

Standard deviation is measurement of variability in a statistical way which shows
amount of variation from the statistical average values. In standard deviation, low
values are indication of data points falling close to average values. High values are
indication of data points are not close to average and are wide spread.
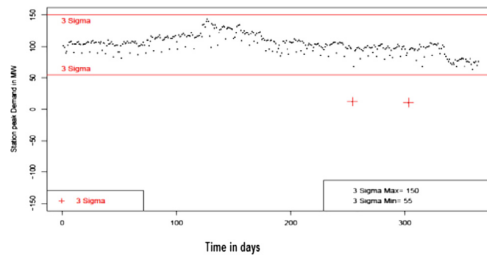
Fig. 8. Detection of anomaly in station demand using 3 sigma method

5.3 Short Term Load Forecasting

In electrical utility, to provide quality and uninterrupted power supply to consumers,
load forecasting is very much essential. Load forecasting is a method of analyzing and
predicting load demands for the future needs. It is important to manage electricity
demands in a scientific and efficient way.

Day Ahead Forecasting Using Seasonal ARIMA Model
ARIMA is an abbreviation which is elaborated as Autoregressive Integrated Moving
Average. This is a statistical model for forecasting. There are multiple methods in
ARIMA model. One of these methods is called Univariate or 'single vector'. Uni-
variate ARIMA is a technique and helps in forecasting the future values of data set
based on its own inertia of a data set. Same is also called as Box-Jenkins ARIMA is
much suitable for long series of data and if the past data provides much stable cor-
relation. On the other hand, if the data series is much shorter and if the correlation
throws much volatility, we can opt for seasonal ARIMA method. Seasonal ARIMA is
abbreviated as SARIMA. Below are the steps involved in the load forecasting.

Step 1: Plotting and Decomposing the Time Series

In 'hourly station demand data' considered for study, it is observed that time series plot is not stationary in mean, as the level changes a lot over time. As periodic components influence the ARIMA model and hence forecasting, it is necessary to check if the data has seasonal component by looking into decomposed data into seasonal, trend and reminder components. 'R' provides 'stl()' function or 'decompose()' function for this.

In this case, Station demand for the period from 1.11.13 to 29.11.13 have been considered and plotted and decomposed (Fig. 9).
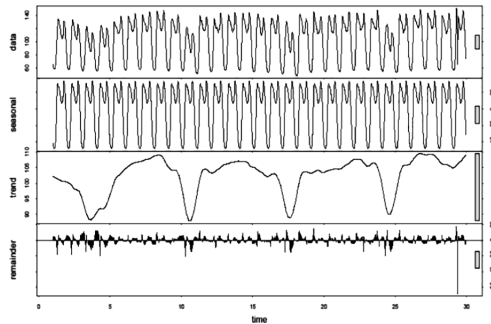
Fig. 9. Decomposition of time series

Step 2: Selecting the Suitable Model and Finding the Order of the Model

Model selected: Seasonal ARIMA model

Order of the model: The (p, d, q) and (P, D, Q) parameters of seasonal ARIMA model is obtained using auto.arima () function available in R. The order of the model is (Fig. 10),
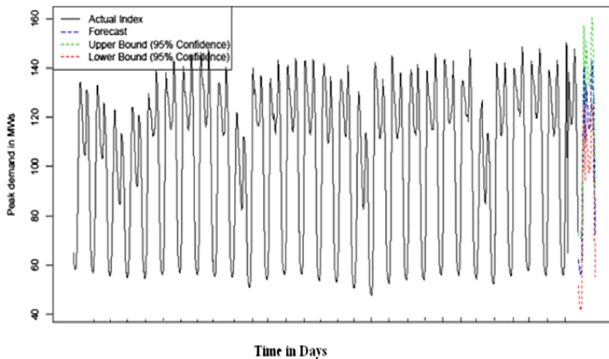
Fig. 10. Day-ahead forecast for station demand ARIMA (1, 0, 2) (0, 1, 2) [24]

Step 3: Plotting the Forecast

The day-ahead forecast obtained from **predict ()** function available in R is shown in
Fig. 11.
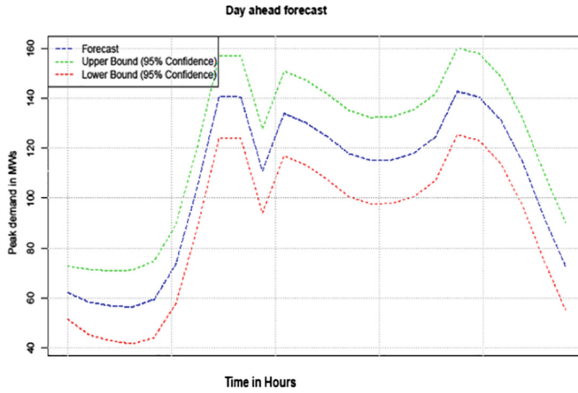
Fig. 11. Forecasted station demand for 30.11.13

Step 4: Model Validation by Plotting Diagnostic Plots

(See Fig. 12).
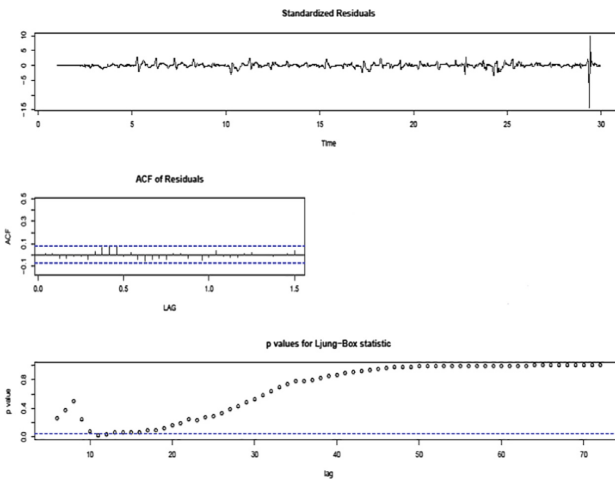
Fig. 12. Diagnostic plots (Color figure online)

Standardised Residuals

Standardized residual is the plot of every residual on the same standardized y-axis. The
distance from the line at 0 is how bad the prediction was for that value.

Residual is the difference of Observed and Predicted values. Positive values for the residual indicate that the prediction was too low and negative values mean the prediction was too high; 0 indicates the prediction was exactly correct.

ACF of Residuals

An ACF plot of the residuals will reveal if there is any autocorrelation in the residuals. The ACF plot shows that the model has identified the patterns in the data quite correctly, as there is no significant spike in the ACF plot.

p-Values for Ljung-Box Statistic

The bottom plot gives p-values for the Ljung-Box-Pierce statistics for each lag up to 10. The dashed blue line is at 0.05. All p-values are above it. That's a good result.

6   Future Scope

The big data analysis has been done in a micro level considering one 220 kV sub-station and connected eight 66 kV sub-stations. This can be extended for the entire electrical grid with more Gigabytes or Terabytes of data. Further, for short term load forecasting, univariate time series model has been considered i.e., forecasting is done considering past demands of sub-stations. The forecasting model can be made multi-variate considering variation of temperature as well.

7   Conclusion

All the domains of the industry are constantly looking for the ways to improve their proficiency. Power Utility with the prime objective of providing quality and reliable power supply is no exception to this, as they are facing one of the biggest challenges of meeting constantly increasing electricity demands and also handling/analyzing ample of diverse data. Utilization of big data analytics for data analysis and load research in power systems offers a promising way to enhance the efficiency. The application of big data analysis at a micro level is presented in this paper.

References

1. Kezunovic, M., Grijalva, S.: Role of Big Data in improving power system operation and protection. In: 2013 IREP Symposium Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid (IREP) (2013)
2. Anil, C.: Benchmarking of data mining techniques as applied to power system analysis. In: Uppsala University Publications (2013)

3. Konopko, J.: Big data solutions for smart grids and smart meters. In: Ryżko, D., Gawrysiak, P., Kryszkiewicz, M., Rybiński, H. (eds.) Machine Intelligence and Big Data in Industry. SBD, vol. 19, pp. 181–200. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-30315-4_16
4. Seppala, A.: Load Research and Load Estimation in Electricity Distribution. Technical Research Centre of Finland (1996)
5. Venables, W.N., Smith, D.M.: An Introduction to R. Network Theory (2009)
6. Prajapati, V.: Big Data Analytics with R and Hadoop. Packt Publishing, Birmingham (2013)
7. White, T.: Hadoop: The Definitive Guide. Shroff Publishers & Distributers Private Limited, Mumbai (2015)

Backward – Forward Algorithm Approach for Computation of Losses in LVDS and Proposed HVDS - Towards Loss Minimization and Voltage Improvement in Agricultural Sector

G. B. Prakruthi[1] and K. T. Veeramanju[2(✉)]

[1] Karnataka Power Transmission Corporation Limited, Bengaluru, India
prakruthi1980gb@gmail.com
[2] Sri Jayachamarajendra College of Engineering (SJCE), Mysuru 570 006, India
veeramanju.kalyan@gmail.com

Abstract. The main objective of this work is to propose High Voltage Distribution System (HVDS) method as an effective approach for improving the voltage profile and reducing the power loss in a power distribution system. Secondly to justify that the proposed method is both technically feasible and economically viable approach for reducing the distribution system losses and also to maintain the voltage profile in distribution system. A software application has been developed using JAVA for pre and post implementation analysis. The process of backward-forward iteration technique for power loss calculation, annual energy savings and pay-back period has been developed. The analysis indicates that, HVDS is an effective and economically feasible approach for distribution system loss reduction and voltage profile improvement for tail end consumers. The problem to supply electricity to the end users with minimum losses in developing countries like India is a big challenge, where about 72% of human resources are inhabited in rural areas. Though investment is made at distribution side, it has not gained much technical impact when compared to generation and transmission systems. The losses in distribution networks are considerably high as the power distribution is done using Low Voltage Distribution System (LVDS). The losses are high in rural areas when compared to urban areas because of the widely spread loads, specially the agricultural loads. To overcome all these problems HVDS method is proposed as a best move to enhance the overall efficiency of the Distribution System. A case study analysis of BESCOM's HVDS scheme is presented in this paper using the developed JAVA based program.

Keywords: Distribution system · Power system
Low Voltage Distribution System (LVDS)
High Voltage Distribution System (HVDS) · Backward-forward iteration

1   Introduction

Energy is the most important component which is necessary for the economic development of any country. Modern society is so dependent upon energy and it has become an important part of our daily life. The availability of considerable amount of energy in the present times has resulted in economic growth of a country. The per capita consumption of energy in a country is an indicator of the standard of living of its people. The per capita consumption of India during 2015–16 is 1075 kWh [1]. The most important form of energy is Electrical energy. Electric power is produced by generating station and this generated power reaches to the consumer end by means of transmission and distribution interconnected systems. But, the units generated at generating station will not wholly reach the consumers. The reasons can be attributed to losses during transmission of power from generating station to consumer end which may be due to losses in the conductors and the efficiency of equipments used for transmission, transformation, sub-transmission and distribution of power. These losses are termed as Transmission & Distribution loss (T&D loss). The T&D loss in India has been officially indicated as around 23% of total generation [2]. The distribution system (DS) is the most vital part of the power system as they form the link between bulk generation and the consumers. Maintenance and operation of distribution system requires a huge amount of investments. The technical impact of the investment of around 30 to 40% in the distribution system is not observable, compared to that seen in generation and transmission systems. Distribution system is a major contributor to the system losses and is about 70% of the total electrical system losses [3]. The high loss in distribution system is attributed to low voltage power distribution. The DS losses are divided as Technical and Non-Technical losses. Technical losses consist of line loss and transformer losses. Non-Technical (Commercial losses) mainly includes theft of energy by direct hooking to LT network, non-payment, improper billing, meter tampering, etc. Commercial losses are mainly associated with LT network. Both these losses add to give high Aggregate losses (AT&C) in DS. The distribution network also has a typical feature that the voltage at nodes reduces if moved away from substation.

   In order to improve the overall performance and efficiency of the electrical system, there is a need to analyse the existing distribution system and to optimize the system so that they can cater future loads with high reliability and decreased losses. To improve the overall efficiency of power delivery requires the distribution utilities to reduce the losses at the distribution level. Considerable research work has been done with a view of reducing loss level in distribution system, mainly in the perspective of active power I^2R loss reduction. Many arrangements like Feeder reconfiguration, Replacement of conductors with higher capacity, Distributed generation implementation, Capacitor placement, HVDS, etc. has been worked out to reduce distribution losses and to improve the voltage profile. One of the effective approaches to improve the efficiency of existing distribution system is implementation of HVDS. The main contribution of this work is to highlight the importance of proposed HVDS method with special reference to the agricultural sector as one of the best technically feasible and financially viable approach for providing reliable and quality supply to consumers.

2 Low Voltage Distribution System (LVDS)

The existing distribution system in India employs large three-phase 11 kV main distribution feeders (primary distribution) with three-phase spur lines and three-phase distribution transformers transforming 11 kV into 0.4 kV. Distribution system with low voltage employs lengthy low tension (LT or Secondary) lines, and multiple loads are fed from a bulk distribution transformer (Mother Transformer-MTR). Particularly, in rural areas where loads are widely dispersed, even to cater power supply to small loads the low tension (LT) lines run for longer distances [4]. Low voltage distribution is done either by three-phase four-wire, three-phase five-wire, single phase three-wire and single phase two-wire low tension lines, depending on the load type. This distribution system involves nearly 2:1 ratio of low and high voltage line lengths [5] or even more in rural areas. This topology of existing LVDS is results in, high system losses, affecting voltage profile and reducing the overall performance of distribution system. Hence, such system is not suitable for Indian conditions, especially for agricultural sector.

2.1 Major Draw Backs Associated with LVDS

The drawbacks of LVDS are as follows: (1) voltage profile at the load ends decreases as power distribution is done at low Voltages to large distances. (2) High quantum of losses because of low voltage profile and high current in LV (Low Voltage). (3) Theft of power at agricultural areas by unauthorized direct hooking on LT(Low Tension) lines. This leads to both technical loss and commercial loss. (4) High Rating Distribution Transformers (DTC) are required to supply power to the consumers and any failure of DTC affects large consumers. (5) High DTC failure rate due to overloading and repair and maintenance requires high expenditure. (6) Loss of standing crops due to delay in replacement of failed DTC, which affects the consumers. (7) Agricultural motor burn-outs are more as the motor tends draw more current at low voltage profile. (8) Monitoring of unauthorized hooking or tapping the bare conductors of LT lines, is more difficult.

2.2 Reasons for High Technical and Commercial Losses [6–8]

The losses in the secondary distribution network have a major contribution for higher loss in distribution system, as the power supply is done at LV 0.4 kV or 0.2 kV. Low voltage (LV) for the distribution of power leads to the high current and thus, resulting in more losses. The other factors which increase the Technical and Commercial loss level at primary and secondary distribution systems are as follows: (a) **Technical Reasons**: (i) Lengthy distribution linesresults in high line resistance and thereby I^2R losses in the line increases. (ii) Non- optimal conductor size will increase the line loss and also causes voltage drop. (iii) Improper location of DTC increases I^2R losses. (iv) Use of over-rated DTC and their under-utilization draw unnecessary high iron loss and the overall system loss increases. (v) Low Voltage appearing at transformers and consumers terminals results in higher currents drawn for the same output, this increases the I^2R loss. (vi) Un-balanced load at DTC causes high neutral currents leaking and loss of power (I^2R loss). (vii) Inadequate reactive power compensation draws more

current and the losses proportional to the square of the current will increase. (viii) Inadequate augmentation of transmission and distribution system results in overloading of the system. (b) **Commercial Reasons**: (i) Loss at consumer meters. (ii) Tampering or bypass of meters. (iii) Theft of power is done by connecting hooks of wires on LT distribution system. (iv) Errors in meter reading and bills.

3 Adoption of HVDS for Loss Reduction and Voltage Improvement

Adoption of three phase HVDS maintains, better voltage profiles to agricultural pump sets, reliability of power supply to the consumers in agricultural sector [9] and also reduces the technical and non-technical losses appreciably. This can be explained by one single illustration that for a 100 kVA load, the amperes at 11 kV is 5 Amps whereas it is 140 Amp at LT voltage of 415 V [10]. In HVDS, power is distributed mainly through high voltage lines (HV or 11 kV). This system employs the combination of 11 kV three-phase and single phase configuration (depending on the type of loads) of small capacity distribution transformers (5 kVA, 10 kVA, 15 kVA, 25 kVA) extending supply to 8 to 10 consumers or 2 to 3 agricultural loads with least LT lines, thereby reducing, losses, overloading, distribution transformer failure and improving the efficiency of the system.

3.1 Conversion of LVDS into HVDS

HVDS is a system, in which 11 kV lines are extended as nearer to the loads as possible and small capacity transformers are erected and supply is provided with minimum LT line. The procedure involved in LVDS to HVDS conversion is as follows: HVDS is constructed by converting the existing LV lines to three phase, three-wire HV lines and replacing low voltage three-phase cross arm by 11 kV V-shape cross arm. Three 11 kV insulators are provided on which the 11 kV wires are re-laid and intermediate poles are erected if required depending on the line span. Each pole on 11 kV side is earthed to provide safety to the human being and animals. Once conversion of LT to HT lines is completed, smaller capacity individual transformer 5 kVA, 10 kVA, 16 kVA is installed either on a single or double pole structure [11], to cater supply to agricultural loads and loads other than agricultural loads depending on field requirement. Finally, the power is delivered to consumers with NO LT or less LT. Unavoidable LT lengths are shielded by ABC to eliminate power theft. In most of the conditions this conversion does not involve the acquisition of any additional land and there will not be any depletion of cultivable or forest land as the conversion is being done on the existing poles. Therefore, there is no requirement of any R.O.W (Right of Way) for the erection of lines.

3.2 Benefits of High Voltage Distribution System

The following are the benefits of HVDS compared to LVDS [12–14]: (i) As High Voltage is transferred nearer to the loads, the power loss will be less. (ii) Voltage profiles at the loads will improve, as the DTC's (Distribution Transformer Centre) are erected

near to the load. (iii) Improving voltage profile reduces the power losses contributing to an improvement in power loss profile. (iv) Improvement in voltage profile near agriculture pump setsreduces motor burnouts. (v) Reduction of unauthorized agriculture connections, as the agriculture consumers will have a feeling of ownership of transformer due to limited connections on it. (vi) Low rating kVA DTCs are enough to supply loads. (vii) In the event of failure of distribution transformer it will affect only a small number of consumers. (viii) As LT overhead line is avoided and load per DTC is restricted, failure on account of over load and LT faults are reduced. (ix) Accidents due to touching of snapped conductors are reduced as the breaker trips at substation. (x) Losses are reduced considerably and power can be supplied to additional loads without additional investment on infrastructure. (xi) High quality of power supply earns total consumer satisfaction. (xii) Reliability of the system can be increased [15].

4 Case Study

A case study analysis of real three phase, three wire secondary distribution system of 433 V, of an 63 kVA MTR (Mother Transformer) in KonalaDoddi village 63 kVA MTR is presented in this paper. The data and the single line diagram have been taken from Bangalore Electricity Supply Company Ltd, Karnataka. This case study includes determination of voltage profile and power losses (I^2R losses) calculation of both existing LVDS and proposed HVDS. Also the annual energy savings and payback period of proposed HVDS scheme are calculated and presented in this paper.
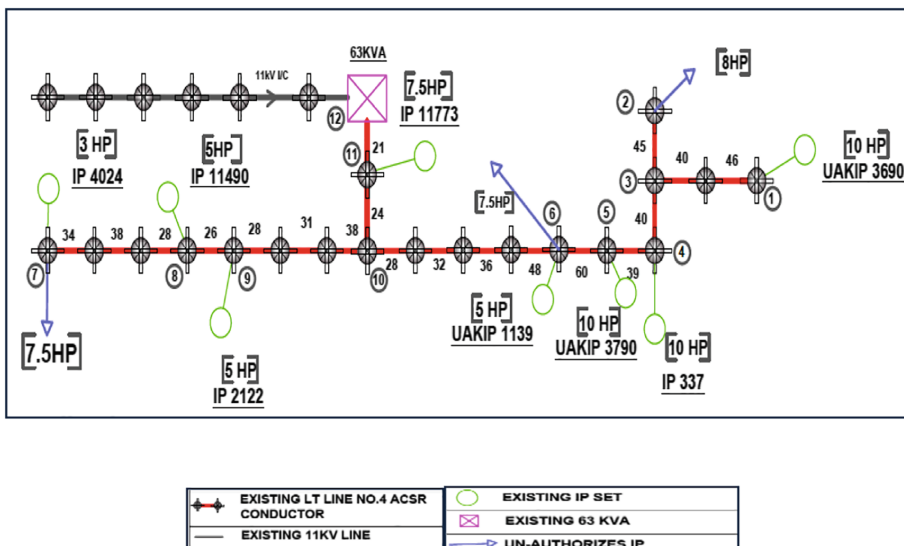
Fig. 1. Single line diagram of existing LVDS
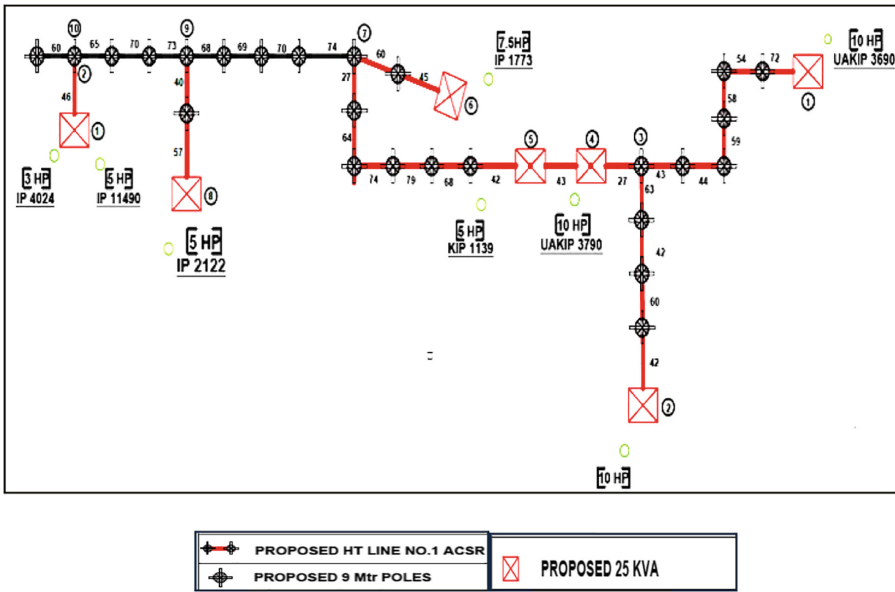
Fig. 2. Single line diagram of proposed HVDS

Data of existing LVDS are shown in Table 1. Data of proposed HVDS system are shown in Tables 2 and 3. The single line diagram (SLD) of existing LVDS is shown in Fig. 1 and the resultant proposed HVDS is shown in Fig. 2.

Table 1. Data of existing LVDS

Branch	From node	To node	Consumer at To node	Total load on branch (HP)	Total load on branch (W)	Branch LT length (Km)
1	3	1	UAKIP3690	10	7460	0.086
2	3	2	Hook(01 no)	8	5968	0.045
3	4	3	-	18	13428	0.040
4	5	4	IP337	28	20888	0.039
5	6	5	UAKIP3790	38	28348	0.060
6	10	6	KIP 1139 & Hook(01 no)	50.5	37673	0.144
7	8	7	IP4024 & Hook(01 no)	10.5	7833	0.1
8	9	8	IP11490	15.5	11563	0.026
9	10	9	IP2122	20.5	15293	0.097
10	11	10	-	71	52966	0.024
11	12	11	IP11773	78.5	58561	0.021

Table 2. Data of proposed HVDS (Circuit-I)

Branch	From node	To node	Transformer at To node		Total load on branch (HP)	Total load on branch (W)	Branch HT length (Km)
			Capacity (KVA)	Number of transformer			
1	2	1	25	1	25	25000	0.046

Table 3. Data of proposed HVDS (Circuit-II)

Branch	From node	To node	Transformer at To node		Total load on branch (HP)	Total load on branch (W)	Branch HT length (Km)
			Capacity (KVA)	Number of transformer			
1	3	1	25	1	25	25000	0.33
2	3	2	25	1	25	25000	0.208
3	4	3	-	-	50	50000	0.027
4	5	4	25	1	75	75000	0.043
5	7	5	25	1	100	100000	0.354
6	7	6	25	1	25	25000	0.105
7	9	7	-	-	125	125000	0.281
8	9	8	25	1	25	25000	0.097
9	10	9	-	-	150	150000	0.208

4.1 Method of Analysis

For the conversion of LVDS into HVDS, the power factor (PF) is taken as 0.85. In existing LVDS, length of LV line runs for 0.682 Km using 4 ACSR conductor having resistance value of 1.37 Ohms/Km The number of authorized and un-authorized consumers on this MTR is 8 and 3 respectively. The total connected load on the MTR is 78.5 HP. In proposed HVDS, The HT line is run 1.21 Km using Rabbit conductor having resistance value 0.5426 Ohms/Km and 7 numbers of 25 kVA transformers are proposed.

Determination of Voltage Profile in Existing LVDS. Using backward-forward sweep method, the voltage profile is determined [16]. The process is as follows: Initially the voltages at all nodes are assumed to be 433 V in case of LVDS.

 Step 1: Backward sweep: The branch currents are calculated from tail end loads till the source using:

$$Ii = \frac{Pi}{\sqrt{3} * Vi * PF} \tag{1}$$

 Step 2: Voltage drops in each branch are determined form the corresponding branch current obtained in Step 1 using:

$$Vdi = Ii * Ri$$

Step 3: Forward sweep: In this moving forward from source towards tail end, the voltages at each node is calculated using the corresponding voltage drops obtained in Step 2 and the voltages are updated. The voltage of source end is always set at 433 V.

Step 4: This iteration process is continued until we get a constant voltage value at tail end or if it is within set tolerance level.

Determination of Voltage Profile for Proposed HVDS. Initially the voltages at all nodes are assumed to be 11 kV in case of HVDS and voltage of source end is always set at 433 V. The voltage profile is determined by backward-forward sweep method as explained above (Step 1 to 4). The voltages obtained here are with respect to primary (HV side).

A short LV line runs from secondary of DTC to consumer. The secondary voltage from DTC to consumer is Vi' and is obtained by

$$Vi' = KVi \tag{2}$$

Where,

$$K = \frac{433}{11000}$$

The forward – backward sweep method can be represented in the form of a flow chart as shown in Fig. 3.
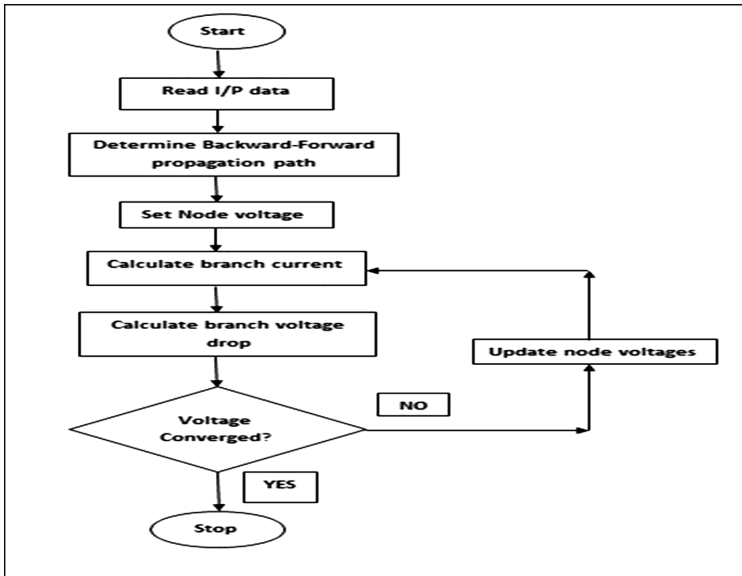
Fig. 3. Flow chart of backward-forward sweep method

Table 4 indicates the comparison of voltage levels (V) at consumer end, of both existing LVDS and HVDS. The comparison of consumer end voltage levels of LVDS and HVDS is shown graphically in Fig. 4 (The order of numbering the consumers is from source to tail end).

Table 4. Voltage profile at consumer end in existing LVDS and proposed HVDS

RR no.	Consumer	LVDS	HVDS
IP 11773	Consumer 1	430.3407	432.92
IP 2122	Consumer 2	424.3208	433
IP 11490	Consumer 3	423.6604	433
IP 4024	Consumer 4	421.9328	432.96
KIP 1139	Consumer 5	415.4218	432.89
UAKIP 3790	Consumer 6	411.5761	432.88
IP 337	Consumer 7	409.7259	432.87
UAKIP 3690	Consumer 8	407.0355	432.87
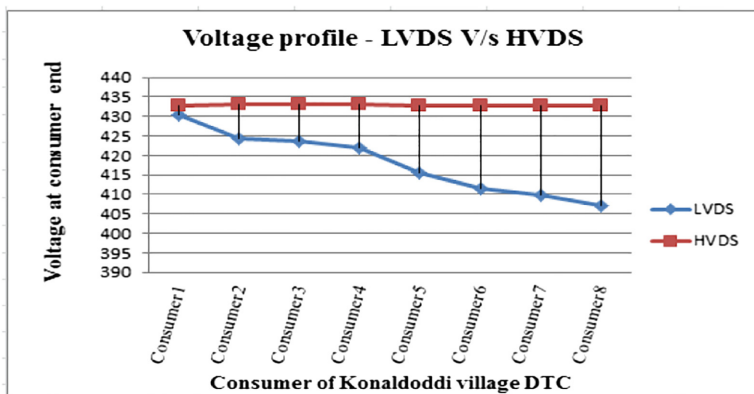
Fig. 4. Voltage profile at consumer end in LVDS and HVDS

Determination of Loss Profile. The loss level in existing LVDS and proposed HVDS system are calculated as follows: For line losses the branch current obtained after 5th iteration are considered for LVDS and the branch currents obtained after 3rd iteration are considered. The power loss of each branch and the total power loss is calculated taking summation of each branch as given by [17, 18].

$$Pi = 3 * Ii^2 * Ri \qquad (3)$$

$$P = \sum Pi \qquad (4)$$

The appropriate values of transformer losses in LVDS and HVDS are chose from Table 5.

Table 5. Particulars of transformer losses

S. no.	KVA rating	No load losses	Load losses
1	5	40	150
2	6.3	40	150
3	10	40	225
4	15	60	275
5	25	110	720
6	63	200	1300
7	100	290	1850

$$P_{DTC} = P_{NL} + P_{FL} \tag{5}$$

Theft loss P_T by direct hooking in existing LVDS is calculated based on the un-authorised load connected in HP.

$$P_T = HP * 746 \tag{6}$$

Total LVDS loss is given by:

$$P_{LVDS} = P + P_{DTC} + P_T \tag{7}$$

Total HVDS loss is given by:

$$P_{HVDS} = P + P_{DTC} \tag{8}$$

Reduction in losses (W) is given by

$$P_R = P_{LVDS} - P_{HVDS} \tag{9}$$

Total reduction in losses per annum (Units) is calculated by,

$$P_A = (P_R \times 6 \times 365) \div 1000 \tag{10}$$

$$\text{Annual Energy Savings} = (\text{Cost per unit } \times P_A) \tag{11}$$

$$\text{Payback period} = (\text{Capital outlay} \div \text{Annual Energy Savings}) \tag{12}$$

The corresponding losses in existing LVDS and proposed HVDS are illustrated in Table 6. The results of the case study are tabulated in Table 7.

The overall system losses of Konala Doddi MTR, gets reduced from **23567.51 W to 5870.164 W** after implementation of HVDS project. Figure 5 shows graphical representation of comparison of power losses in LVDS and proposes HVDS.

Table 6. Comparison of losses

Sl. no.	Particulars	Existing LVDS (W)	Proposed HVDS (W)
1	Line loss	4909.51	60.16388
2	No-load loss	200	770
3	Full-load loss	1300	5040
4	Theft loss	17158	-
5	Total Loss	23567.51	5870.164

Table 7. Results of case study

Sl. no.	Particulars	Remarks
1	% net reduction in losses	75.092
2	Annual savings (Rs)	150377.89
3	Capital outlay of proposed system (Rs)	409907
4	Payback period	5.39 years
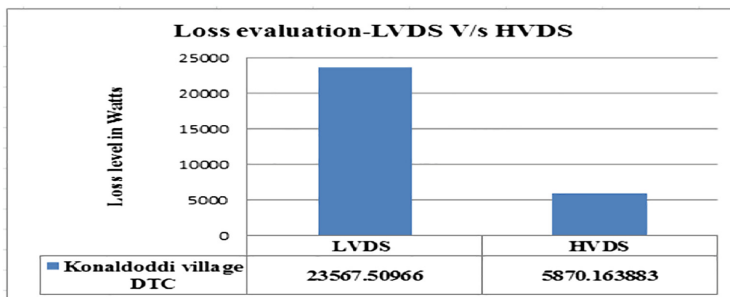
Fig. 5. Comparison graph of power loss of LVDS and HVDS

4.2 Case Study Results of 10 Numbers of 63 kVA MTRs

Analysis results for 10 numbers of 63 kVA MTRs including the above case study is also presented in this paper. Table 8 gives list of 10 MTRs.

Figure 6, indicates the graph showing the voltage profile improvement at tail end consumers in case of all 10 MTRs. Figure 7, represents a graph showing comparison of losses in existing LVDS and proposed HVDS, of all 10 MTRs. Annual energy savings and Payback period for proposed HVDS scheme, of all 10 MTRs is shown in Fig. 8.

4.3 Benefits Achieved After Implementation of HVDS

The following are the benefits achieved after implementation of HVDS for above MTR: (i) The failure of equipment will be least, as no overloading and no meddling with LT line. (ii) Failure of equipment affects only 2 to 3 IP consumers. (iii) The registered consumers take the ownership and will not allow others to hook onto LV

Table 8. List of 10 MTRs

Feeder name	Name of the MTR
Thokasnadra	Balagere village
Thokasnadra	Donasnadra village
Thokasnadra	Konala Doddi village
Thokasnadra	M.maniyamballi village
Doddala Halli	Ullagallu gate
Doddala Halli	Therina Doddi village
Shivanahalli	Shivanahalli mata
Thokasnadra	Ajjegowdan Walse village
Doddala Halli	Kunur village
Doddala Halli	Karirayana Doddi
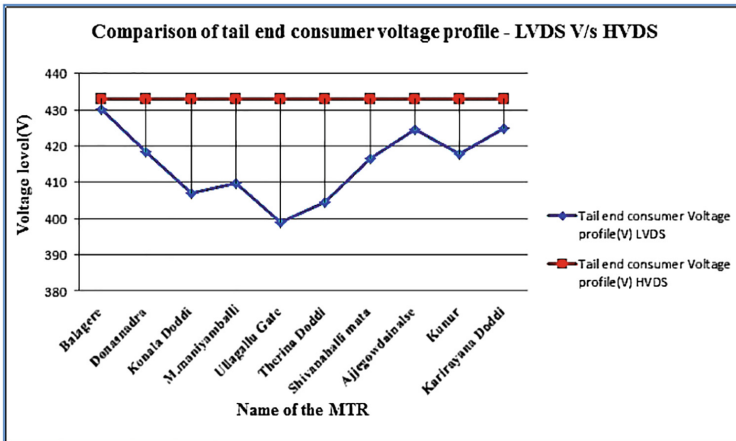
Fig. 6. Voltage profile improvement at tail end consumer
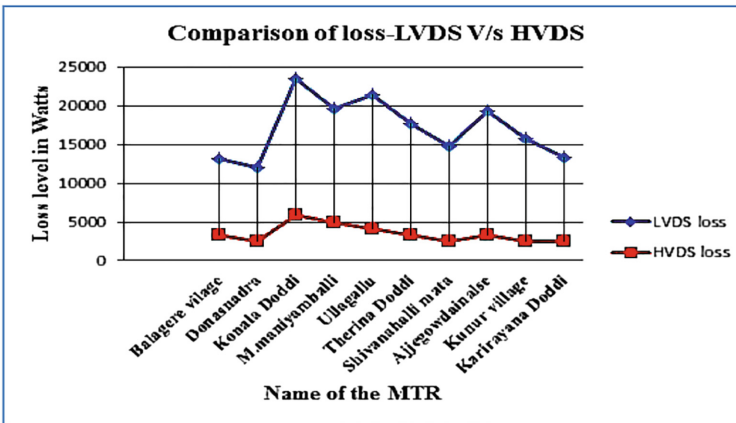
Fig. 7. Comparison LVDS and HVDS losses of all 10 MTRs
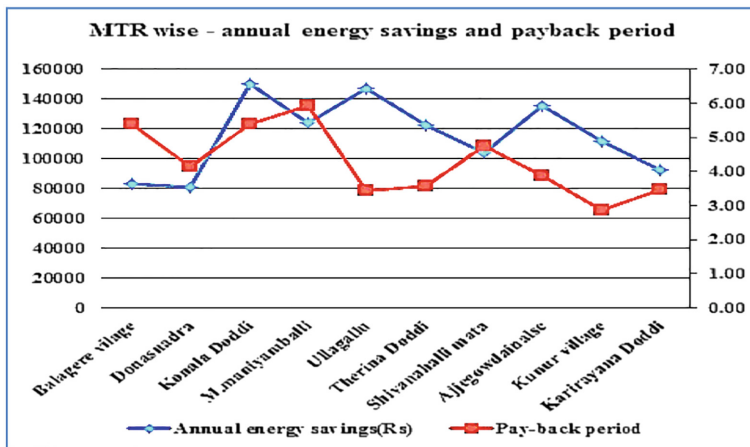
Fig. 8. Annual energy savings and payback period, all 10 MTRs

network. (iv) The motor burnouts are reduced because of improved voltage profile. (v) Interruptions are reduced and power quality is improved.

5   Conclusion

The fact is that, all energy supplied to a distribution utility does not reach the end consumers as a there will substantial amount of loss of energy. Loss Minimization in power networks has assumed greater significance, in light of the fact that enormous amount of generated power is continuously wasted as losses. The pressure of enhancing the overall capability of power delivery has forced the Indian power utilities to reduce the loss, particularly at the distribution level by adopting various loss reduction techniques. Based on the analysis results of proposed HVDS project work of BESCOM, it is observed that by implementation of HVDS, there is an improvement of voltage profile in distribution system and the distribution system technical and non-technical losses gets reduced. However, even though the capital cost is marginally higher because of use of more number of small capacity transformers; it has been proved in the analysis that the investment on conversion from LVDS to HVDS can be easily recovered by the utility, by way of loss reduction in a short span of time. The analysis based on the backward forward approach in computing concludes that, the proposed HVDS scheme is an effective approach for, reducing technical and non-technical losses occurring at LT side of the distribution system and to improve the system performance in agricultural sector. This proposed method in agricultural sector presents one of the best technically feasible and financially viable method for providing reliable and quality supply to consumers.

References

1. Government of India Power Sector, Central Electricity Authority New Delhi (2017)
2. Transmission and Distribution Losses (Power). TERI report
3. Ramesh, L., Madhusudhanaraju, M., Chowdhury, S.P., Chowdhury, S.: Voltage profile improvement through high voltage distribution system. Chennai and Dr. MGR University, SEISCON 2011, Chennai, Tamil Nadu, India (2011)
4. Bansal, I., Gill, H.S., Gupta, A.: Minimization of losses by implementing high voltage distribution system in agricultural sector. IOSR J. Electr. Electron. Eng. (IOSRJEEE) **1**(5), 39–45 (2012)
5. Kapuree, V.S., Mahajan, K.M.: Loss reduction by improving ratio of HT/LT line in electrical distribution system. IJOER Peer Rev. Int. J. **4**(1) (2016)
6. Ghosh, S.: Loss reduction and efficiency improvement: a critical appraisal of power distribution sector in India. IJMER **2**(5), 3292–3297 (2012)
7. Chen, S.X., Eddy, Y.S.F., Gooi, H.B., Wang, M.Q., Lu, S.F.: A centralized reactive power compensation system for LV distribution networks. IEEE Trans. Power Syst. **30**(1), 274–284 (2015)
8. Chibuisi, I., Okonba Brown, J.: Modeling of high voltage distribution system using Simulink. Int. J. Trend Res. Dev. (IJTRD), **2**(6) (2015)
9. Spandana, K., Varsha Reddy, A.: Restructuring of a low voltage distribution system into a high voltage distribution system for an improved voltage and power loss profile. In: International Conference and Utility Exhibition 2014 on Green Energy for Sustainable Development (ICUE 2014), Thailand (2014)
10. Thakur, R., Chawla, P.: High voltage distribution system (HVDS) - an alternate for improvement of voltage drop profile. IJETMAS **3**(1), 197–203 (2015)
11. Narsi Reddy, V., Krishnarjuna Rao, S., Nagendra Kumar, S.: Comparative study on loss & cost minimization by using high voltage distribution system. Imp. J. Interdisc. Res. (IJIR) **2** (11) (2016)
12. Sampath Kumar, S.A., Vasudaven, V., Antony, J., Raju, M.S., Ramesh, L.: Minimization of power losses in distribution system through HVDS concept. In: International Conference on sustainable Energy and Intelligent System, Chennai, India, pp 86–89 (2011)
13. Jahnavi, W.V.: Reduction of losses in distribution system using HVDS with real time application. IJAREEIE **3**(9) (2014)
14. Solanki, A.S., Pal, R.K., Aslam, K.: Rehabilitation of electrical power losses by implementing high voltage distribution system. In: National Conference on Synergetic Trends in engineering and Technology (STET-2014) International Journal of Engineering and Technical Research, Special Issue (2014)
15. Vidyasagar, E., Devender Rao, K., Prasad, P.V.N.: Reliability improvement of high voltage distribution system over low voltage distribution system. In: National Conference on Power Distribution, CPRI, Bangalore (2012)
16. Ulinuha, A., Masoaum, M.A.S., Islam, S.M.: Unbalance power flow calculation for a radial distribution system using forward-backward propagation algorithm
17. Dembra, A., Sharma, A.K.: Improvement in voltage profile and loss minimization using high voltage distribution system. IJEER **2**(3), 11–20 (2014)
18. Jain, S., Singh, R.: Enhancement of the distribution system by implementing LT - less distribution technique. IJSRP **3**(10) (2013)
19. Agüero, J.R.: Improving the efficiency of power distribution systems through technical and non-technical losses reduction

20. Magadum, R.B., Timsani, T.M.: Minimization of power loss in distribution networks by different techniques. IJSER **3**(5) (2012)
21. Ramesh, L., Chowdhury, S.P., Chowdhury, S., Natarajan, A.A., Gaunt, C.T.: Minimization of power loss in distribution networks by different techniques. Int. J. Electr. Electron. Eng. **3**(9), 521–527 (2009)
22. Kour, G., Sharma, R.K.: Different techniques of loss minimization in distribution system. Int. J. Enhanc. Res. Sci. Technol. Eng. **2**(2) (2013)
23. Peponis, G.J., Papadopoulos, M.P., Hatziargyriou, N.D.: Distribution network reconfiguration to minimize resistive line losses. IEEE Trans. Power Deliv. **10**(3), 1338–1342 (1995)
24. Siti, M.W., Nicolae, D.V., Jimoh, A.A., Ukil, A.: Reconfiguration and load balancing in the LV and MV distribution networks for optimal performance. IEEE Trans. Power Deliv. **22**(4), 2534–2540 (2007)
25. Borges, C.L.T., Falcao, D.M.: Impact of distributed generation allocation and sizing on reliability, losses, and voltage profile. In: 2003 IEEE Bologna Power Tech Conference Proceedings, Bologna, 23–26 June 2003, vol. 2 (2003)
26. Kiran, I.K., Laxmi, J.: Shunt versus Series compensation in the improvement of power system performance. Int. J. Appl. Eng. Res. **2**(1), 28–37 (2011)
27. Michiline Rupa, J.A., Ganesh, S.: Power flow analysis for Radial distribution system using backward-forward sweep method. Int. J. Electr. Comput. Energ. Electron. Commun. Eng. **8**(10), 1540–1544 (2014)
28. 24 × 7 Power for All – Karnataka: A Joint Initiative of Government of India and Government of Karnataka. Document

Analysis of Segmentation Methods on Isolated Balinese Characters from Palm Leaf Manuscripts

Deepak Kumar$^{(\boxtimes)}$, K. Vatsala , Sushmitha Pattanashetty ,
and S. Sandhya

Dayananda Sagar Academy of Technology and Management (DSATM),
Udayapura, Kanakapura Main Road, Bengaluru 560082, India
dipkmr@gmail.com, vatsala.5996@gmail.com,
sushpattanashetty@gmail.com, ssandhya196@gmail.com

Abstract. Segmentation of an image is a complex process when image has to be divided into constituent and meaningful parts. In case of a clean image, the segmentation process is an easy task due to distinct foreground and background within the image. Whereas, the images captured or scanned from palm leaf manuscripts, the segmentation process is more complex and complicated. In this paper, we have used the individual characters from the palm leaf manuscript for segmentation. We have chosen multiple segmentation algorithms to perform segmentation of isolated Balinese characters. We evaluate the performance of all the algorithms on AMADI_Lontarset dataset. From our analysis, we observe that a global thresholding approach provides good segmentation on the set of analyzed images over the local thresholding approach.

Keywords: Palm leaf manuscript · Isolated character segmentation
Segmentation · Balinese script · Performance evaluation

1 Introduction

Earliest form of storing information was carried out by writing on palm leaves. The writing mainly covers wisdom and heritage from ancient age on manuscripts. Palm leaf manuscripts are prepared from the dried palm leaves. The leaves are susceptible to deterioration such as aging, stains, and insect bites. To avoid deterioration, smoke treatment is performed. The palm leaves used for writing are derived from Borassus species (Palmyra palm) or Ola leaf (leaf of Coryphaumbraculifera or the talipot palm). The earliest use of palm leaf for writing dates back to the 5th century BC [1].

The amount of information available for users has increased rapidly with the advancement in technology. Ancient documents are stored and retrieved using modern imaging tools [2]. Our key objective is to perform analysis on the historical palm leaf documents. Hence, it is essential to preserve and digitize these documents. Several superior methods are available to capture the image of the leaf using higher resolution cameras. Additionally, image processing techniques have contributed for preserving the information in palm leaf manuscripts in an effective way [3]. Even then, Multiple

artifacts are invariably added into the digital image while scanning [4]. Our research work involves in analyzing the digital documents for retrieval and restoration of ancient wisdom.

There are several deformations seen in Palm leaf manuscripts. Overcoming all of these offers us fresh challenges in document image analysis system. An example page from palm leaf manuscript is as shown in Fig. 1.
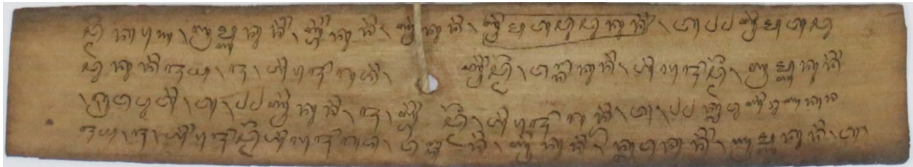
Fig. 1. A palm leaf manuscript image from AMADI_Lontarset dataset [5]. We can observe the deformations such as varying space between letters and lines within the image.

Palm leaf was used as a medium for writing for several centuries before the advent of paper. The manuscripts contain information about art, literature, science, technology, astronomy, medicine, and music. The manuscripts are mostly preserved in museums, temples and universities in many countries. There may be a possibility that the manuscripts might have been damaged before preservation. A partially damaged palm leaf image is as shown in Fig. 2. A special stylus is used to engrave letters on the palm leaf. These manuscripts are rubbed with charcoal powder to make the script readable just before reading.
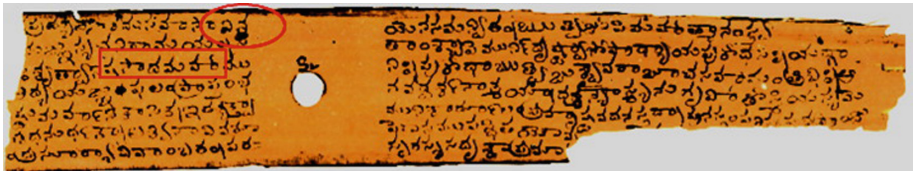
Fig. 2. A sample palm leaf manuscript which is partially damaged [4].

2 Literature Survey on Segmentation Methods

Segmentation of an image in document image processing field is known as binarization. In the literature, several segmentation methods have been published. A few of them are selected and used for the analysis of the segmentation results. We have broadly classified the segmentation methods into two types namely global threshold based segmentations and local threshold based segmentations. We have experimented using Su's, Sauvola's, Niblack's, Kittler's, Kapur's and Otsu's method. We have also experimented segmentation using power law transform (PLT) method [6], nonlinear enhancement and selection of plane (NESP) method [7], and midline analysis and propogation of segmentation (MAPS) method [8].

2.1   Otsu's Method

A segmentation method was proposed by Otsu, which is considered as the most efficient global thresholding method in the year 1979 [9]. It uses the histogram of an image for the calculation of Fischer discriminant value. It considers the pixel value of an image as a threshold. The histogram is divided into two parts to calculate the threshold value. This method does not look for the edges and selects the region with uniform pixel values for segmentation. Hence, this method does not work well in case of non-uniform background. A moving window is used to compute the local threshold for a region, which moves over the regions and calculates the threshold of each region [10].

The discriminant value used for calculating the threshold is:

$$\sigma^2(k) = max \frac{[\mu_T w(k) - \mu(k)]^2}{w(k)[1 - w(k)]} \tag{1}$$

Where,

$$w(k) = \sum_{i=1}^{k} p_i \tag{2}$$

$$\mu(k) = \sum_{i=1}^{k} ip_i \tag{3}$$

$$\mu_T = \sum_{i=1}^{L} ip_i \tag{4}$$

Here, 'L' is the number of gray levels, 'p_i' is the normalized probability distribution, 'k' denotes the level, 'μ' is the class mean level and 'w' denotes the probabilities of class occurrence.

2.2   Kapur's Method

Kapur proposed an entropy measure to calculate a threshold value for binarization of a document image [11]. It is an extended version of Pun's method. In this method the foreground entropy '$H_f(k)$' and the background entropy '$H_b(k)$' are calculated by dividing the document image [10]. The maximum is considered as the optimal threshold.

$$H(k) = max\big(H_f(k) + H_b(k)\big) \tag{5}$$

$$H_f(k) = \sum_{i=1}^{k} \frac{-p_i}{p(k)} log\left(\frac{p_i}{p(k)}\right) \tag{6}$$

$$H_b(k) = \sum_{i=k+1}^{L} \frac{-p_i}{1-p(k)} log\left(\frac{p_i}{1-p(k)}\right) \tag{7}$$

$$p(k) = \sum_{i=1}^{k} p_i \tag{8}$$

2.3 Kittler's Method

Kittler proposed a method based on the window size [12]. Whole image is considered as a single window and Sobel image is obtained for the threshold calculation [10]. The threshold is calculated using the edges falling within the same window.

$$T = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} g(i,j) * f(i,j)}{\sum_{i=1}^{M} \sum_{j=1}^{N} g(i,j)} \tag{9}$$

Here, 'g(i, j)' is the gradient image of the input image 'f(i, j)' [10].

2.4 Niblack's Method

A local threshold is calculated for every pixel in the image. This is done by making use of a rectangular window. This window is moved over the entire image [13]. For the threshold calculation, the pixels falling within the window are used [10].

$$T(x,y) = \mu(x,y) + k_n * \sigma(x,y) \tag{10}$$

$$\mu(x,y) = \frac{1}{N_w} \sum_i \sum_j f(x+i, y+j) \tag{11}$$

$$\sigma^2(x,y) = \frac{1}{N_w} \sum_i \sum_j (f(x+i, y+j) - \mu(x,y))^2 \tag{12}$$

Here, the values for 'k_n' and the window size are initially fixed. In our experiment, these values are $k_n = -0.2$ and $N_w = 8$ for all the images [10].

2.5 Sauvola's Method

An improvement was proposed by Sauvola for Niblack's method [14]. The modified threshold is given as:

$$T(x,y) = \mu(x,y) * \left[1 + k_s * \left(\sigma\frac{(x,y)}{R} - 1\right)\right] \tag{13}$$

Here, 'R' is the dynamic range. The parameter values are $R = 128$, $N_w = 25$ and $k_s = 0.2$ for calculating the threshold [10].

2.6 Su's Method

Su's method uses the local image statistics [15] for segmentation of document image. The local variations in background is reduced using the following formula,

$$D(x, y) = \frac{f_{max}(x, y) - f_{min}(x, y)}{f_{max}(x, y) + f_{min}(x, y) + \epsilon} \tag{14}$$

Here, '$f_{min}(x, y)$' and '$f_{max}(x, y)$' are the local minimum and maximum intensities [10], respectively. ϵ is a small positive number [10]. A global threshold such as Otsu is applied to segment the image $D(x, y)$ and the high contrast image is obtained. The formula for classification of pixel is

$$B(x, y) = \begin{Bmatrix} 1 & N_e \geq N_{min} \wedge f(x, y) \leq E_{mean} + E_{std}2 \\ 0 & otherwise \end{Bmatrix} \tag{15}$$

Where, E_{std} and E_{mean} are the standard deviation and the mean [10].

$$E_{mean} = \frac{\sum_{neighbor} f(x, y) * (1 - E(x, y))}{N_e} \tag{16}$$

$$E_{std} = \sqrt{\frac{\sum_{neighbor} (f(x, y) - E_{mean}) * (1 - E(x, y))^2}{N_e}} \tag{17}$$

Where, N_e is the number of high contrast pixels [10]. When a high contrast pixel is detected, $E(x, y)$ is set to 0 else 1. The neighborhood window size is 15 pixels and $N_{min} = 25$ pixels [10].

2.7 PLT Method

Power law transform (PLT) is a type of gray level transformation which is commonly known as gamma transform. Basically, this transformation was used in transmission of video. The adjacent characters in an image are combined during binarization for low contrast images. The characters merging is minimized before applying a threshold to the image. The characters contrast is improved by power law transformation which assists in better segmentation [6]. The power law transformation expression is

$$s = cr^{\gamma} \tag{18}$$

where 'r' and 's' are the gray values; c and γ are positive constants.

2.8 NESP Method

Nonlinear enhancement and selection of plane (NESP) method uses improvement over PLT method [7]. This method uses different color planes for segmentation such as Red, Green, Blue, Intensity, and Luminous. Out of five segments, the segment which has

highest discrimination factor is chosen as the best segmentation for the image. The discriminant factor is calculated as:

$$d_c^2(k) = max \frac{[\mu_{ct}w_c(k) - \mu_c(k)]^2}{w_c(k)[1 - w_c(k)]}, \ c \in [R, G, B, I, L]$$
(19)

Where,

$$w_c(k) = \sum_{i=1}^{k} p_{ci}$$
(20)

$$\mu_c(k) = \sum_{i=1}^{k} i p_{ci}$$
(21)

$$\mu_{ct} = \sum_{i=1}^{L} i p_{ci}$$
(22)

Here, the normalized probability distribution 'p_{ci}' [7].

2.9 MAPS Method

Midline analysis and propogation of segmentation (MAPS) method was proposed by Deepak et al. [8]. The middle row is selected and segmented. For segmentation, Niblack's threshold is used with a window. Then, the labels from the segmentation are used for labeling the other pixels in the image. The impact of degradations is reduced in this approach. This method assumes that degradation is not present in the middle row [8]. The threshold value calculated for the middle line is:

$$T_i = \mu_i + k_n * \sigma_i$$
(23)

$$\mu_i = \frac{1}{N_w} \sum_{j \in N_w} x_{i+j}$$
(24)

$$\sigma_i = \frac{1}{N_w} \sum_{j \in N_w} (x_{i+j} - \mu_i)^2$$
(25)

Here, 'x_i' is the gray value. The values for 'k_n' and N_w are 0.1 and min(height, width), respectively.

Bayesian binary classification is used for labeling the non middle row pixels. The posterior probability is calculated for each pixel $x_{i,j}$ as [8, 16]:

$$p(C_0/x_{i,j}) = \frac{p(x_{i,j}/C_0)p(C_0)}{p(x_{i,j}/C_0)p(C_0) + p(x_{i,j}/C_1)p(C_1)}$$
(26)

3   Performance Evaluation

The salient purpose of this paper is to evaluate the segmentation methods and perform analysis on them [17, 18]. In order to perform evaluation, we need a kind of measure which fits to our experiment. We have used cost function to rank the segmentation methods. The cost function for each segmentation method is calculated as:

$$C_i = \sum_{x=1}^{M} \sum_{y=1}^{N} [f(x, y) - b_i(x, y)]^2 \tag{27}$$

where 'f(x, y)' is the gray scale image and 'b_i(x, y)' is the segmented image from the ith method.

The calculated cost is used to rank the methods. The method which has the lowest cost is ranked as 1 and the method which the next lowest cost is ranked as 2 and the process is repeated until all the methods are ranked for an image. In case, if the two methods have same cost, then the two methods will be assigned with a same rank and the method which has the next lowest cost function is assigned with a rank that is incremented by 2. For example, if the two methods have same lowest cost, then the two methods are ranked as 1 and the method which has next lowest cost is ranked as 3. Skipping one rank place due to same cost for the two methods.

3.1   Dataset

AMADI_Lontarset dataset [5] is used for the performance evaluation. We have selected six training images from each class for evaluation. A total of 486 images were selected from 11,710 to determine better segmentation approach for these images. Sample images are as shown in Fig. 3.
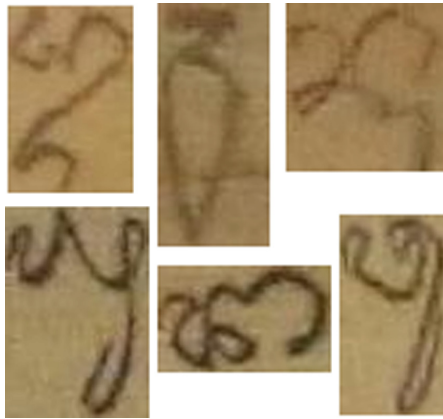
Fig. 3. Sample isolated Balinese character images from AMADI_Lontarset dataset [5].

4 Results and Discussion

The segmentation results for a sample image from the dataset is as shown in Fig. 4. It is observed from the figure that the global thresholding methods are good at segmenting the characters. Otsu's method and MAPS method have the lowest cost calculated for the performance evaluation.

Tables 1 and 2 provide the tabulated cumulative rank for the images considered for the experiment. Kittler's method and MAPS method in Table 1 have low values indirectly indicating that the global thresholding methods have achieved better segmentation score. Figure 5 shows the count of the number of images for which the method has lowest cost based on the individual image wise analysis. We can infer that none of the local thresholding methods had lowest cost for all the tested images from Fig. 5.

Fig. 4. Segmentation of a sample image from the dataset using different segmentation methods. Top row: Original image, results of Kapur's method, Otsu's method, Kittler's method, and Niblack's method, respectively. Bottom row: Results of Sauvola's method, Su's method, PLT method, NESP method, and MAPS method, respectively.

Table 1. Cumulative sum of ranks for global thresholding methods on a set of 486 images.

Global thresholding methods	Cumulative sum of rank
Kapur's method	1251
Otsu's method	2176
Kittler's method	**938**
PLT method	1300
NESP method	2245
MAPS method	**974**

Table 2. Cumulative sum of ranks for local thresholding methods on a set of 486 images.

Local thresholding methods	Cumulative sum of rank
Niblack's method	4190
Sauvola's method	3719
Su's method	3566
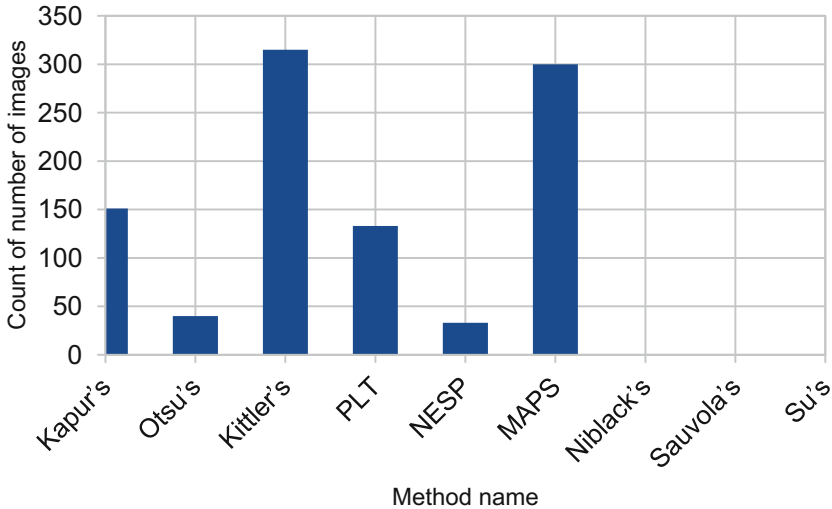
Fig. 5. The graph shows the count of the number of images for which the method has the lowest cost.

5   Conclusion and Future Work

We have worked on Challenge 3 train images of Balinese Script in Palm leaf Manuscript Images with 92 classes using nine binarization methods. From the performance evaluation, it is evident that the global thresholding methods are better than the local thresholding methods.

We do not have the ground truth for the test set. We couldn't perform recognition of the binarized images. Our future work is intended in developing a classifier to recognize the segmented characters. We further use the gray scale image and the binarized image for feature extraction to train the classifier.

We observed that the character images are of different sizes and training a classifier for images with different sizes is difficult. As future work, we are working on building patches based classifier which may provide more accuracy than any other methods.

References

1. Shi, Z., Setlur, S., Govindaraju, V.: Digital enhancement of palm leaf manuscript images using normalization techniques. In: Proceedings of the 5th International Conference on Knowledge-Based Computer Systems, Hyderabad, India, 19–22 December 2004
2. Chamchong, R., Fung, C.C.: Optimal selection of binarization techniques for the processing of ancient palm leaf manuscripts. IEEE (2010)
3. www.acharya.gen.in:8080/applications/palm_leaf.php. Accessed 25 Aug 2017
4. Vijaya Lakshmi, T.R., Sastry, P.N., Rajnikanth, T.V.: A novel 3D approach to recognize Telugu palm leaf text. Eng. Sci. Technol. Int. J. **20**(1), 143–150 (2017)

5. Burie, J.-C., Coustaty, M., Hadi, S., Kesiman, M.W.A., Ogier, J.-M., Paulus, E., Sok, K., Sunarya, I.M.G., Valy, D.: ICFHR2016 competition on the analysis of handwritten text in images of balinese palm leaf manuscripts. In: 15th International Conference on Frontiers in Handwriting Recognition, pp. 596–601, Shenzhen, China (2016)
6. Kumar, D., Ramakrishnan, A.G.: Power-law transformation for enhanced recognition of born-digital word images. In: Proceedings of the 9th SPCOM (2012)
7. Kumar, D., Anil Prasad, M.N., Ramakrishnan, A.G.: NESP: nonlinear enhancement and selection of plane for optimal segmentation and recognition of scene word images. In: Proceedings 20th DRR (2013)
8. Kumar, D., Anil Prasad, M.N., Ramakrishnan, A.G.: MAPS: midline analysis and propagation of segmentation. In: Proceedings 8th ICVGIP (2012)
9. Otsu, N.: A Thresholding selection method from gray-level histogram. IEEE Trans. SMC **9**, 62–66 (1979)
10. Kumar, D., Anil Prasad, M.N., Ramakrishnan, A.G.: Evaluation of document binarization using eigen value decomposition. In: Proceedings of the 20th DRR (2013)
11. Kapur, J.N., Sahoo, P.K., Wong, A.K.C.: A new method for gray-level picture threshold using the entropy of the histogram. Comput. Vis. Graph. Image Process. **29**, 273–285 (1985)
12. Kittler, J., Illingworth, J., Föglein, J.: Threshold selection based on a simple image statistic. Comput. Vis. Graph. Image Process. **30**, 125–147 (1985)
13. Niblack, W.: An Introduction to Digital Image Processing. Prentice Hall, Englewood Cliffs (1986)
14. Sauvola, J.J., Pietaikinen, M.: Adaptive document image binarization. Pattern Recogn. **33** (2), 225–236 (2000)
15. Su, B., Lu, S., Tan, C.L.: Binarization of historical document images using the local maximum and minimum. In: Proceedings of the 9th International Workshop on Document Analysis Systems (DAS 2010), pp. 159–166 (2010)
16. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification. Wiley, Hoboken (2001)
17. Trier, Ø.D., Taxt, T.: Evaluation of binarization methods for document images. IEEE Trans. PAMI **17**, 312–315 (1995)
18. Trier, Ø.D., Jain, A.K.: Goal-directed evaluation of binarization methods. IEEE Trans. PAMI **17**, 1191–1201 (1995)

Energy Harvesting from Dairy and Hospital Wastewater Using Microbial Fuel Cell (MFC)

C. Shakunthala$^{(\boxtimes)}$ and Surekha Manoj

Vidya Vikas Institute of Engineering and Technology, Mysuru, India
shakunthala.mys@gmail.com,
surekhamanoj.vviet@gmail.com

Abstract. The present experiment is focused on the production of electricity using Double Chambered Microbial Fuel Cell (DCMFC). In the present experimental study Distillery and Sugar industrial wastewater was used as substrate, which were collected from Mymul Nandini Dairy, Nazarbad, Mysore, India and JSS Hospital, Near Agrahara, Mysore, India, were used. The Biochemical Oxygen Demand (BOD) of JSS Hospital wastewater was found to be 62.64 mg/L. The maximum voltage and current obtained from JSS Hospital wastewater were found to be 825 mV and 2.1 mA, respectively, by using Aluminium as electrode. Experiments were also conducted using Dairy wastewater which yielded a maximum voltage of 1000 mV and current of 4.5 mA by using Copper electrode. Higher BOD and COD (Chemical Oxygen Demand) concentrations of wastewaters was found to yield greater current production, on the other hand dilution of wastewater resulted in lower current production.

Keywords: Microbial Fuel Cell · Micro-organisms · Electricity
Substrate · Organic matter

1 Introduction

In the recent decades, consumption of energy within the world has an emerging trend [1]. Energy sources are classified into two types' renewable sources and non-renewable sources, in which non-renewable sources of energy, which include an enormous portion of energy consumption, could be categorized into two major classifications: nuclear and fossil energy [2]. Fossil fuels negatively influence the nature owing to the emission of carbon dioxide (CO_2). Consumption of fossil fuels has severely imperilled human life through its drastic aftermaths, such as global warming and atmospheric pollution. However, miscellaneous countries around the world have made significant efforts to find a piece of convincing solution for energy crisis by revolving the eyes into renewable energy sources such as solar energy, energy produced from wind and water. As an outcome of these efforts, one of the future alternative energy sources is Fuel Cell (FC), which generates energy using high value metal catalysts [1]. The prime advantage of this types of energy generator lead to-no emissions of environmental polluting gases (such as SO_2, NO_2, CO_2 and CO), higher efficiency, no existence of transportable parts [2]. MFCs are the kind of FCs uses an active microorganism as a biocatalyst in an anaerobic anode compartment for production of bioelectricity. A MFC transforms organic energy,

existing in a bio-convertible substrate directly into Electrical energy. To reach this, bacteria are used as a catalytic agent and to convert substrate into electrons and protons which will pass through on Peripheral Circuit and Proton Exchange Membrane (PEM), separately [3]. MFCs are used to generate less excess sludge as compared to the aerobic treatment process [3]. Sediment fuel cells are being developed to monitor eco-friendly systems such as bays, rivers, and oceans [4]. MFCs are special types of bio-fuel cells, producing electricity through consuming microorganisms. Liu has conducted experiments on Single Chamber Microbial Fuel Cells (SC-MFC) using an enhanced cathode assembly [5]. Patra has conducted experiments using low cost single chambered fuel cells, which return electricity and produce domestic water from wastewater. Banik have carried out studies on the development of SC-MFC using wastewaters under anaerobic condition with two sets of electrode configurations [7]. Logan has demonstrated the significance of thickness of anode on electrical energy generation in a SC-MFC with different electrode distances with 28 ml capacity [6, 8, 9].

2 Materials and Methods

Microbial Fuel Cells (MFC) have been defined as bioreactors, which transform the energy stored in biological compounds present in wastewater into electric power through catalytic activity of micro-organisms under anaerobic conditions. Many researchers have worked on the generation of power from wastewater using different electrodes under varied operational conditions. In this work, laboratory experiments were conducted to generate power from sugar and distillery industrial wastewater by using Aluminium and Copper electrodes in DC-MFC. The detailed discussion on the materials used and methods adopted have been discussed in the following sections.

2.1 Materials

Fabrication of MFC necessitates the use of materials that are often utilized in the construction of conventional fuel cells. It is important to note that, microorganisms are the vital components of the MFC. The materials which have been used in the present MFC study have been discussed in the following sub-sections.

In this study, Double Chamber MFCs were fabricated with acrylic Plexi glass was used to conduct the laboratory experiments. Electrodes such as, Aluminium and Copper electrodes were used. In this experimental study Distillery and Sugar industrial wastewater were used as substrate, which were collected from Mymul Nandini Dairy, Nazarbad, Mysore, India and JSS Hospital, Near Agrahara, Mysore, India. For the preparation of the Salt-bridge, Potassium Chloride and Agar media were used. Mediators such as, Potassium Permanganate, Phosphate buffer and Potassium Ferri cyanide of laboratory grade was used to enhance the electron transfer. For conducting the laboratory experiments, chemicals such as, sulphuric acid, hydrochloric acid, was used.

2.2 Methods

Laboratory experiments were planned to conduct under various conditions such as: different electrode materials, MFC module and different wastewater. A detailed discussion

on the wastewater analysis as well as the experimental procedure on the generation of power from industrial wastewater has been discussed in the following sub-sections.

The Double Chamber Microbial Fuel Cell (DC-MFC) reactors used in the present study were made of acrylic Plexi glass. Laboratory experiments were conducted using DC-MFC reactor dimensioning $(10 \times 10 \times 20)$ cm with a working volume of 2 L capacity. Plexi glass DC-MFC reactors were used because of its high durability and stability compared to normal glass material. Usage of plastic material has been avoided in the fabrication of DC- MFC reactor, as it is known to be reactive with chemicals present in the reactor. In the present study, different types of electrodes were used, these include: Aluminium electrodes and Copper electrodes. The dimensions of the electrodes were $(8 \text{ (height)} \times 4 \text{ (width)} \times 0.5 \text{ (thickness))}$ cm. Both the electrodes were kept in the DC-MFC reactor at a distance of 5 cm from each other. In order to have effectiveness of the electrodes, electrodes were dipped in distilled water over a period of 24 h before start-up of the experiment. Membrane is defined as a layer of material, which assists as a selective barrier wall between two ways, which is resistant to specific particles, particles and substances. Synthetic membranes include anion, cation and ultrafiltration membranes. These membranes are most costly and are susceptible to clogging and hence, agar-salt bridge was taken as proton exchange membrane (PEM).

The salt bridge was prepared by taking 9.6 gm of KCl and 1.3 gm of agar with 25 ml of distiller water and allowed to boil for about 2 min. The hot solution was then poured into the hollow tube having 5 cm length and internal 1.5 cm diameter. The setup was further allowed to cool for nearly 2 h, and the salt bridge was thus ready for use.

Micro-organisms are generally used in MFC to degrade the wastewater in the absence of oxygen and to transfer the electron to the anode. The ability of many microorganisms possesses to relocation of the electrons resulting from the metabolism of organic matters to the anode. These microorganisms are steady and have great coulombic efficiency. Shewanella putrefaciens [1], Geobacteraceae sulfurreducens [4] and Geobacter metallireducens [7] bacterial strains have showed effective growth on the anode surface and transmission of electrons directly to electrode across the membrane. These types of microbial cultures are costly and the growth and maintenance of these cultures are also found to be very difficult.

Hence, in this work culture obtained from Dairy and Hospital generally contains certain anaerobic microbes. An ideal catalyst can be able to cross the cell membrane, be able to grasp electrons from the electron carrier of the electron track chains, should possess a high electrode reaction rate, good solubility in the anolyte, non-biodegradable, non-toxic to micro-organisms and should be of low cost. In this study, Phosphate buffer, Potassium Ferri-cyanide and Potassium Permanganate were used as catalysts to maintain pH of wastewater between 7.0 and 7.7. Phosphate buffer was prepared by taking 80.2 ml of Potassium Phosphate di-basic and 19.8 ml of Potassium Phosphate mono-basic, which was diluted with 900 ml of distilled water, and 0.1 mg Potassium Ferri-cyanide solution was prepared by taking 3.292 gm of Potassium Ferri-cyanide ($K3[Fe(CN)6]$) crystals in 100 ml distilled water. On the other hand, 600 μg Potassium Permanganate solutions was prepared by taking 0.1 gm of Potassium Permanganate crystals in 1 litre of distilled water. In order to increase the electron transfer efficiency Phosphate buffer and Potassium Permanganate was used.

In the present experimental study Distillery and Sugar industrial wastewater were used as substrate, which were collected from Mymul Nandini Dairy, Nazarbad, Mysore, India and JSS Hospital, Near Agrahara, Mysore, India. Glucose has been added for the growth of micro-organisms present in the wastewater. The wastewater collected from both the industries was analysed for Nitrate, Phosphate, BOD, COD and pH concentrations.

2.3   Experimental Procedure

Three Double Chambered MFC units have been fabricated using acrylic Plexi-glass and used for conducting laboratory experiments with a working volume of 2 L capacity each. A 5 mm diameter hole has been bored in the middle of each top of the reactor and 30 cm long conductor has been inserted through the hole. Another hole is drilled on same horizontal level, on walls of both reactors facing each other, where a hollow tube of 5 cm long is inserted, which acts as a salt bridge. The experiment was conducted with different types of electrodes such as, Copper (Cu) and Aluminium (Al) electrodes. Agar Salt-bridge was used as Proton Exchange Membrane (PEM), which was prepared by taking 9.6 gm of KCl (Potassium Chloride) and 1.3 gm of agar medium. Anaerobic sludge of 10 gm was taken into the anodic section, which leads to have a better electron transfer to the cathode by degrading the wastewater. Wastewater samples were collected from Mymul Nandini Dairy and JSS Hospital, Mysore.

2.4   Analysis

DC-MFC experiments have been carried out in different phases over a period of 3–5 days. However, voltage and current have been measured at an interval of one hour for 1st day and every two hour during 2nd and 3rd days. The influent and effluent COD, BOD applications and pH were monitored as per Regular Methods for analysis of water and wastewater (Fig. 1).
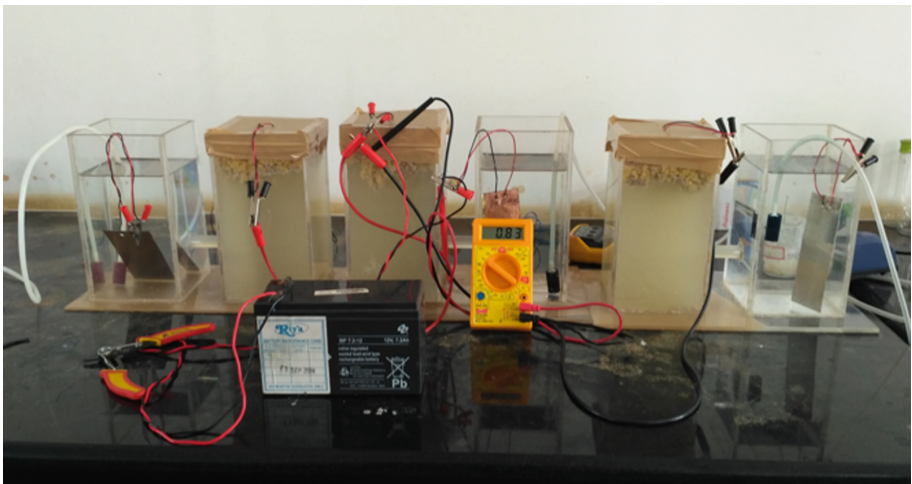
Fig. 1. Experimental setup of Double-Chambered Microbial Fuel Cell (DC-MFC)

3   Results and Discussion

In the degradation of Dairy wastewater, it was found that, Dairy wastewater using Copper electrode generated more voltage (~ 1000 mV) compared to Hospital wastewater (~ 800 mV) using Aluminium electrode. The reason might be due to, BOD and COD of the Dairy wastewater was found to be more when compared to Hospital wastewater, which leads to movement of more electrons. It was also observed that, at the initial stage of the experiment, voltage increases, then drops and finally again reduces and maintains constant. Figures 2, 3, 4 and 5 represent the variation of voltage, current, current density and power versus time for Dairy and Hospital wastewater in MFC reactor.
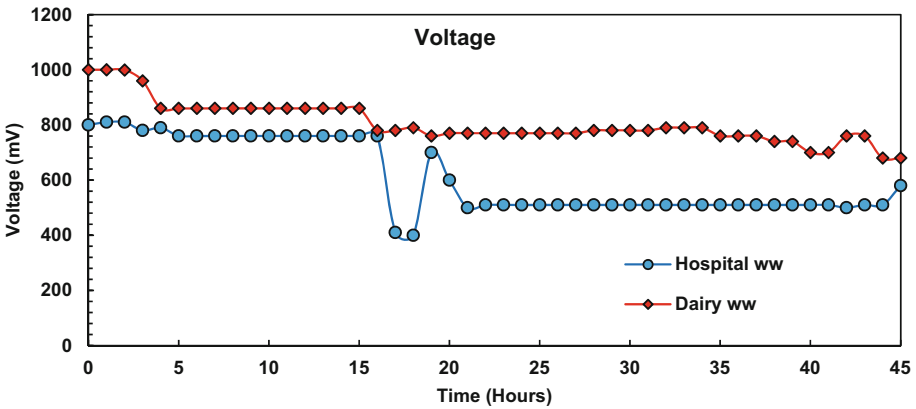
Fig. 2. Variation of voltage with time in double chambered microbial fuel cell.
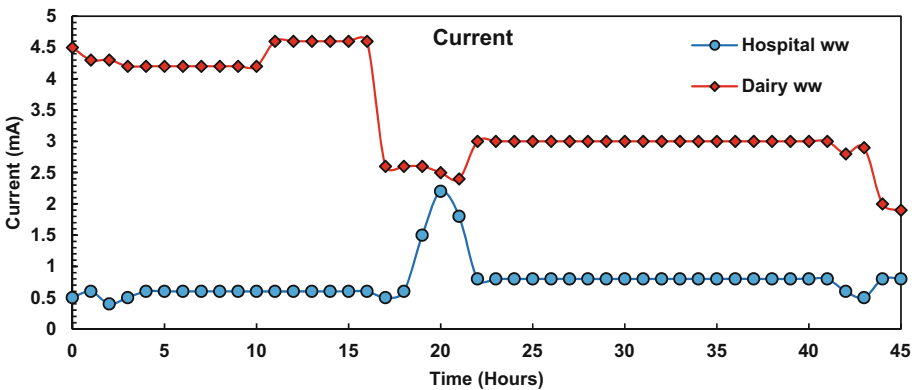
Fig. 3. Variation of current with time in double chambered microbial fuel cell.
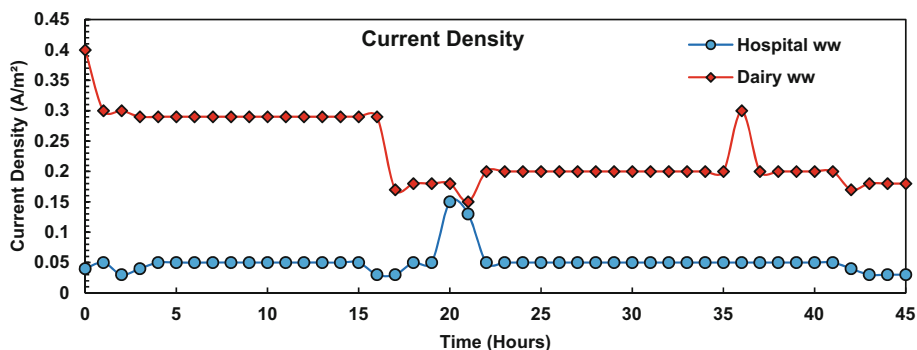
Fig. 4. Variation of current density with time in double chambered microbial fuel cell.
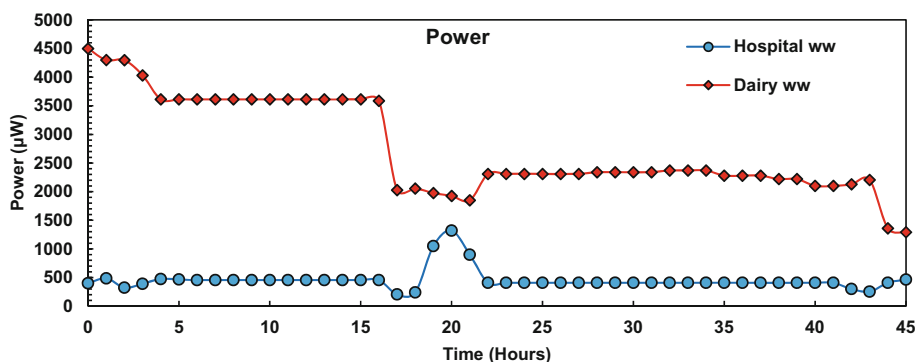
Fig. 5. Variation of power with time in double chambered microbial fuel cell.

From the Fig. 3, it was found that, Dairy wastewater using Copper electrode generated more current ($\sim$4.5 mA) when compared to Hospital wastewater using Aluminium electrode. This may be due to; Copper electrodes are more favourable to transfer electrons to the cathode than Aluminium electrodes. However, a maximum current density of approximately 0.4 A/m^2 (Fig. 4) was observed for Dairy wastewater, which was found to be greater than the current density (0.05 A/m^2) generated from Hospital wastewater using Aluminium electrode. It was also observed that, at the initial stage of the experiment, current density was found to be maximum and then drops and finally maintains the constant value. Power generation from the MFCs were also calculated (Fig. 5) and found that, Dairy wastewater using copper electrode generated more power ($\sim$4700 µW) compared to Hospital wastewater ($\sim$500 µW).

4   Conclusion

This paper represents the implementation of a generation of electricity that assists a utilization of wastewater like industrial as well as hospital wastewater, which contain high BOD and COD values. The voltage and current values obtained in DC-MFC using

Dairy wastewater was 1000 mV and 4.5 mA, respectively, which was found to be higher than the voltage and current values of 825 mV and 2.1 mA, respectively, obtained from Hospital wastewater. It was also found that, Copper electrodes were more effective in electron transfer and electricity generation when compared to Aluminium electrodes.

References

1. Kim, H.J., Park, H.S., Hyun, M.S., Chang, I.S., Kim, M., Kim, B.H.: A mediator-less microbial fuel cell using a metal reducing bacterium, Shewanella, putrefaciens. Enzyme Microb. Technol. **30**, 145–152 (2002)
2. Logan, B., Cheng, S., Watson, V., Estadt, G.: Graphite fiber brush anodes for increased power production in air-cathode microbial fuel cells. Environ. Sci. Technol. **41**, 3341–3346 (2007)
3. Bond, D.R., Lovley, D.R.: Electricity production by Geobacter sulfurreducens attached to electrodes. Appl. Environ. Microbiol. **69**, 1548–1555 (2003)
4. Fan, Y., Hu, H., Liu, H.: Sustainable power generation in microbial fuel cells using bicarbonate buffer and proton transfer mechanisms. Environ. Sci. Technol. **41**, 8154–8158 (2007)
5. Patra, A.: Low-cost, single-chambered microbial fuel cells for harvesting energy and cleansing wastewater. J. U. S. SJWP **10**, 72–85 (2008)
6. Banik, A., Jana, N.K., Maiti, B.R., Ghosh, T.K.: Development of microbial fuel cells and electrode designs with waste water anaerobes. Greener J. Biol. Sci. **2**, 013–019 (2012). ISSN 2276-7762
7. Ahn, Y., Logan, B.E.: Domestic wastewater treatment using multi-electrode continuous flow MFCs with a separator electrode assembly design. Appl. Microbiol. Biotechnol. **97**, 409–416 (2013)
8. Cheng, S., Liu, H., Logan, B.E.: Increased performance of single chamber microbial fuel cells using an improved cathode structure. Electrochem. Commun. **8**, 489–494 (2006)
9. Ahn, Y., Logan, B.E.: Altering anode thickness to improve power production in microbial fuel cells with different electrode distances. Am. Chem. Soc. **27**, 271–276 (2012)

Cognitive Computing in Video Analytics

Anomalous Event Detection in Videos Using Supervised Classifier

K. Seemanthini and S. S. Manjunath$^{(\boxtimes)}$

Dayananda Sagar Academy of Technology and Management, Bangalore, India
be.outstanding@gmail.com, mnj_ss2002@yahoo.co.in

Abstract. Observing and modeling human behavior and activity patterns for detecting anomalous events has gained more attention in recent years, especially in the video surveillance system. An anomalous event is an event that differs from the normal or usual, but not necessarily in an undesirable manner. The major challenge in detecting such events is the difficulty in creating models due to their unpredictability. Most digital video surveillance systems rely on human observation, which are naturally error prone. Hence, this work validates the rising demand of analysis of video surveillance system. The system being proposed here is of minimum requirements with a competitive computational power when compared to the existing ones.

The main objective of this research work is to build up a framework that recognizes small group of human and to detect the event in the video. A combination of feature extraction using Histogram of Oriented Gradient (HOG) and feature reduction with Principle Component Analysis (PCA) is proposed in this work. The knowledge base and video feed for test cases are classified using the Support Vector Machine (SVM) to categorize the event as either anomalous or not based on various parameters.

The experimental result demonstrates that this approach is able to detect anomalous events with a competitive success rate. The framework can be used to identify various events such as anomalous detection of events, counting people, fall detection, person identification, gender classification, human gait characterization etc.

Keywords: Foreground extraction · Human group extraction (HOE)
Visual saliency · Event detection · Classification

1 Introduction

1.1 Overview

Detecting human beings and identifying actions in videos of a surveillance system is gaining more importance due to its wide variety of applications in detecting anomalous events, counting the number of people in a dense crowd, identification of human, traffic safety and surveillance, sports analysis, gender classification, human gait characterization, fall detection for elderly people, etc.

Latest surveillance cameras are installed all around the world daily, as webcams, for surveillance and other purposes. Most of the digital video surveillance systems depend on human observers for detecting and identifying specific activities in the video scenes.

But, there are several limitations in the human capability to monitor simultaneous events in a video surveillance system. Hence, automated human event analysis in video surveillance system has become one of the most effective and attractive research topics in the area of computer vision and pattern recognition.

However, most existing multi-person tracking methods are still limited to special application scenarios. They require multi-camera input, scene specific knowledge, a static background, or depth information, or are not suitable for online processing.

Moreover, there may be both individual and gathering activity in the same scene, it is much harder to speak to and receive such situations. Punctuation models have been generally utilized as a part of the complex visual occasion acknowledgment lately. To apply sentence structure in models or occasion acknowledgment, normally low-level components are firstly removed from features and after that characterized to an arrangement of terminal images, i.e., visual occasion primitives. Hence the proposed work detects and identifies both individual and group events, hence it overcomes the drawback of existing system, and it is also based on cognitive linguistic method which uses unsupervised learning method.

Detecting human object is a difficult task from a machine vision perspective as it is motivated by a wide variety of possible appearance due to changing articulated pose, lighting, clothing and background, but prior knowledge on these limitations can improve the performance of detection. The proposed system detects and captures motion information of moving targets for accurate object classification. Unsupervised classifiers are used for learning method and labels are known, hence the instance such as kick, hug, punch, or any such features are extracted and events are detected. The classified object is being used for high level analysis.

1.2   Objectives

The main objective of this research work is to build up a framework that recognizes the small human group and to detect the event in the video. This framework is utilized for robotized little human gathering occasion discovery inside of social or open spot environment furthermore serves to recognize a fording wrongdoing, in places like Railway station, traffic, collages, office, etc.

1.3   Problem Statement

Detecting human beings and identifying actions in videos of a surveillance system is gaining more importance due to its wide variety of applications in detecting anomalous events, counting the number of people in a dense crowd, identification of human, traffic safety and surveillance, sports analysis, gender classification, human gait characterization, fall detection etc.

The proposed work is used to represent both individual and multiple individuals in an event, hence it overcomes the drawback of existing system, and it is based on cognitive linguistic method which uses unsupervised learning method.

The proposed system is able to identify the events automatically in the video surveillance system. Thus, it reduces the human interaction with the video surveillance system and reports the alerts as the events detected.

1.4   Proposed Methodology

The image is given as an input to the training database. The obtained RGB images are further preprocessed using mathematical morphological method to reduce noise and later converted to grayscale. Features are extracted using HOG descriptor and reduced using PCA. The resultant is stored in a file which is trained using SVM classification. On the other hand the testing dataset is converted to frames, preprocessed, and their absolute difference is evaluated to distinguish background from foreground. Further morphological operation takes place to reduce noise, followed by Feature extraction using HOG and PCA and classified using the SVM Classifier. The key techniques used are

1. Preprocessing using morphological operations.
2. Feature Extraction and Reduction using HOG and PCA.
3. Classification using SVM.

1.5   Applications

Automated anomaly detection has a wide variety of applications. It has huge potential in the field of video surveillance system. Even though video surveillance cameras are installed everywhere, the availability of human resources to monitor the footage is poor. Hence, an automated system will aid in overcoming such human errors. Events such as trespassing can be alerted immediately when an automated system is placed.

Detection of non-human objects in unexpected places aids in betterment of security measures. It helps in person counting in densely crowded places such as those shown in Fig. 1. An automated anomaly detection system may aid in fall detection in the homes

(a) Stadium

(b) Shopping Mall

(c) Airport

(d) University Pass

Fig. 1. Analysis of different scenario through anomalous event detection.

of the elderly. Traffic safety is the major applications of anomaly detection. Detection of speeding vehicles or reporting drivers breaking the law immediately can be achieved using an anomaly detection system. Another growing field is in sports analysis where an automated system might alert the referee or judge in case of actions which may otherwise be overlooked.

2   Literature Survey

Zhaozhuo Xu [1] has introduces a Human-Object Interaction model, and are able to establish methods and systems to recognize events that are dangerous. In this approach, the process of event understanding is based on identifying dangerous objects in possible areas predicted by human body parts. The accuracy of dangerous human events understanding is improved when human body parts estimation is combined with objects detection.

Dongping Zhang [2] presents an approach to identify group level crowds and detect any abnormal activities in them. It incorporates particle motion information calculated using a set of sample images with long trajectories and other properties, into identifying small human crowds in foreground images while in motion. Science of Human behaviour is studied and employed to detect normal and abnormal activity. Attributes such as orientation, velocity and crowd size are used to distinguish between normal and abnormal behaviour.

MyoThida [3] has presented a review of crowd video analysis in this paper. Automation of surveillance has become in crowded places such as shopping malls, railway stations and airports. Providing intelligent solutions to these places is of high priority to computer researchers. The paper provides a thorough review of the existing automation techniques for analyzing complex and crowded scenes.

The merits and demerits of the various modern methods are discussed in detail. Tracking individuals in a crowd is a major topic. It is a highly complex task due to interactions with various other objects present in the crowd.

M. Sivarathinabala [4] proposes an intelligent video surveillance system, which can be remotely monitored and alerts the user in a situation that the system may interpret as an anomaly. The main focus is on monitoring a single person in situations such as a burglary.

A live video is captured and reduced to images. The images undergo preprocessing. Human behaviour analysis plays an important role to detect any anomalous human activity. This is done by comparing existing sample templates with the processed image. If found, the image is stored in the system and an alert is sent as specified by the user either to MMS, SMS or email. The live video is then compressed and a key frame is specified to directly retrieve the required part of the video.

This paper concludes by providing an automated method for surveillance that not only identifies an anomaly but also triggers an alert to the user. It helps in retrieval of

the suspected video by holding key frame values and help in extracting of images of individuals before and after the incident.

Manoranjan Paul [5] throws light on the need for accurately detecting anomalies in videos and its applications in surveillance technologies. Detecting human beings and their actions accurately in a video has various applications such as person identification, fall detection for elderly people, event classification and gender classification.

The authors use the benchmarks set by few existing datasets for comparison and providing their assessment. An intelligent system can capture and detect moving objects in a video. In this study the authors focus on detecting only human beings in general. This in itself is a complex task due to the number of various attributes each person may have such as, clothing, pose, lighting and background.

Detecting objects in a surveillance video is a challenging task due to the low resolution of the video. This paper discusses different methods of object identification and object classification. The various benchmarks are discussed and the applications of human detection in surveillance videos are reviewed.

C. Stauffer [6] proposed a computer vision algorithm for detecting or analyzing the motion of people in crowds. Computer vision algorithm divides background in regions and track the crowds and analyses every movement of people.

D. Ryan [7] develops a scene independent approach that can count the no of people in the crowd. A scene independent counting system can easily be deployed at different place. The counting is been done using a global scaling factor to relate crowd size from one scene to another.

Condition of providing the right heuristic ranking to the individuals, to avoid confusing them with one another. Hence, achieve robustness by finding optimal trajectories over many frames while avoiding the combinatorial explosion that would result from simultaneously dealing with all individuals.

3 System Design

3.1 Overview

The following image represents the system architecture which is further broken down into a clear flowchart in the later segments. The architecture as described in the image consists of testing phase and training phase. In the training phase, the training video is added into the knowledge base after preprocessing, feature extraction and reduction. This dataset is classified into either normal or abnormal event using SVM Classifier. The testing set of the video, which goes through the same operations are classified as abnormal or normal by comparing it to the sample frames of the training videos that are classified already. The output classifies each frame to be either "Normal" or "Abnormal" (Fig. 2).

Training Phase   **Testing Phase**
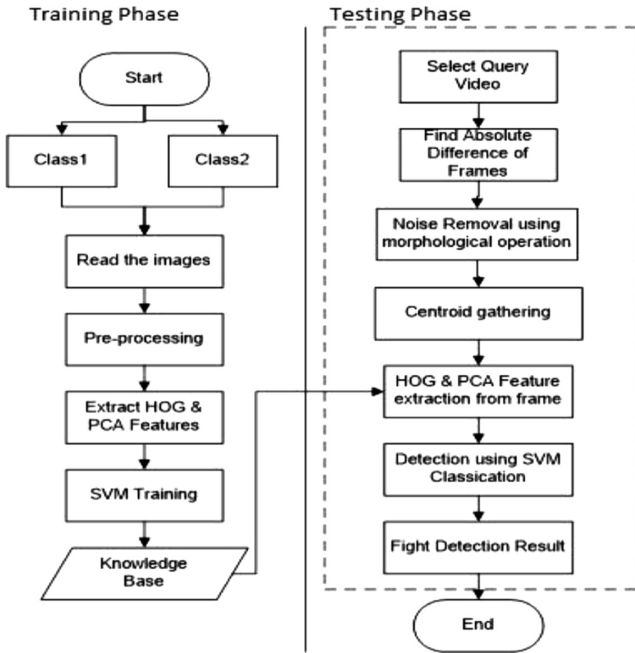
Fig. 2. Overall flow chart of the proposed system

3.2 Preprocessing

Initially the given video is converted into frames. The converted frames are used for further processing. In pre-processing unnecessary noise in the frames are eliminated using morphological operations. In order to get a more accurate difference between the background and the foreground, the image needs to have lesser noise [7].

3.3 Feature Extraction

The main purpose of feature extraction is to extract the image component and to separate the foreground from the background through HOG feature extraction and further reduce the obtained attributes by the method of PCA. This helps in providing faster time for analysis due to a better predictive model, with many similar attributes reduced to a single attribute (Figs. 3, 4 and 5).

3.3.1 Histogram of Oriented Gradients

Histogram of oriented gradients is a feature descriptor used in image processing for detecting objects. This technique counts occurrences of gradient orientation in localized portions of an image. The HOG descriptor is most popularly used for detecting humans in images. The flow diagram of the HOG descriptor is given below (Fig. 6).
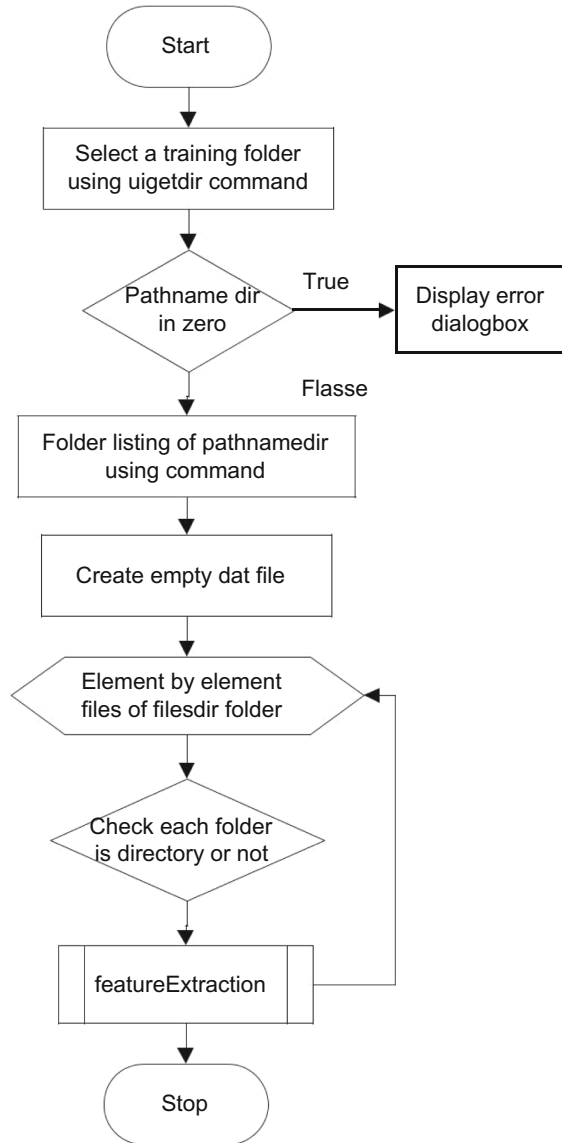
Fig. 3. Feature collection

3.3.2 Principal Component Analysis

PCA is mathematically defined [8] as an orthogonal linear transform that transforms the data to a new coordinated system such that the greatest variance by some projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on.
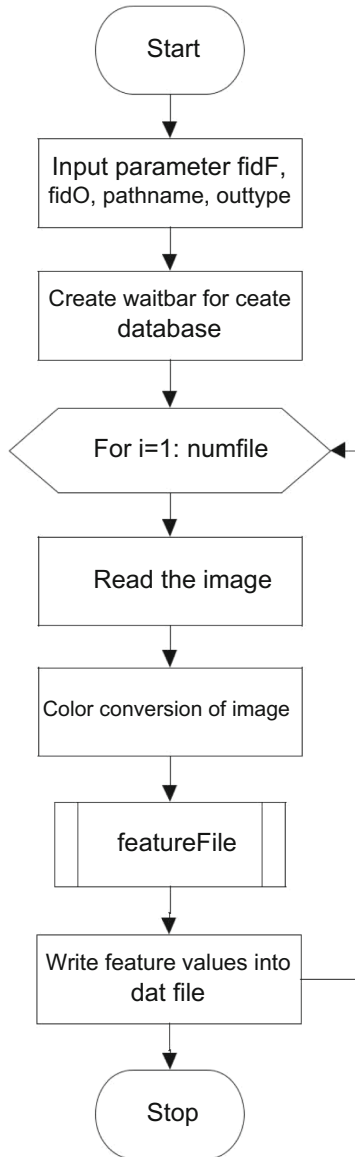
Fig. 4. Feature extraction

PCA applied to image processing in this particular project to reduce the orientation values of HOG that are stored. This helps in improved performance and faster detection and analysis of objects in the frames.
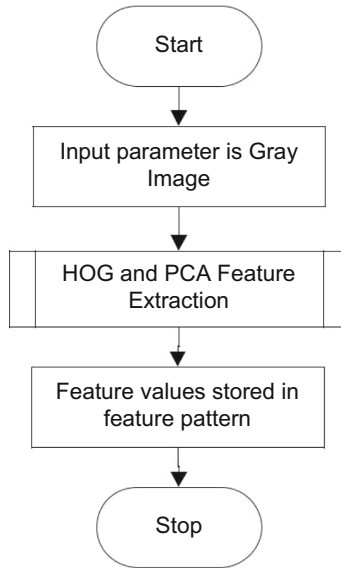
Fig. 5. FeatureFile subroutine that performs HOG and PCA feature extraction/reduction
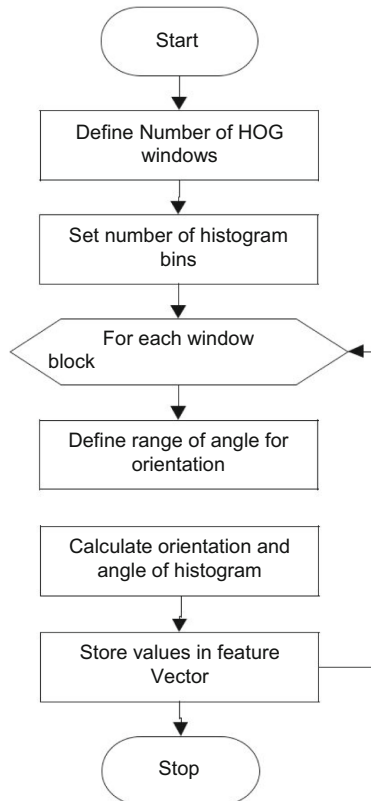
Fig. 6. HOG feature extraction

3.4   Support Vector Machine Classifier

Support Vector Machines (SVM) is the most popularly used classification method [11]. It has a wide variety of applications such as classification of text [8], recognizing the facial expression [10], in the analysis of genes [7] and many others. SVM is one of the technique for constructing a linear classifier, which produces a classifiers based on the theoretical foundation [12].

3.4.1   SVM Classification Using "KERNAL TRICKS"

Kernel techniques is used in SVM to solve linearly inseparable problems which transforms data to a high dimensional space. But training and testing large data sets consumes more time. Hence, we can train and test the large data sets using linear SVM without kernels.

The Experimental results proves that the proposed method is beneficial for large-scale data sets. Hence the proposed method can be successfully applied to natural language processing (NLP) applications [6].

3.4.2   Radial Basis Kernel Function for SVM Classification of Images

Radial basis kernel function (RBF) is the most commonly used kernel function in machine learning. RBF kernel is used for various kernelized learning algorithms. In particular, it is most commonly used SVM classification [7].

Consider two samples $\mathbf{x}$ and $\mathbf{x}'$, which represents a feature vectors in some *input space*, The RBF kernel is defined as

$$K(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right)$$

$\|\mathbf{x} - \mathbf{x}'\|^2$ indicates the squared Euclidean distance between the two feature vectors.

σ is a free parameter. An equivalent, definition involves a parameter $\gamma = \frac{1}{2\sigma^2}$:

$$K(\mathbf{x}, \mathbf{x}') = \exp(-\gamma\|\mathbf{x} - \mathbf{x}'\|^2)$$

As shown in Fig. 7, we use SVM training loading the extracted features which are stored in features.dat. The output frames loaded in output.dat is compared to the trained features that are stored in the knowledge base.

3.5   Data Flow

The following image represents the data flow diagram of the proposed work. The architecture as described in the image consists of a testing phase and a training phase.

In the training phase, the training video is preprocessed, features are extracted and added into the knowledge base. This dataset is classified into either normal or abnormal
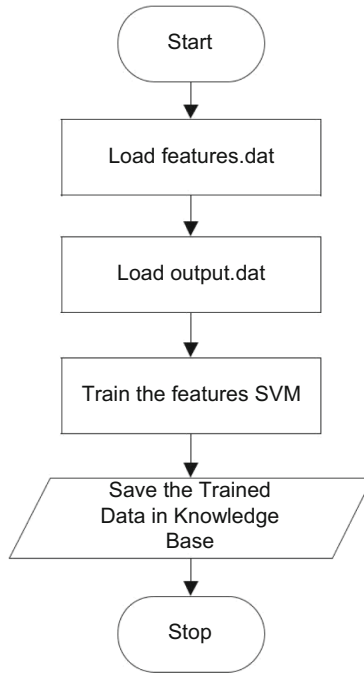
Fig. 7. SVM training

event using SVM Classifier. Even in the testing phase, the video is preprossed and the features are extracted. Later, The SVM classifier, classifies each frame to be either "Normal" or "Abnormal".

The Fig. 8 shows the two phase of data flow diagram. Initially, the image is read in RGB Frame format. During preprocessing, the frames are converted to gray scale format. This resultant format is used for extracting features using HOG and PCA methods. The extracted features are trained using SVM and the trained data is stored in the knowledge base and used for classifying the data under the SVM.

In the testing phase, the input video is broken down into frames and preprossed using morphological operations. Then the foreground objects are extracted by performing the background substraction. This is done by finding the Absolute difference of the frames. Later Feature extraction takes place on the noise free images. The feature extraction procedures include HOG & PCA methods. These features fall under a classification trained under the SVM, as explained earlier.

In the current project, the SVM classifies data as either Normal or Anomalous (Abnormal). The detected region and recognition result is displayed along with the frame.

Fig. 8. Data flow diagram

4 Result Analysis

4.1 Discussion

The dataset used in this project was shot in a Canon D750 camera at 55 mm focal length. The video resolution is adjusted to 380 × 240 pixels at a frame rate of 15. Three different scenarios are evaluated in the datasets. The first two datasets were used to depict anomalous event such as punching and kicking. This also depicts normal scenarios of handshake and hug. The third dataset is of a typical fall detection even that may occur at any old age homes.

4.2 Performance of Our System

The system, when tested with the datasets mentioned above, an error rate of **0.27** was obtained. The error rate was calculated using the formula:

Error Rate = No of False Negatives/(No of False Negatives + No of True Positives)

Here False Negatives refer to the anomalous events that were not identified. True Positives are the anomalous events that were identified correctly.

Table 1 provides a detailed overview of the performance. The SVM classifier was successfully able to classify most of the event as normal or abnormal. The dataset was provided as frames to the classifier with specification for classification process (Fig. 9).

Table 1. Comparison of error rate for the three datasets

	Dataset 1	Dataset 2	Dataset 3	Total
True positives	33	16	20	69
False positives	1	0	7	8
False negatives	8	5	12	25
Err. rate	0.20	0.24	0.38	**0.27**
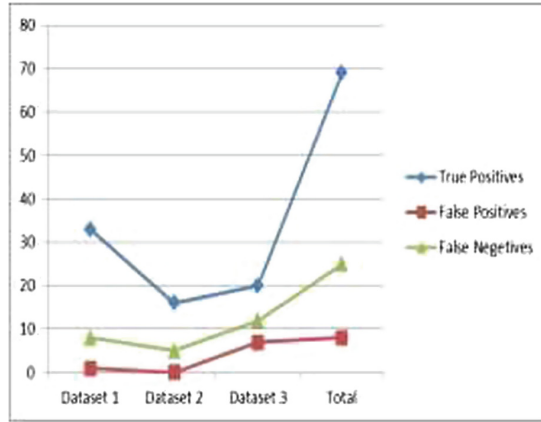
Fig. 9. Performance comparisons for three datasets

4.3   Discussion of Result

The dataset 3 produced the highest number of false negatives because the person of interest's movement and placement towards the camera made it difficult for the classifier to identify the event. The feature extraction was also not optimal due the constant changes the objects of the environment.

It is observed that when this dataset was removed from evaluation the error rate was reduced to only **0.21**. This shows the importance of a static background for our system, which is also a drawback and suggestion for future enhancement.

Table depicts the performance when dataset 3 is removed (Table 2).

Table 2. Comparison of error rate without the challenging video

	Dataset 1	Dataset 2	Total
True positive	33	16	49
False positives	1	0	1
False negatives	8	5	13
Err. rate	0.20	0.24	**0.21**
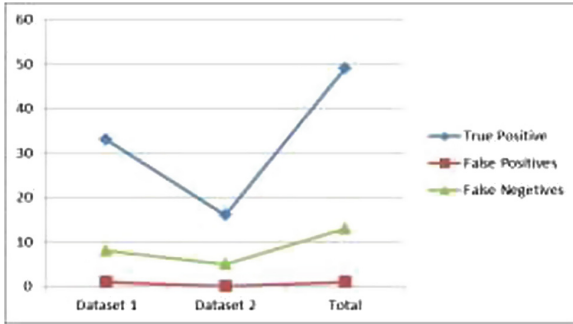
Fig. 10. Performance comparisons without the challenging video.

Fig. 11. Dataset 1: shows sequence of normal and anomalous events

The performance of the system is calculated, with and without the challenging video in Figs. 11 and 12. Figure 13 provides a detailed description of the three datasets (Figs. 10, 14, 15, 16 and 17).
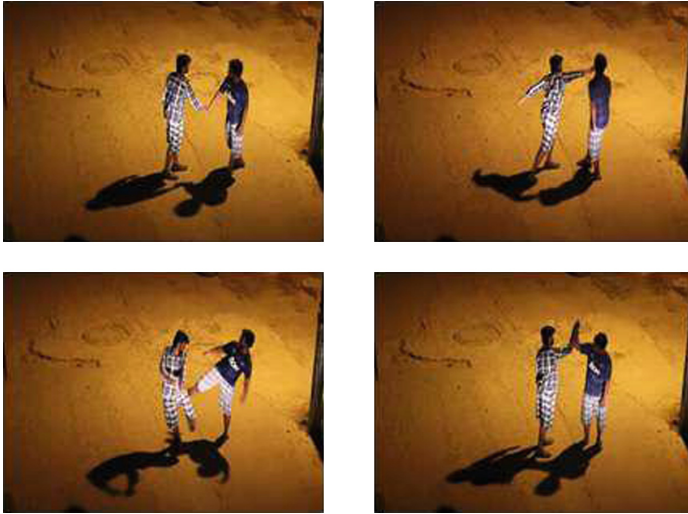
Fig. 12. Dataset 2: handshake, punch, kick, high five events

Fig. 13. Dataset 3: false negative event followed by fall detection event

4.4   Output
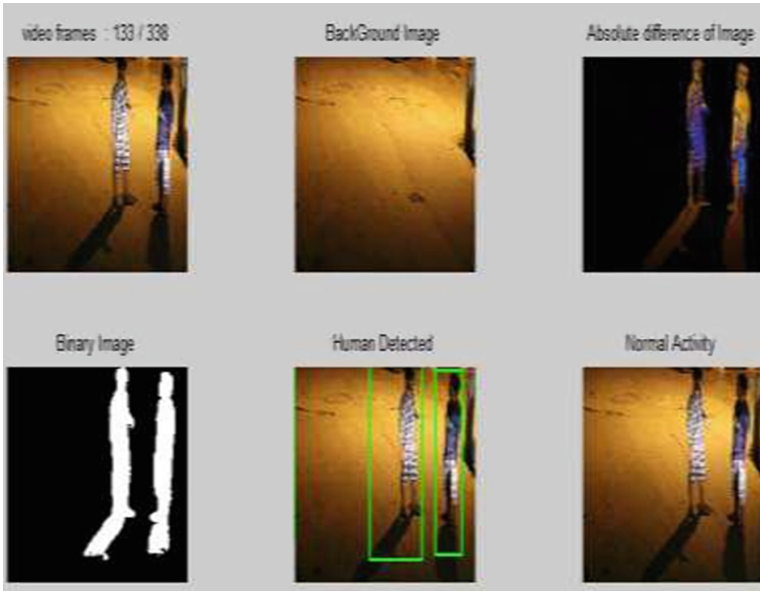
Fig. 14. Normal event recognition as seen in dataset 1

Fig. 15. Anomalous (or abnormal) event recognition as seen in dataset 1
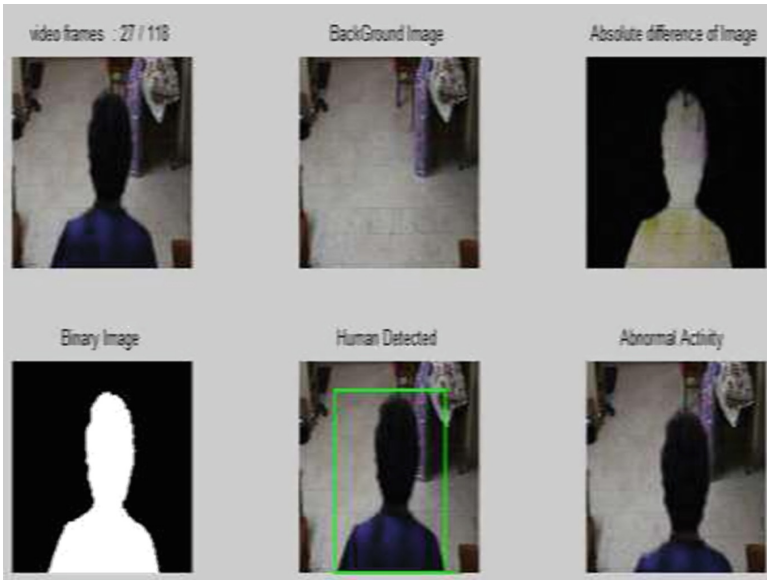
Fig. 16. False negative events in dataset 3

Fig. 17. Fall detection in dataset 3

5   Conclusion

The automated human event analysis in video surveillance system has become one of the most effective and attractive research topics in the area of computer vision and pattern recognition. The increasing computational power, provides a great environment for improving the existing systems.

The proposed work has provided the satisfactory results as expected. The dataset used in this implementation was taken and designed for a static camera. An error rate of 0.27 was achieved when tested with the given datasets. A better error rate of 0.21 was achieved when a challenging video was removed from the dataset. This method can be further improved and implemented for real time surveillance systems. Hence the study for anomalous event detection can grow further.

References

1. Andriluka, M., Roth, S., Schiele, B.: People tracking by detection and people detection by tracking. In: IEEE Computer Vision and Pattern Recognition (2008)
2. Leibe, B., Schindler, K., Cornelis, N., Gool, L.V.: Coupled object detection and tracking from the static cameras and moving vehicles. IEEE Trans. Pattern Anal. Mach. Intell. 30(10), 1683–1698 (2008)
3. Li, Y., Huang, C., Nevatia, R.: Learning to associate: HybridBoosted multi-target tracker for crowded scene. In: IEEE Computer Vision and Pattern Recognition (2009)
4. Fradkin, D., Muchnik, I.: Support vector machines for classification. DIMACS Series in Discrete Mathematics and Theoretical Computer Science
5. Michel, P., Kaliouby, R.E.: Real time facial expression recognition in video using support vector machines. In: Proceedings of ICMI 2003, pp. 258–264 (2003)
6. https://www.quora.com/How-does-one-decide-on-which-kernel-to-choose-for-an-SVM-RBF-vs-linear-vs-poly-kernel answer by Charles H Martin, followed up on https://charlesmartin14.wordpress.com/2012/02/06/kernels_part_1/
7. Vapnik, V.: The Nature of Statistical Learning Theory, 2nd edn. Springer, Heidelberg (2013). https://doi.org/10.1007/978-1-4757-3264-1
8. Suriani, N.S., Hussain, A., Zulkifley, M.A.: Sudden event recognition: a survey. Sensors 13(3), 9966–9998 (2008)
9. Lin, W., Sun, M.-T., Poovendran, R., Zang, Z.: Group event detection for video surveillance. In: IEEE, pp. 2830–2833 (2009)
10. Candamo, J., Shreve, M., Goldgof, D.B.: Understanding transit scenes: a survey of human behavior-recognition algorithms. IEEE Trans. Intell. Transp. Syst. 11, 206–224 (2010)
11. Zaidenberg, S., Boulay, B., Garate, C., Chau, D.P.: Group interaction and group tracking for video-surveillance in underground railway stations. In: International Workshop on Behavior Analysis and Video Understanding (2011)
12. Zhang, Y., Ge, W., Chang, M.-C., Liu, X.: Group context learning for event recognition. In: Applications of Computer Vision. IEEE, pp. 249–255 (2012)
13. Zhang, D., Chen, F., Tong, C.: Particle motion based abnormal event detection in group-level crowd. J. Converg. Inf. Technol. 7(14) (2012)
14. Arbat, S., Sinha, S.K., Shikha, B.K.: Event detection in broadcast soccer video by detecting replays. Int. J. Sci. Technol. Res. 3(5), 282–285 (2014)

15. Tran, D., Yuan, J., Forsyth, D.: Video event detection: from subvolume localization to spatiotemporal path search. IEEE Trans. Pattern Anal. Mach. Intell. **36**(2), 404–416 (2014)
16. Kumar, A.N., Suresh Kumar, C.: Abnormal crowd detection and tracking in surveillance video sequence. Int. J. Adv. Res. Comput. Commun. Eng. **3**(9) (2014)
17. Pooka, N.S.: Suspicious group event detection for outdoor environment. Int. J. Mod. Trends Eng. Res. **0X**(0Y) (2015)
18. Berclaz, J., Fleuret, F., Fua, P.: Robust people tracking with global trajectory optimization. In: IEEE Computer Vision and Pattern Recognition (2006)
19. Chang, Y.-W., Hsieh, C.-J., Chang, K.-W., Ringgaard, M., Lin, C.-J.: Training and testing low-degree polynomial data mappings via linear SVM. J. Mach. Learn. Res. **11**, 1471–1490 (2010)
20. Guyon, I., Weston, J., Barnhill, S., Vapnik, V.: Gene selection for cancer classification using support vector machines. Mach. Learn. **46**(1–3), 389–422 (2002)
21. Jolliffe, I.T.: Principal Component Analysis. Springer Series in Statistics, 2nd edn. Springer, New York (2002). https://doi.org/10.1007/b98835. XXIX, 487 p. 28 illus. ISBN 978-0-387-95442-4
22. Joachims, T.: Text categorization with support vector machines: learning with many relevant features. In: Nédellec, C., Rouveirol, C. (eds.) ECML 1998. LNCS, vol. 1398, pp. 137–142. Springer, Heidelberg (1998). https://doi.org/10.1007/BFb0026683
23. Paul, M., Haque, S.M.E., Chakraborty, S.: Human detection in surveillance videos and its applications - a review. J. Adv. Sig. Process. **2013**, 176 (2013)
24. Sivarathinabala, M., Abirami, S.: An intelligent video surveillance framework for remote montioring. Int. J. Eng. Sci. Innov. Technol. IJESIT **2**(2), 297–301 (2013)
25. Thida, M., Yong, Y.L., Climent-Pérez, P., Eng, H.-l., Remagnino, P.: A literature review on video analytics of crowded scenes. In: Atrey, P.K., Kankanhalli, M.S., Cavallaro, A. (eds.) Intelligent Multimedia Surveillance, pp. 17–36. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41512-8_2

Classification and Clustering of Infected Leaf Plant Using K-Means Algorithm

Prathyakshini[(✉)] , Akshaya , and C. V. Aravinda

NMAM Institute of Technology, Nitte, Karkala, India
prathyakshini@nitte.edu.in

Abstract. Identifying the plant leaf images will play a key role in preventing great loss in the yield and quantity of production in farm. Study of leaf disease means the scientific study of disease in plant leaves which is caused by infectious organisms and physiological conditions. Keeping track of health and identifying of disease of plant leaves in early stage is very crucial in agriculture. Identifying plant leaf disease manually is very difficult as it needs lot of work, only the expertise will be able to identify the plant disease. So we mainly prefer image processing so as to detect leaf disease. Detection of disease will include several steps such as image-acquisition, image-preprocessing, image-segmentation, feature-extraction and also categorization.

Keywords: Image acquisition · Image segmentation · Feature extraction

1 Introduction

Agriculture plays a significant role in Indian economy. A large amount of employment opportunities are offered by agriculture. Farmers select different varieties of crops. Plant diseases might cause a large amount of loss in agricultural products. There should be continuous monitoring of plant health so that crop yield could be increased. Previously the identification of diseased plants was done manually, which needs lot of time. So in order to increase the agricultural growth, there should be an effective method to identify the disease on plants. Diseases often occur on different parts of plants such as stem, leaves and fruit. Using image processing techniques, disease on the leaf can be recognized also it can be categorized under the common disease types [1].

2 Literature Survey

Zang et al. [3] proposed a method of fuzzy feature selection based on the fuzzy-curves and surfaces for disease of cotton leaves. There are 2 steps in which the research is carried out. Initially the very small set-of significant features is taken out from the original features. And the false features are removed making use of fuzzy curves. In the next step the feature which is dependent on the important features are segregated. The technique has high classification success rate.

Meunkaewjinda et al. [4] proposed grape leaf disease which is mainly consists of three main parts. Initially from the complex background the input grape leaf color extraction is done. The grape leaf disease color extraction is done in the next step. In order to recognize the colors of grape leaf analysis of back-propagation neural network is used. For the segmentation purpose MSOFM and GA and for classification purpose SVM is used [2]. Finally to classify and detect the type of grape leaf disease SVM is applied. The classification is based on the three classes such as rust disease, Scab disease, and no disease. This method gives good result even if there is a limitation of picking ambiguous color pixels from the background image.

Using k-means clustering algorithm the problem in the leaf is found. Segmentation is done on the desired set of images and values are extracted [5]. K-Means clustering method gives accurate result steps in disease detection.

3 Steps in Leaf Disease Detection

Step 1: Image Acquisition

The leaf images are captured through the camera. Basically the plant leaf images that are considered are color images, which is in RGB that is Red, Green and Blue.

Step 2: Image Pre-processing

The noise is the image can be removed in this particular step. In order to obtain the concerned region of the image, image clipping can be used. Also the image is more enhanced and displayed such that the image becomes clearer. This contrasted image is now clearer with respect to the affected area when compared with the original image. The RGB images can be converted in to gray scale image using the Eq. (1)

$$f(x) = 0.2989*R + 0.5870*G + 0.144.*B \tag{1}$$

In order to enhance the leaf images, the intensities of the images are distributed by applying histogram equalization.

Step 3: Image Segmentation

In segmentation, image is divided into different parts of same features .In order to do this different techniques such as k-means clustering, otsu' method, converting RGB image into HIS model can be used which converts RGB image into HIS model.

(a) Segmentation using Boundary and spot detection algorithm: In order to segment the image, the RGB image can be transferred to HIS model. In order to identify the infected regions on the leaf, boundary detection method is used. For this, 8 connectivity of pixels are considered and boundary detection algorithm is being adapted [6].

(b) Segmentation using K-means clustering: This algorithm is used in order to differentiate the objects based on the N set of clusters [8]. The classification is made by considering the mean distance in the clusters.

Based on certain heuristic or randomly just take a cluster and initialize the same. Assign all the pixels to the cluster such that distance between pixel distances to the

cluster is minimized. Calculate the average of all the pixels in the cluster to get the cluster centre again step 2 and 3 is repeated till the convergence is reached.

The symptoms of the disease are identified in the image by changes in color of the leaf. Wherever the color is green is identified as healthy part where as other than green is considered as the disease.

Step 4: Feature extraction: In order to identify an object, feature extraction plays a very important role. Features are extracted based on some feature dimension. Feature dimension is measured using the Euclidean distance. Dissimilarity is seen if the distance between two points is very large. Using the Eq. (2) Euclidean distance is measured which is the difference between the points is taken which shows whether there is similarity or not.

$$d(Xi, Xj) = |Xi - Xj| / \sqrt{M(i) - M(j)} \tag{2}$$

Where $d(Xi, Xj)$ is the Euclidean distance, $M(i)$ is the mean distance of Xi and Xj.

Once the features are extracted, it is trained using neural network. The features such as texture, morphology, color etc. are used in plant disease detection. It is very difficult part when it comes to identify the diseased parts in the leaf image, as the region of affected area differs from one diseased leaf image to another. Texture tells us how the color is spread across the leaf image, hardness as well as roughness of the image. This can be used to detect diseased area in the plant leaf.

(i) Color co-occurrence method: In this approach the color and texture are considered in order to get the separate set of features from the image.

Texture Features - Various properties like contrast, correlation, Energy, Homogeneity are computed for the leaf image.

Contrast: the contrast value between pixel and neighbor over the complete image which is given in the Eq. (3) Range $= [0 \, (\text{size} \, (\text{SGDM}, 1)–1)^2]$. For a constant image contrast is zero.

$$\text{Contrast} = \sum_{i,j=0}^{N-1} (i, j)^2 \, C(i, j) \tag{3}$$

Energy: The sum of squared elements is returned as the sum. Range [0 1] For a constant image energy is 1 and this can be done using Eq. (4)

$$\text{Energy} = \sum_{i,j=0}^{N=1} C(i, j)^2 \tag{4}$$

Homogeneity: The homogeneity for the leaf image is obtained as shown in Eq. (5)

$$\text{Homogeneity} = \sum_{i,j=0}^{N-1} C(i, j)/(1 + (i - j)^2) \tag{5}$$

Correlation: To measure how the pixels are coordinated with each neighboring pixels. Correlation is calculated using the Eq. (6).

$$\text{Correlation} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{\{i \times j\} \times P(i, j) - \{\mu_x - \mu_y\}}{\sigma_x \times \sigma_y} \tag{6}$$

Disease Classification:

(i) Using Artificial neural network: features are extracted and classified from the image by making use of neural network as the steps depicted in Fig. 1. Multiclass support vector machine is being used for classification. The features such as texture, color, morphology, edges etc. are being used to detect the plant leaf disease [7].
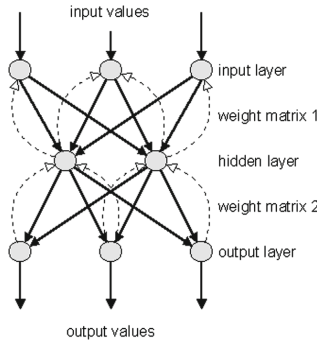
Fig. 1. Back propagation network

(ii) Using Back Propagation: The trained neural network with known weights is used for the computation of output values for the images that does not occur in the learning datasets.

4 Results

Different types of leaf diseases could be identified and so initially leaf image from the stored set of images is chosen in order to identify the disease in the leaf. Once the image is selected from the available dataset, image pre-processing is done so that smoothness of the image is achieved as well as the image is contrasted so that diseases are can be identified easily. Image is segmented using k-means clustering method. Cluster showing the disease by highlighting the affected area also one more cluster showing greenness which implies the leaf area is healthier as shown in the Fig. 2.

By selecting the cluster numbers, we can get disease category under which the leaf problem comes. Suppose if the cluster 2 is selected which is affected by the powdery mildew disease as shown in the Fig. 3.
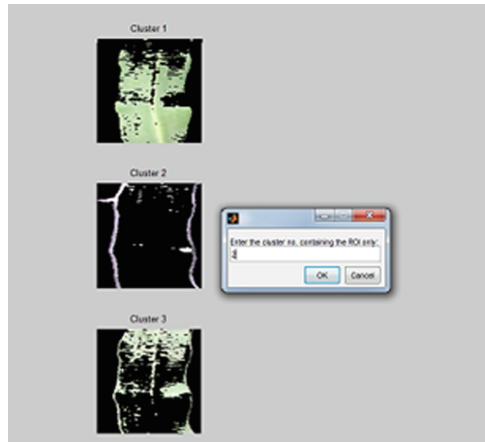
Fig. 2. K-means clustering (Color figure online)
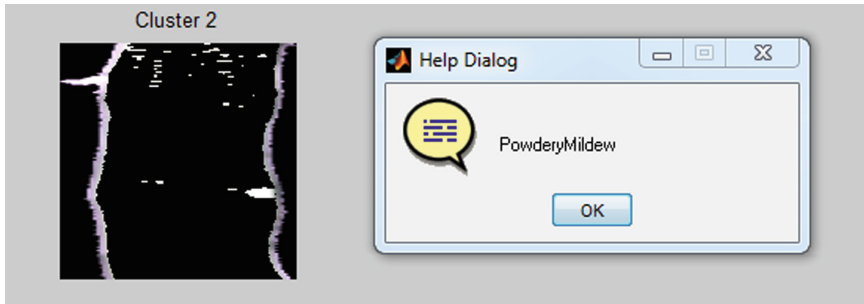
Fig. 3. Disease classification

Once the disease area is classified, affected region under the cluster selected is calculated. And the accuracy of the whole data set is obtained as shown in the Fig. 4.

The features after being extracted are trained. It is being stored for each pixel. In this regard there are different features are extracted from the leaf image. After training it using neural network it is saved in a file which as shown in Fig. 5.

For the experiment purpose leaf images around 750 to 800 are considered and checked to see whether the leaf is infected or not. We have considered three different types of disease namely, Powdery Mildew, Septoria, Wheat rust. In this, K means algorithm is used to form different clusters of the leaf. The system shows an overall accuracy of 96% for the data set which is taken in to consideration.
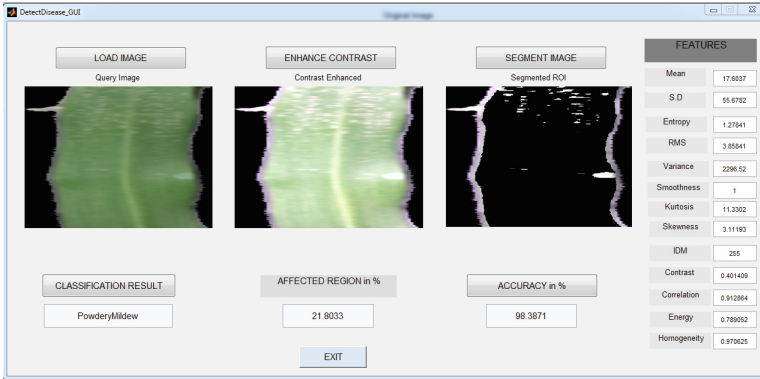
Fig. 4. Accuracy obtained over the complete datasets

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.0789	0.9783	0.7626	0.9749	14.8439	47.8117	1.7099	5.5748	2.1507e+03	1.0000	15.5978	3.6320	255
2	0.4668	0.8657	0.7967	0.9592	14.1501	48.1396	1.3658	4.3136	1.6322e+03	1.0000	15.7654	3.6744	255
3	0.3676	0.9102	0.7573	0.9625	16.4441	51.4194	1.6679	5.3404	2.3050e+03	1.0000	13.7926	3.4025	255
4	0.5412	0.7510	0.5382	0.9222	17.9717	37.6635	2.5829	7.4037	1.3068e+03	1.0000	10.4951	2.5883	255
5	0.5128	0.7103	0.8947	0.9717	17.1185	35.5205	2.8432	10.4505	1.1622e+03	1.0000	27.6033	4.6820	255
6	0.6976	0.8739	0.4873	0.9104	31.5604	56.4596	2.9830	8.1140	2.8443e+03	1.0000	4.4008	1.6129	255
7	0.4886	0.9580	0.2687	0.9403	71.8528	83.0729	5.1204	11.4616	5.6827e+03	1.0000	1.8270	0.6497	255
8	0.4309	0.8966	0.7660	0.9656	17.4376	52.4639	1.8789	5.7289	2.0524e+03	1.0000	12.8361	3.2736	255
9	0.5761	0.9092	0.7104	0.9584	23.8136	60.2088	1.6734	5.4362	3.2303e+03	1.0000	6.9582	2.3349	255
10	0.7462	0.9098	0.5279	0.9007	40.0473	73.8575	2.9119	7.5330	4.4652e+03	1.0000	3.9932	1.5873	255
11	0.8894	0.8263	0.8185	0.9651	16.4181	55.6534	1.3002	4.3228	2.8415e+03	1.0000	12.3204	3.3010	255
12	0.4140	0.9702	0.4106	0.9730	76.6184	97.9821	3.8439	9.3642	6.3323e+03	1.0000	1.5171	0.5990	255
13	0.0863	0.9491	0.8848	0.9934	8.5755	35.4527	0.7176	2.5151	1.1104e+03	1.0000	18.7524	4.1059	255
14	1.0451	0.8167	0.6192	0.9209	26.9492	60.8740	2.4818	6.9560	3.4713e+03	1.0000	6.6288	2.2075	255
15	0.4131	0.8459	0.8424	0.9766	10.6032	41.4724	1.1945	3.8289	1.5944e+03	1.0000	22.3461	4.4208	255
16	1.0066	0.7952	0.7960	0.9543	16.5970	54.7251	1.2976	4.6009	2.7719e+03	1.0000	12.3711	3.2826	255
17	1.3198	0.8648	0.4802	0.9194	44.1780	76.8477	3.3220	8.6895	5.4942e+03	1.0000	3.4116	1.4309	255
18	0.2745	0.8710	0.8596	0.9791	9.7420	38.4817	0.8864	3.5148	1.3888e+03	1.0000	19.3455	4.1094	175
19	0.5655	0.8692	0.6730	0.9549	21.6041	53.8695	2.3224	6.2055	2.4847e+03	1.0000	9.9314	2.7350	255
20	0.2625	0.8792	0.7989	0.9765	13.2757	42.4137	1.2711	4.3011	1.6140e+03	1.0000	12.9619	3.2818	255
21	0.3318	0.8662	0.7344	0.9575	18.0845	41.4308	3.8170	9.7440	1.1588e+03	1.0000	22.6075	4.2802	255
22	0.8420	0.7639	0.7044	0.9356	18.2989	47.7490	3.3386	7.1461	2.2016e+03	1.0000	13.6090	3.2690	255
23	0.4115	0.9814	0.3505	0.9448	128.9061	118.6973	3.6237	12.5744	1.0666e+04	1.0000	1.0988	0.0159	255
24	0.6671	0.9169	0.4803	0.9117	41.2289	71.5043	3.2702	8.1947	8.4091e+03	1.0000	3.8337	1.5028	255
25	0.2959	0.7321	0.9029	0.9781	5.7064	29.9149	0.6154	2.5528	844.1867	1.0000	39.5829	5.9285	255
26	0.9301	0.8328	0.5082	0.8891	30.4763	59.5952	3.0559	8.2184	3.3415e+03	1.0000	6.3927	2.0485	255
27	0.5309	0.9387	0.6026	0.9679	37.8254	74.9052	2.5919	6.9272	4.4153e+03	1.0000	4.0426	1.6585	255
28	0.2977	0.9788	0.4226	0.9748	67.4126	94.2979	3.6332	9.5141	7.9912e+03	1.0000	1.8924	0.8231	255

Fig. 5. Trained features

5   Conclusion

For the favorable crop production, accurately detecting and categorizing of the leaf disease is necessary and it can be achieved through leaf image processing technique. Various methods are considered for image segmentation. Also feature extractions and classifying methods for infected plant leaf images are employed. Back propagation method, SVMs is effectively used for classification of diseased leaves. Using image processing techniques, the diseased plant leaves can be effectively identified and classified various plant diseases.

References

1. Parikh, A., Raval, M.S., Parmar, C., Chaudhary, S.: Disease detection and severity estimation in cotton plant from unconstrained images. In: IEEE International Conference on Data Science and Advanced Analytics (DSAA) (2016)
2. Padol, P.B., Yadav, A.A.: SVM classifier based grape leaf disease detection. In: IEEE Conference on Advances in Signal Processing (CASP) (2016)
3. Zhang, Y.C., Mao, H.P., Hu, B., Li, M.X.: Features selection of cotton disease leaves image based on fuzzy feature selection techniques. In: IEEE Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, Beijing, China (2007)
4. Meunkaewjinda, A., Kumsawat, P., Attakitmongcol, K., Srikaew, A.: Grape leaf disease detection from color imagery using hybrid intelligent system. In: Proceedings of ECTI-CON (2008)
5. Badnakhe, M.R., Deshmukh, P.R.: Infected leaf analysis and comparison by Otsu threshold and k-means clustering. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **2**(3), 449–452 (2012)
6. Phadikar, S., Sil, J.: Rice disease identification using pattern recognition. In: Proceedings of 11th International Conference on Computer and Information Technology (ICCIT 2008), Khulna, Bangladesh, pp. 25–27 (2008)
7. Jhuria, M., Kumar, A., Borse, R.: Image processing for smart farming: detection of disease and fruit grading. In: Proceedings of the 2013 IEEE Second International Conference on Image Information Processing (2013)
8. Husin, Z.B., Aziz, A.H.B.A., Shakaff, A.Y.B.M., Farook, R.B.S.M.: Feasibility study on plant chili disease detection using image processing techniques. In: Third International Conference on Intelligent Systems Modelling and Simulation (2012)

Human Action Detection and Recognition Using SIFT and SVM

Praveen M. Dhulavvagol[1]([✉]) [ID] and Niranjan C. Kundur[2] [ID]

[1] KLE Technological University, Hubballi, India
praveen.md@bvb.edu
[2] JSS Academy of Technical Education, Bangalore, India
niranjanckt@gmail.com

Abstract. Human action detection and recognition is the most trending research topic in applications like surveillance of videos, analysis of sports videos and many applications which involve human computer interaction. Many researchers are working on different algorithms to improve the accuracy of human detection. Identifying the actions of human from the given video is a challenging task. In the proposed paper combination of two different techniques is applied i.e. SVM and SIFT techniques are used to identify and recognize the human actions in a given video or image. To extract local features of the given video SIFT based technique is used. In this techniques initially we extract features based on the interest points at a particular point or frames, Mainly SIFT techniques involves 4 basic steps Scale-space extreme detection, Key-point localization, Orientation assignment and Key-point descriptor. Once the key features are extracted they are further classified using SVM classifier. In the results and discussion we perform the comparative analysis of these two techniques on a standard KTH dataset with running and hand clapping actions. The experimental results determine the overall accuracy of 82% for the actions: running and hand clapping actions.

Keywords: SIFT · Scale-space · SVM · Action recognition · Key-points

1 Introduction

Action detection and recognition is active research area for many emerging video surveillance applications in our day today life mainly used for threat identification, monitoring. The primary goal is to identify and recognize the action performed in a given video or image dynamically. Action recognition plays important role in various applications like military video surveillance, video content analysis in sports videos also in many video retrieval and human-computer interaction applications. With the advancement of the technologies action recognition is also used in computer vision domain in recognizing the different human activities such as hand waving, jumping, walking, sitting, standing etc. It can also be used in video content analysis such as scenes with cluttered, moving background, variation in appearance of people, viewpoints, orientation etc. so to detect and recognize the actions from the above discussed applications is a challenging task so we need to apply appropriate feature extraction and

classification techniques, considering all the above facts has lead into the development of a fully automated human action recognition system for a given video. The system should be capable of classifying a human actions accurately is a challenging task due to problems, such as changes in scale, obstruction, cluttered background, viewpoint and tracking the dynamic motion of a human in the video, so different classes of classification are used based on analyzing the key features. According to the survey actions can be broadly classified into three categories:

i. Static action for single person: In this case we are considering the actions like standing, sitting, and all the static pose given by the human. Here the input dataset for our application are videos.
ii. Dynamic action for single person: Here in this case all the actions are dynamic, motion based videos can be considered as input for this class. The actions mainly considered are walking, running, jumping etc. for detection of such actions we require video as an input dataset along with the appearance we need to consider the movement of the person the popular datasets are KTH dataset and Weizmann datasets.
iii. Multiple people's interactions: This type of actions is relatively complex to recognize, interactions usually include handshaking, hugging, kicking and punching actions. The popular dataset for such actions are UT interaction dataset.

1.1   Action Recognition Process

Action recognition mainly consists of three main stages i.e. Video processing, feature extraction and classification (Figs. 1 and 2).
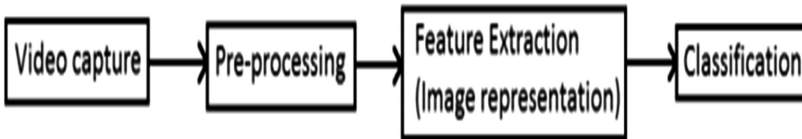
Fig. 1. Stages of action recognition

1.1.1   Video Processing
Prior to video processing a video need to be captured and decoded according to the programming environment, next step is to preprocess the video to eliminate noise and perform foreground and background subtraction.

1.1.2   Feature Extraction
In this stage we extract the key features of the image using standard key feature extraction techniques, In our case we have chosen SIFT (Scale Invariant Feature Transform) for detection and recognition of humans in a given video mainly we need to consider global features because for feature extraction and classification we need global
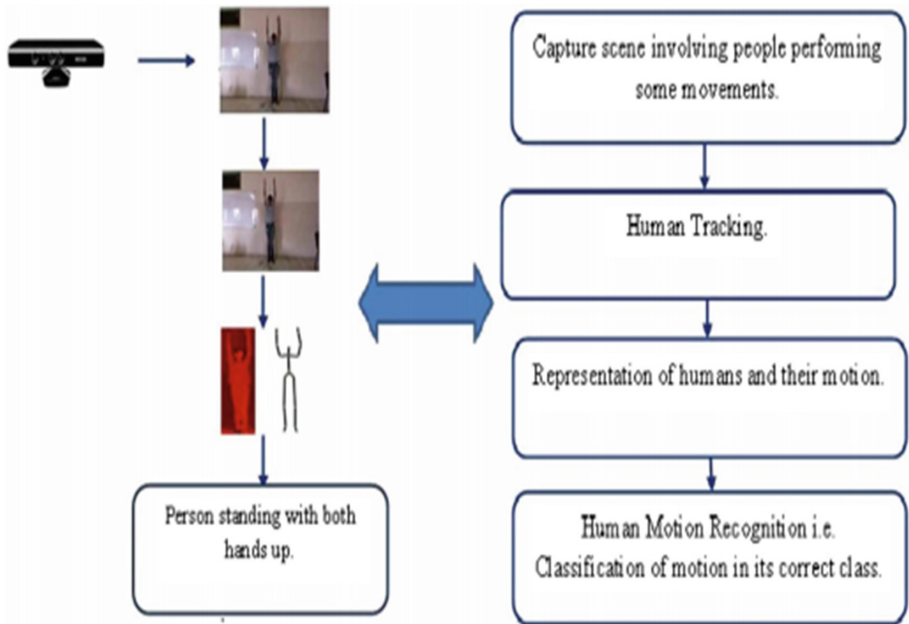
Fig. 2. General framework for human motion recognition

representation of the action. The steps carried out to extract the features are, first we need to localize the human by performing background subtraction, then in second step the ROI (region of interest) need to determine from the images/frames ROI is used to calculate the image descriptor values, since the RoI are more sensitive to viewpoint, occlusion and noise in the given video.

1.1.3 Classification

In this step we classify the actions according to the class labels, there are different classification techniques available among which one to use depends on the feature vectors weather they are of fixed size or varying size along with is it also depends on the properties of the image for example discriminative or generative, probabilistic or statistical classifiers. This means selection of classifier technique depends on the second stage i.e. feature extraction.

2   Related Work

The main focus in action recognition will be on the type of video representations, features with different key formats, different feature extraction techniques are available need to choose the efficient and relevant one, considering all the above facts a survey is carried out, The main aim is to detection of humans in the given video and to extract the

image features and classification of those features mapping to appropriate class labels. In Vision based human action classification a label is assigned for image sequences and action labels, the assignment of labels is a challenging task because of the variation in the inter-personal differences, motion performance, recording settings [2].

2.1   Inter and Intra Class Variations

The performance analysis differs for different actions depending on the variations involved for a particular actions, for example the normal walking action may have different variation like slow walking, fast walking or running this may differ in terms of speed and length, and similar type of observations are there in other action [1]. So an efficient classification algorithm needs to be chosen for classification. For increasing number of actions and class the challenge will be more because similar type of actions may overlap with each other. So considering one single domain and then performing the action classification will be a good alternative according to the authors of paper [3].

2.2   Selection of Human Motion Dataset

According to the literature survey the standard dataset considered for implementation and testing purpose are the standard KTH dataset, This dataset has six actions i.e. boxing, clapping, hand waving, walking, running and jogging, In this proposed paper we are mainly focusing on two actions i.e. hand clapping and running, Here each of these actions involves some variations i.e. walking involves three variations slow walking, fast walking and running so to recognize the particular action we need to consider the viewpoint duration which allows us to recognize a particular action, The background in all these images need to be static [3].

In case of Weizmann human action dataset ten different actions dataset are available but for our experimentation purpose we are considering only two actions walking and hand clapping. in this dataset the background and viewpoint of the image are static for robust evaluation two different sets of action sequences are recorded, we have considered two actions, the Fig. 3 shows the actions considered in which one set shows walking actions picture in different angels and second set shows front-view parallel walking actions [4].

UCF sports action dataset consists of nearly 150 sports actions such as kicking, weightlifting, diving, skating, golf swinging, kicking, running, swinging a baseball bat and walking, there is variation in the performance for different set of classes due to the appearance, background and foreground illuminations along with the view point the Fig. 3 shows the above discussed sample data sets [3, 4].

2.3   Feature Extraction and Representation

Once the appropriate dataset is selected the next phase is to extract the features from the image or video, different feature extraction techniques are available such as SIFT, HoG, NWFE and LKT (Lucas-Kanade-Tomasi). As proposed in reference paper [10]. Extracting the features from the given video or image involves four steps.
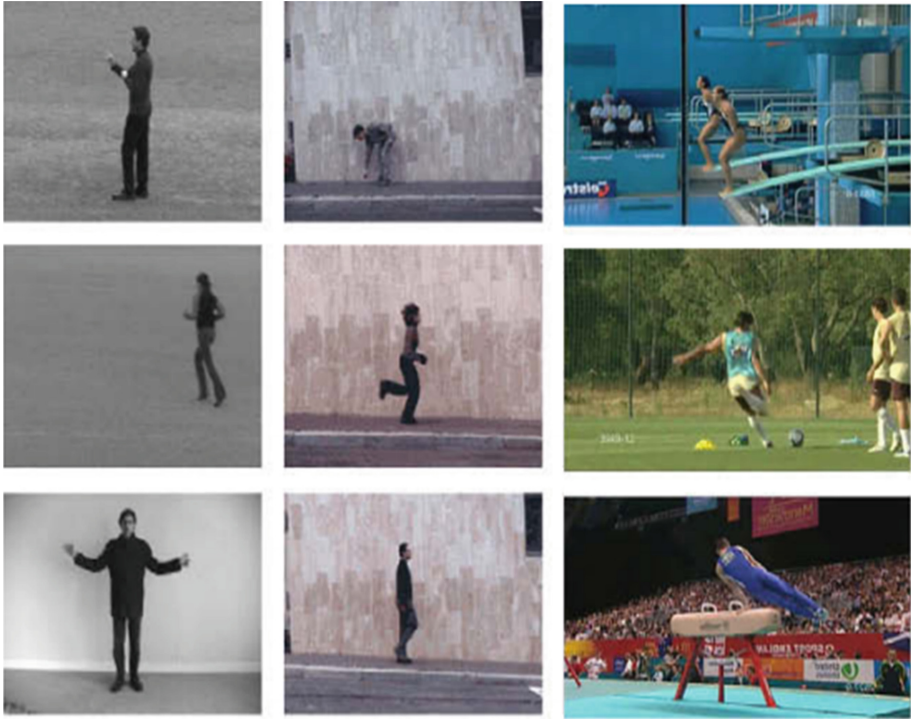
Fig. 3. Examples of KTH, Weizmann, UCF sports actions

2.4   Action Detection and Classification

Once the key features are extracted the next step is detection and classification. To achieve better performance, it is essential to select the relevant and efficient classification algorithm such as SVM, K-NN and HMM. The HMM techniques are used in speech and video applications were time sequential data is involved [11]. The HMM techniques with left-to-right state transition are considered for action recognition. But most of the human actions will be having quasi-period cycles of body movements, so its complex to model using left-to-right HMMs.

2.5   Support Vector Machine (SVM) Classifier

SVM classifier is most popular classifier, which classifies the given data based hyperplane representation i.e. on margin-based it is mainly used in pattern recognition applications. The main goal of SVM is to calculate the optimized hyperplane which will maximize the margin of two classes. On this margin the points will be scattered and these data points are called support vectors, Local space-time features can also be

extracted be using SVM classifier which is generally used to recognize the human actions. However the author in [12] has used nonlinear SVM classifier consisting of multi-dimensional Gaussian kernel for action recognition (Fig. 4).
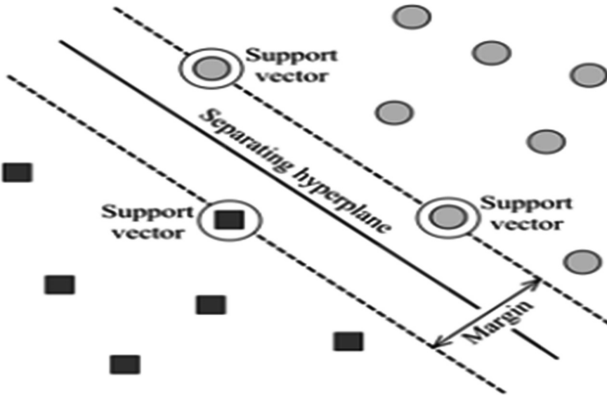
Fig. 4. Linear classification of SVM classifier

K-NN Classifier: K Nearest Neighbour classifier uses a constant value k which is predefined (K) [13], considering this value we try to identify the nearest neighbour and then try to assign label for the data points which fall within this class. In most cases K-NN is used in multi model classification.

3   Implementation

The Fig. 5 shows the overall system model which mainly consists of two phases that interact with the system: Training dataset and test data. Human Action Recognition System consists of three modules: Video pre-processing, Feature extraction and SVM classification. Our training dataset consists of videos of actions running and hand-clapping. Firstly we carry out pre-processing of videos. In the pre-processing step we convert videos into frames. Then these frames are subjected to gray scale conversion as it helps in identifying significant edges or other features and it saves the processing time. Colored images influence the speed of the process. These gray scale images are given as input to SIFT (Scale Invariant Feature Transform) algorithm. The first step of SIFT technique is to extract key features, second step is detection and recognition of human action.

3.1   Feature Extraction

Given an input video the first step is to extract the key features, which basically involves four steps:

TRAINING DATA
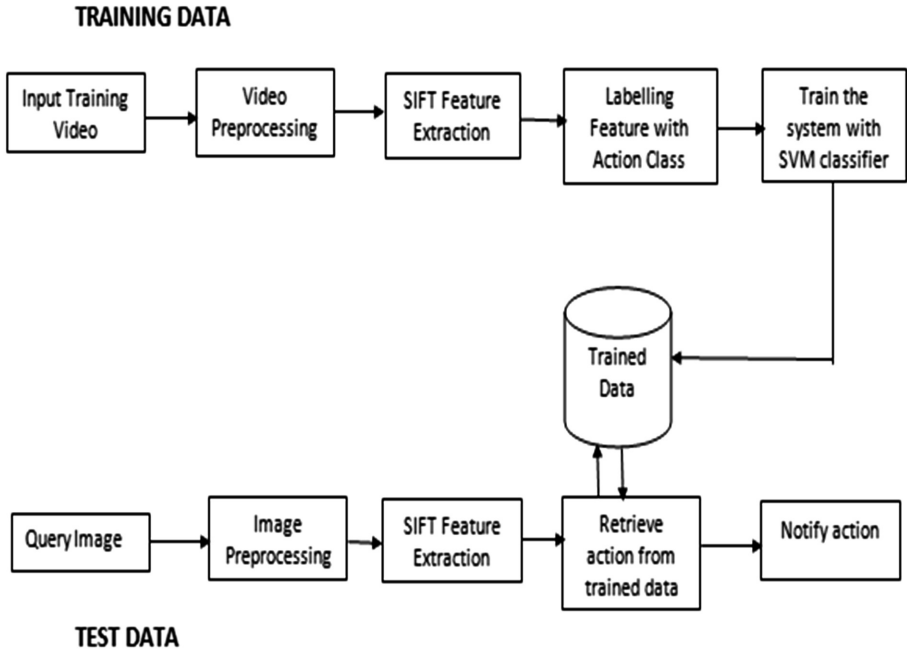
Fig. 5. System architecture

3.1.1 Scale Space Extrema Detection

In this step we create a scale space of images and construct a set of progressively blurred images. Next we consider the difference to get a DoG (Difference of Gaussian) pyramid that occurs at multiple scales. The function to define the scale space

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

Here $G(x, y, \sigma)$ is Gaussian variable-scale, $*$ is convolution operator, and $I(x, y)$ is the input image.

Localizing the scale-space extrema, $D(x, y,$ and $\sigma)$ is computed by considering the difference between two images, one with scale k times the other. Hence the scale-space extrema function is:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma)$$

Next step is to identify the local maxima and minima we need to determine local extrema in the scale space and select the key points. For each key point in a 16 * 16 window we determine the histograms of gradient direction and create a feature vector.
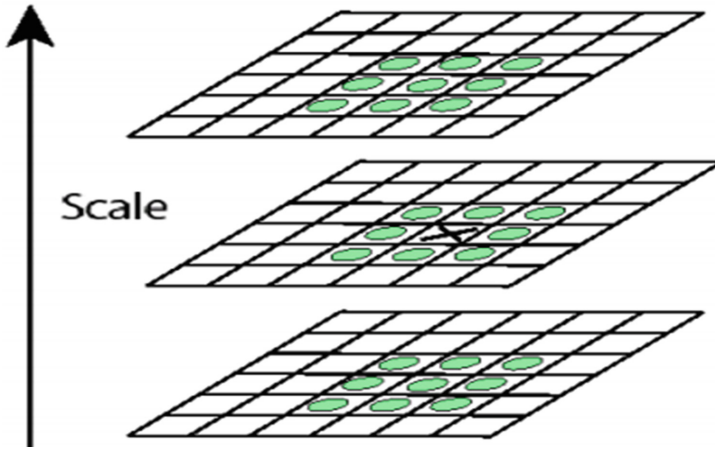
Fig. 6. Keypoint localization

3.1.2   Keypoint Localization

The Scale-space extrema method results in more number of keypoints, all the keypoints are not useful so we need to perform localization and eliminate some of the keypoints which are unstable, sensitive to noise and have low contrast such points will be eliminated. In the Fig. 6 we can observe that each point is compared with its 8 neighborhood points of the current image, and 9 neighbors each in the scales above and below. Generally in this method we try to find the robust extremum (maximum/minimum keypoints both in space and in scale. Basically we need to use DoG pyramid to determine maximum values and then using any of the edge detection technique we need to eliminate "edges" and pick only corners.

The function for determining the scale f = kernel * image

Kernels are:

Laplacian

$$L = \sigma 2(Gxx(x, y, \ \sigma) + Gyy(x, y, \ \sigma))$$

Difference of Gaussians (DoG):

$$DoG = G(x, y, \ k\sigma) - Gyy(x, y, \ \sigma))$$

In Laplacian method location of extremum, z is given by the following equation:

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x$$

If the z value falls below the threshold value then that such point will be discarded and rest of the points are considered for feature extraction. The Fig. 7 shows the sample image for determining the extrema at different scale.
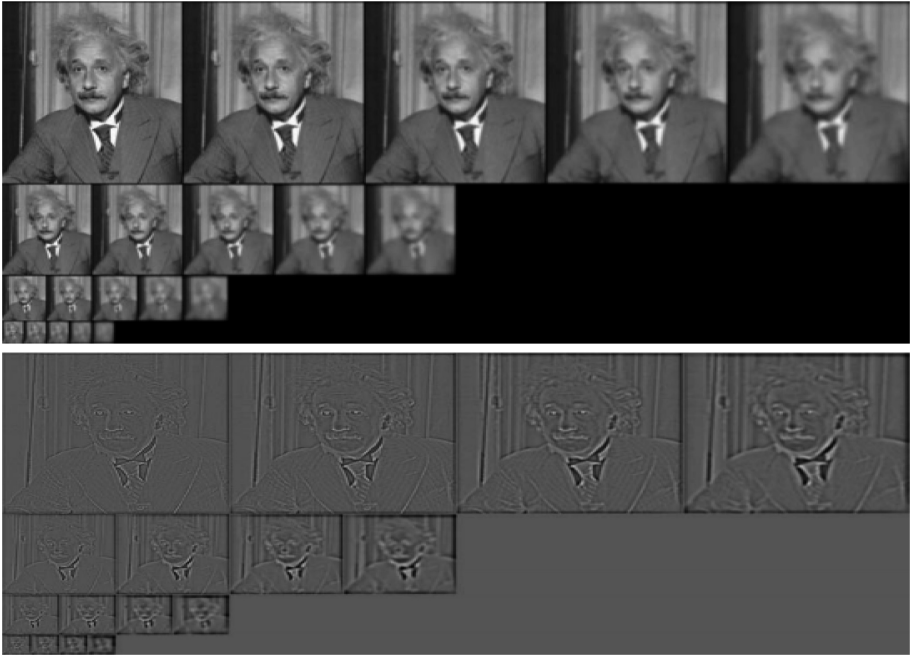
Fig. 7. Finding extrema at different scale

3.1.3 Orientation Assignment

Construct the histogram of local gradient directions at selected scale and assign each keypoint with one or more orientations based on local image gradient directions. Each keypoint now specifies stable 2D coordinates (x, y, scale, orientation). The keypoints which are obtained from localization are assigned with consistent orientation considering the image properties, Orientation can be calculated considering the keypoint scale technique for the selection of the Gaussian smoothed image value L, next we need to compute the gradient magnitude m using the function:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2}$$

Next we compute the orientation θ(x, y) considering the pixel differences:

$$\theta(x,y) = a\tan 2(L(x,y+1) - L(x,y-1), L(x+1,y) - L(x-1,y))$$

3.1.4 Keypoint Descriptor

Once the keypoints locations at particular scales and assigned orientations are determined, next we need to compute descriptor vector for the local image regions and keypoints which are invariant with respect to the viewpoint and illusion. The gradient

value obtained using orientation assignment is used to construct the keypoint descriptor. Once the descriptor values is calculated next step is to rotate the gradient value upwards upto the orientation of keypoints and then weighted by a Gaussian with variance of 1.5 * keypoint scale.

The Fig. 8 shows the 484 gradient window with histogram of 484 samples per window in 8 different directions, here in this case the Gaussian weightage around centre (σ) value is 0.5 times that of the scale keypoint, so totally it is 4 * 4 * 8 = 128 dimensional feature vector. The Fig. 9 shows the sample example of SIFT technique processing pipeline which is discussed in above section. The extracted features are labeled with action class and then the system is trained with SVM classifier. The SVM creates a hyper plane for classifying the data into a high dimensional space for separating data with different labels. Using training dataset a model is built. Now the test data is processed with the same steps as that of training dataset but for predicting the action. These test images are those that are not trained earlier. Finally both the files consisting of labeled features with action classes of training dataset and test data are given as input to SVM classifier. The SVM classifier predicts the action based on the similarity measure between training data and test data and then detects the action.
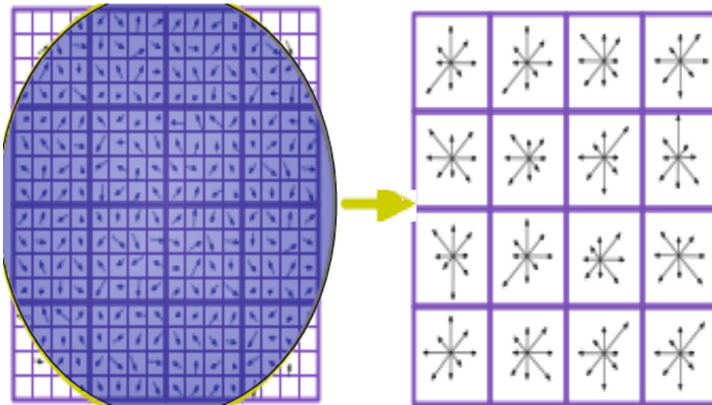
Fig. 8. Image gradient and keypoint descriptor

3.2 Video Pre-processing

Input for video preprocessing module is KTH video dataset and output is gray scale images. This module converts the input video to frames and perform image enhancement (i.e. Noise removal etc.). Based on standard convention, 26 frames are generated per second. So the number of frames generated per video depends on the length of the video. Certain frames that are generated are not informative i.e. the frames that doesn't contain complete human postures are to be eliminated. Hence these key-frames are converted to gray scale images (Fig. 10).
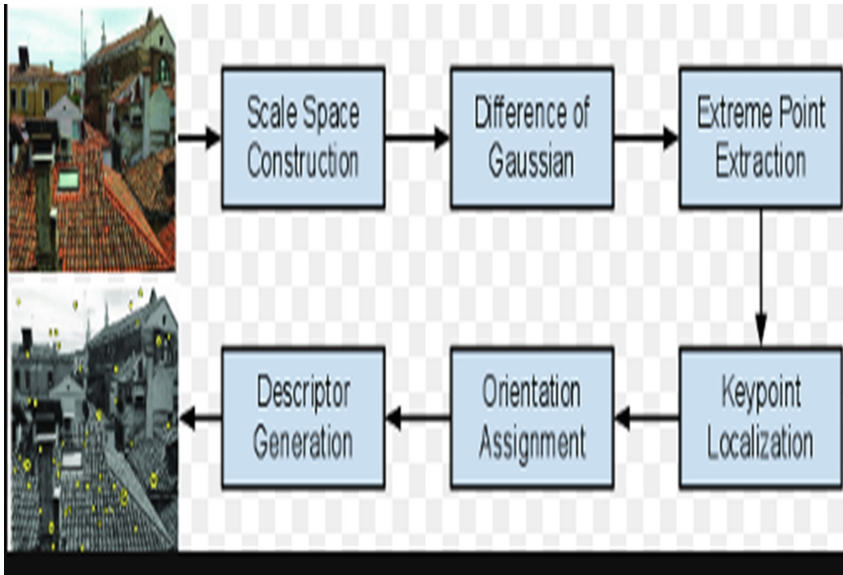
Fig. 9. SIFT algorithm processing steps
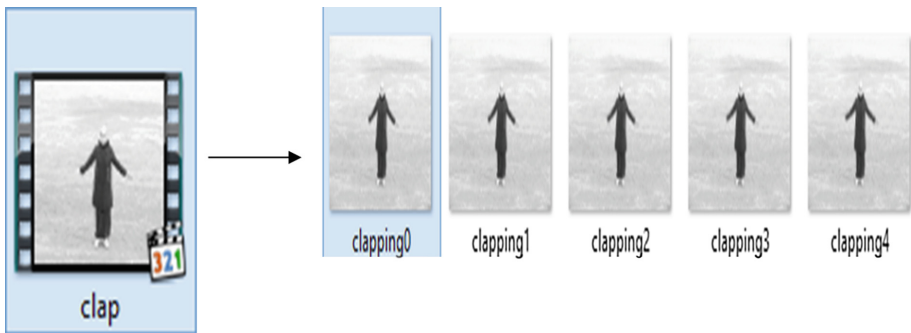
Fig. 10. Video to frames extraction

The input for SIFT technique is a gray scale images and output is local features of the image. For the given input image keypoints are determined considering the scale-space extrema, the low contrast points and edge response points along an edge are rejected. SIFT performs the following steps for extracting local features: Represent the image as the DoG pyramid, to create this pyramid the image should be scaled to 4 different sizes also called octaves. For each octave a set of images is created each with different degree of Gaussian blurred, each image is extracted with one degree less blurred, after doing this for each octave consolidate all octave and put them together to form difference of Gaussian pyramid (DoG). The Fig. 11 shows the process to determine the difference of Gaussian pyramid.
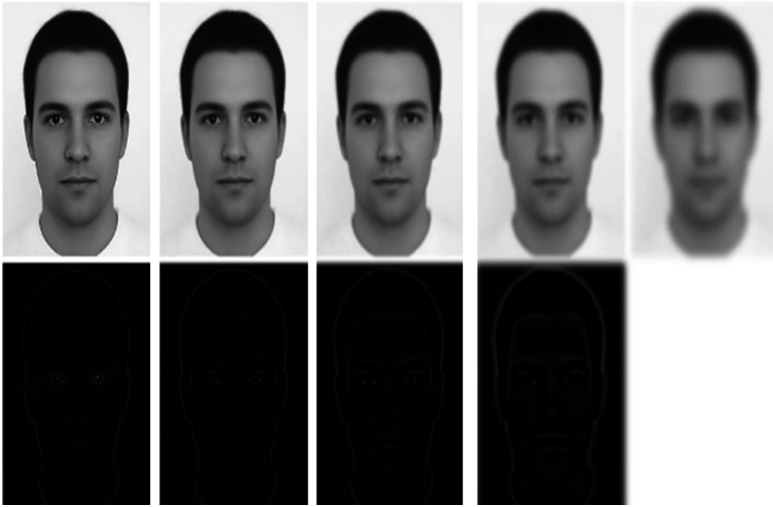
Fig. 11. Scale space extrema detection process

Next step is keypoint localization were we identify the keypoints in scale space representation this will have too many points so we need to discard flat areas and edges keeping corners. Next discard unstable extrema and keep extrema that are found at several scale space planes, at this point we should have good set of key points which will be used for orientation assignment. In orientation assignment step the keypoints are oriented based on histogram of gradients, this is done to achieve rotation invariance, at each pixel we try to find derivatives and that will give us the magnitude and direction.

Considering the interest of points and looking at the orientation around each of the points along all the directions and next we try to build the histogram of 36 bins. We construct the orientation histogram consisting of orientation histogram with 36 bins covering 360° is built. Next it is weighted by gradient magnitude and gaussian-weighted circular window with equal to 1.5 times the scale of keypoint. Assign weights for each of the pixels based on the gradient magnitude next step we transform the data to its relative to the assigned orientation, scale and location providing invariance to these transformations.

3.3 Key-Point Descriptor

At this step we will consider location, orientation and scale as keypoints. Consider the neighborhood around the future points and take the gradient directions of other points and quantize them into 8 different directions to form a histogram the Fig. 12 shows the 16 * 16 neighborhood around the point and 4 * 4 histogram array with 8 orientations forming 128 dimensions (Fig. 13).
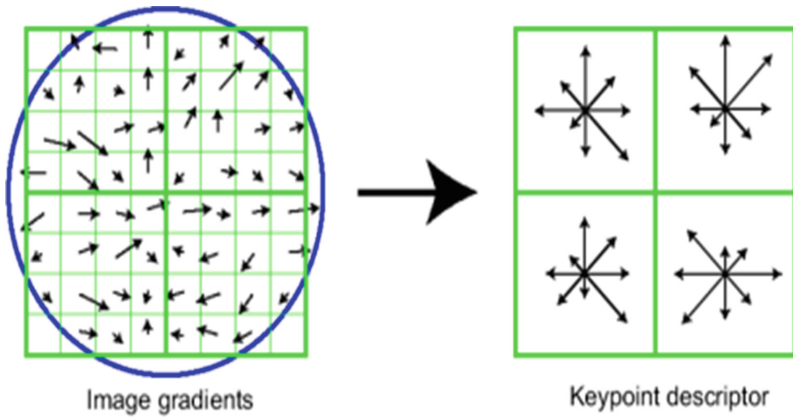
Fig. 12. Keypoint descriptor extraction process
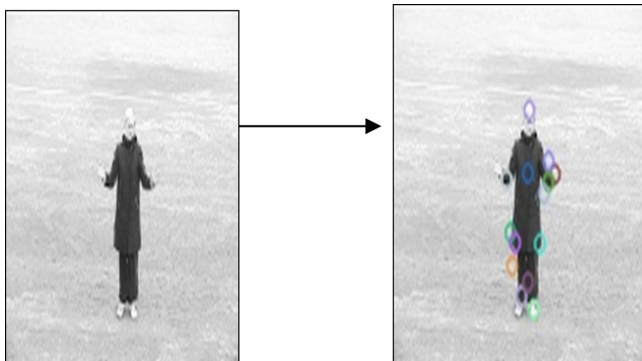
Fig. 13. SIFT keypoints extraction

3.4 SVM Classifier

A two-layer SVM classifier is used at the first level the keypoint descriptors derived from SIFT technique are used for classification, A key-point features, labels are given as input for the classifier. The goal of SVM (Support Vector Machine) is to maximize the margin of training data by finding the optimal separating hyper plane. Among all hyper planes, SVM selects the hyperplane as optimal hyperplane where the distance of hyper plane is as large as possible. Generally SVM is applied on two types of data: Separable data and non-separable data [8].

4   Results and Discussions

See Fig. 14.

Fig. 14. Computed SIFT key points

4.1   Confusion Matrix

We have drawn a confusion matrix for running and hand clapping actions for KTH and Weizmann dataset. We have trained 3 videos as a sample dataset and we have considered 100 images of KTH and Weizmann for each action as our test data. Out of 100 running images of KTH dataset, 74 were recognized correctly as running and 27 images were not recognized. Out of 100 running images of Weizmann dataset, 62 were correctly recognized as running and 38 images were not recognized. Out of 100 hand clapping images of KTH dataset 89 were correctly recognized as hand clapping and 11 images were not recognized. Out of 100 hand clapping images of Weizmann dataset 78 were correctly recognized as hand clapping and 22 images were not recognized. The system is 81% efficient.

Table 1. Test case results

Datasets	Classes	Videos	Efficiency
KTH	• Number of action classes = 2 • Verbs: Running and Hand Clapping.	• 50 videos (10 training, 4 testing). • Resolution = 160x120. • Black and white videos. • Static camera	• 73% for running and • 89% for Hand clapping
WEIZMANN	• Number of action classes = 2 • Verbs: Running and Hand Clapping.	• 20 videos (evaluate by leave one out cross validation). • Resolution = 180x144. • Static camera	• 62% for Running and • 78% for Hand clapping

The Table 1 shows the dataset considered for testing purpose the number of action classes initially considered were 2 i.e. hand clapping and running, in the Table 4 we had considered 5 different action classes were considered and tested, for KTH dataset we had considered 50 videos which were having a resolution of 160×120 which were captured from static camera and these 50 videos were processed, the efficiency of the system was 73% for running action and 89% for hand clapping, similarly for Weizmann dataset we had considered 20 sample videos with resolution of 180×144, the efficiency of the system was 67% for running and 78% for hand clapping.

Table 2. Confusion matrix for KTH data set

KTH dataset	Actions	
	Running	Hand clapping
Running	0.74	0.27
Hand clapping	0.11	0.89

Table 3. Confusion matrix for Weizmann data set

Weizmann dataset	Actions	
	Running	Hand clapping
Running	0.62	0.38
Hand clapping	0.22	0.79

The Tables 2 and 3 shows the confusion matrix for KTH & Weizmann dataset, here we have considered two actions as test cases running and hand clapping actions for KTH and Weizmann dataset. We have trained 3 videos for each action and given to SVM classifier. We have considered 100 images of KTH and Weizmann for each action as our test data. Out of 100 running images of KTH dataset, 73 were recognized correctly as running and 27 images were not recognized. Out of 100 running images of Weizmann dataset, 62 were correctly recognized as running and 38 images were not recognized. Out of 100 hand clapping images of KTH dataset 89 were correctly recognized as hand clapping and 11 images were not recognized. Out of 100 hand clapping images of Weizmann dataset 78 were correctly recognized as hand clapping and 22 images were not recognized. The system is 81% efficient.

The performance analysis of proposed approach and the related works are all done considering both KTH and Weizmann dataset which is discussed in Tables 4 and 5.

Table 4. Confusion matrix using our approach

	Running	Hand clapping	Hand waving	Walking	Jogging
Running	0.90	0.08	0.0	0.0	0.0
Hand clapping	0.06	0.88	0.0	0.0	0.2
Hand waving	0.0	0.0	0.86	0.3	0.0
Walking	0.0	0.0	0.13	0.83	0.0
Jogging	0.7	0.0	0.0	0.0	0.86

Table 5. Performance comparison with other methods

Method	KTH	Weizmann
Proposed method	97.89	96.66
Bregonzio et al.	94.33	95.66
Tran et al.	95.66	0.00
Kaaniche and Bremond	94.6	0.00

5   Conclusion

Action recognition in computer vision and video surveillance applications has increasing demand, in this paper combination of SIFT and SVM techniques are used for feature extraction and action recognition, the dataset considered for testing purpose are KTH and sports, Weizmann dataset there are totally 12 actions in KTH dataset, in the proposed paper we have tested the application considering two actions running and hand clapping, on this input dataset SIFT algorithm is applied to extract the feature vectors which will be further used to classify the action using SVM classifier. The proposed method is very robust considering the different variations in human actions. In future we can consider other complex actions and activities involving multiple people and cluttered background for recognizing actions. We can apply more robust algorithms for feature extraction and classification.

References

1. Poppe, R.: A survey on vision-based human action recognition. J. Image Vis. Comput. **28**(6), 976–990 (2009). Human Media Interaction Group, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente

2. Kellokumpu, V., Zhao, G., Pietikäinen, M.: Recognition of human actions using texture descriptors. Mach. Vis. Appl. **22**(5), 767–780 (2009). https://doi.org/10.1007/s00138-009-0233-8

3. Bermejo Nievas, E., Deniz Suarez, O., Bueno García, G., Sukthankar, R.: Violence detection in video using computer vision techniques. In: Real, P., Diaz-Pernil, D., Molina-Abril, H., Berciano, A., Kropatsch, W. (eds.) CAIP 2011. LNCS, vol. 6855, pp. 332–339. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23678-5_39

4. Schuldt, C., Laptev, I., Caputo, B.: Recognizing human actions: a local SVM approach. In: ICPR Proceedings of the 17th International Conference, vol. 3 (2004). https://doi.org/10.1109/icpr.2004.1334462

5. Ahmad, M., Lee, S.W.: Human action recognition using shape and CLG-motion flow from multi-view image sequences. Pattern Recogn. **41**(7), 2237–2252 (2008)

6. Patil, R.A., Sahula, V., Mandal, A.S.: Facial expression recognition in image sequences using active shape model and SVM. In: Fifth UKSim European Symposium on Computer Modeling and Simulation (EMS), pp. 168–173 (2011)

7. Lai, K.-T., Hsieh, C.-H., Lai, M.-F., Chen, M.-S.: Human action recognition using key points displacement. In: Elmoataz, A., Lezoray, O., Nouboud, F., Mammass, D., Meunier, J. (eds.) ICISP 2010. LNCS, vol. 6134, pp. 439–447. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-13681-8_51

8. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**, 91–110 (2004)

9. Hoang, L.U.T., Ke, S., Hwang, J., Tuan, P.V., Chau, T.N.: Quasi-periodic action recognition from monocular videos via 3D human models and cyclic HMMs. In: Proceedings of IEEE International Conference on Advanced Technologies for Communications (ATC), Hanoi, Vietnam, pp. 110–113 (2012)

10. Scheldt, C., Laptev, I., Caputo, B.: Recognizing human actions: a local SVM approach. In: Proceedings of the 17th IEEE International Conference on Pattern Recognition (ICPR), Cambridge, UK, vol. 3, pp. 32–36 (2004)

11. Kumari, S., Mitra, S.K.: Human action recognition using DFT. In: Proceedings of the Third IEEE National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), Hubli, India, pp. 239–242 (2011)

12. Hu, Y., Cao, L., Lv, F., Yan, S., Gong, Y., Huang, T.: Action detection in complex scenes with spatial and temporal ambiguities. In: Proceedings of the 12th IEEE International Conference on Computer Vision (ICCV), pp. 128–135 (2009)

13. Bregonzio, M., Xiang, T., Gong, S.: Fusing appearance and distribution information of interest points for action recognition. Pattern Recogn. **45**(3), 1220–1234 (2012)

Novel Real-Time Video Surveillance Framework for Precision Pesticide Control in Agribusiness

Nayana G. Bhat[(✉)] and Guruprasad M. Bhat[(✉)]

Jyothy Institute of Technology, Tataguni, Bangalore, India
nayana.gb@ciirc.jyothyit.ac.in,
guruprasad.bhat@jyothyit.ac.in

Abstract. Agribusiness has turned out to be considerably more than just a method to bolster consistently developing populaces. Plants have progressed toward becoming an imperative wellspring of vitality, and are a basic piece in the puzzle to take care of the issue of an unnatural climatic change. The economic changes of the mid 1990s set the phase for an improved and developing part by the private area of agriculture. Economic development quickened, especially after 2000, with major modernizing impacts on agriculture, cultivating and agri-food esteem chains [1]. Private agribusiness organizations are at the front line of overwhelming interest in agricultural R&D and mechanical development [2]. To improvise and expand agribusiness, one of the step is to control the pests on growing plants. We are proposing a framework for controlling pests on the growing plants of tomato in agribusiness using Real-Time Video Surveillance and image processing technique. In the proposed work, video is captured, converted to suitable color models, and pre-processed. From the obtained video, pests are quantified. Pesticides can be sprayed on the plants of tomato for controlling pests and improve yield in agribusiness.

Keywords: Begomoviruses in whitefly · Video processing technique
Automated tests management · Image analysis · Object detection
Object segmentation · Object extraction

1 Introduction

Agriculture keeps on playing remarkable part in Indian Economy. One noteworthy issue for agriculturists is to diminish or devastate the advancement of vermin influencing harvest yields. The chief method or methodology utilized routinely for the identification and acknowledgment of yield pests by numerous agriculturists in real parts of the world is perception with the bare eyes. This procedure requires persistent checking of the yield stems and leaves, which are troublesome, work escalated, inappropriate and costly for expansive farmsteads. Issues related with this technique is continual witnessing, which is extremely troublesome, sprinkling pesticides in improper amount and these pesticides will cause skin malevolence. Whitefly move quickly to tomato and other vegetable harvests, transmitting distinctive Begomoviruses, causing significant economic losses and nutrition values of vegetables. In this specific condition, diagnosing bugs in a

correct and supportive way is completely basic. A couple of bugs don't have any detectable signs related, or those appear to be exactly when it is past the point where it is conceivable to act. In those cases, generally some sort of current examination, when in doubt by techniques for capable flexible instruments, is fundamental. A run of the mill approach for this condition is the utilization of remote recognizing strategies that look at multi and hyperspectral picture gets. The frameworks that grip this approach as frequently as conceivable use pushed picture managing propelled picture taking care of contraptions to satisfy their goals.

Hence, in this paper, we intend to design layout and develop a system that can screen the whitefly bug action in light of which the pesticides can be showered normally to keep the bugs in unequivocal levels. The essential focus of our undertaking is to utilize savvy framework.

2   Related Work

Ailment in plants have been to a great extent contemplated in the logical region, basically concentrating on the natural attributes of diseases [3]. For example, food grade plants, like tomato [4] was influenced by diseases was demonstrated [5]. The issue of plant maladies is an around the world issue likewise identified with nutrition security of food. Research is being carried to eradicate pests on various growing crop yields. As we know, many crops are less yielded because of unfavorable environment, fungal, bacteria's and viral infections [5]. In 1983, A device was developed for finding the disease severity on plants using digital image processing [6], the image was fed to microcomputers, which could not be further modified. The image was able to visualize and can be used to further process. Depending on the RGB color distribution the diseases on rubber leaf was diagnosed using advanced neural network for 600 samples, where they could have achieved 80% result [7]. In [8], photos of damaged and good maize plants were taken in various stages and in three distinctive light forces. The proposed calculation had two phases: the handling and the image investigation. Amid the principal organize, the images were handled to make twofold images, where the leaves were divided from alternate pixels. At the second stage, the pictures were subdivided into squares and delegated 'harmed' or 'non-harmed' contingent upon the quantity of items found in each piece. The calculation accurately ordered 94.7% of 720 pictures. Self-decision machine vision system for continuous observing of lettuce plant wellbeing and development in controlled-condition plant-generation frameworks has been developed by [9]. Top foreseen covering zone, vitality, entropy, and homogeneity textural parameters were the most extreme empowering pointers for promising area of calcium inadequacy in the lettuce. The created method perceived calcium-deficient lettuce plants one day sooner than identification by human eyes. [10] projected a strategy for evaluating the reality of organisms related sickness in sugar stick. With the proper quantity and combination of pesticide by evaluating disease seriousness utilizing image processing technique. Simple threshold and Triangle thresholding strategies are utilized to segment the leaf range and sore region individually. At long last illnesses are categorize by figuring the remainder of sore region and leaf region. The precision of the analysis is observed to be 98.60%. Research demonstrates that this strategy to compute leaf infection seriousness is quick and exact. For finding the diseases in paddy plantation [11] consolidated the techniques like, scanner

image securing, examining enhancement, furthermore, progressed cognitive vision. Contaminations that have ascended in the past two decades and that purpose of containment the making of basic vegetable collects in tropical, subtropical, and quiet ranges are as often as possible transmitted by the whiteflies Begomoviruses, Bemisia tabaci and Trialeurodes vaporariorum. Components driving the improvement of whitey-transmitted scatters join genetic changes in the contamination, changes in the vector masses, and long detachment transport of plant material or vector dreadful little animals on account of trade of vegetables and resplendent plants [12]. In [13], depicted method of Pomegranate organic product sicknesses identification and illnesses are: Alterneria, Bacterial Curse and Anthracnose. Pre-processing included picture resizing, sifting and morphological operations. RGB, La*b, HSV and YCbCr are utilized to make groups in division. In highlight extraction shading, morphology and surface highlights are extricated and Gabor channel is utilized as a part of surface and morphology for getting limit of picture. Shape vectors are removed from solid organic product picture and minimum separation classifier (MDC) is connected for preparing and grouping of ailing or non-unhealthy images. In [14], work has been done on tomato plants using SVM algorithm and segmentation to detect disease and achieved 93% of the result. Complexity, speed and size are some of the drawbacks of SVM algorithm. As complexity increases cost increases. The proposed work is a cost-effective system utilize Simulink, Hardware description language to detect and quantify the whitefly on tomato plants.

3   Experimental Setup

See Fig. 1.
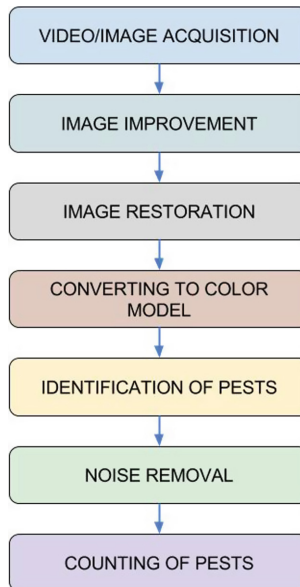
Fig. 1. Flowchart to detect and count the white flies on the leaf of tomato plants.

Utilizing multi-utility, high quality wide angle lens of ROBO K20 image/video captured (Fig. 2).

Fig. 2. Camera input images

Image is improved or enhanced for the Interested features using spatial domain (Fig. 3).

Fig. 3. Enhanced image

Using deconvolution process, noise is removed from the image and restored (Fig. 4).

Fig. 4. Restored image

The acquired image is converted into appropriate color model i.e.; YCbCr, HSV color models using MATLAB. Conversion parameter are utilized to determine the shading spaces for changing over between images. Here the images taken from ROBO K20 Webcam is specifically taken in YCbCr and it is changed over into RGB utilizing shading space transformation (Fig. 5).

| Component Image | Cb component Image | Cr component Image |
| H component Image | S component Image | V component Image |
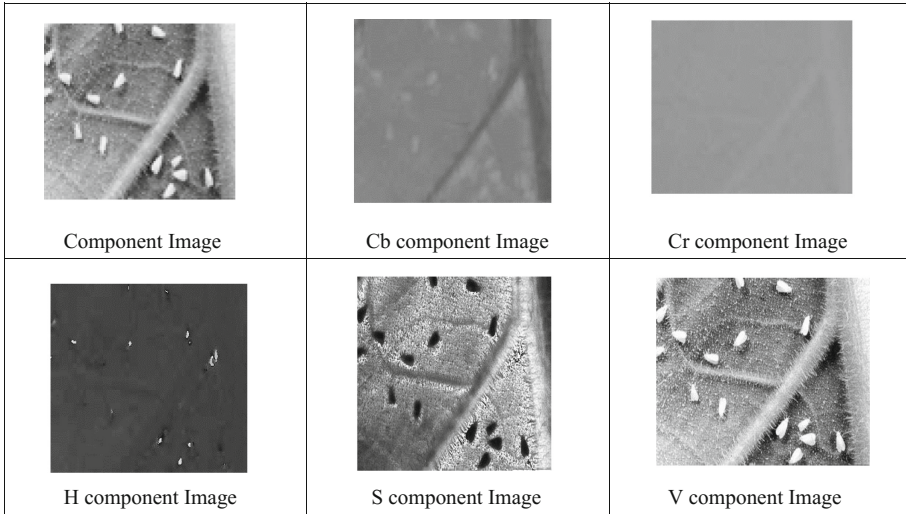
Fig. 5. YCbCr, HSV color images

Segmentation performed to streamline and additionally change the portrayal of a pictrue into something that is more significant and less demanding to break down. To be accurate, image segmentation is done by assigning a name to each pixel in an image to such an extent that pixels with a similar mark share certain attributes. The changed over image is then divided utilizing MATLAB to distinguish the fly. The image is segmented using MATLAB function, f = im2bw (cab, x); We have chosen CBS component image as it is brightness independent and the converted image is being subjected to segmentation in MATLAB. The segmented image is shown in Fig. 6.

Cb image.   Segmented image.
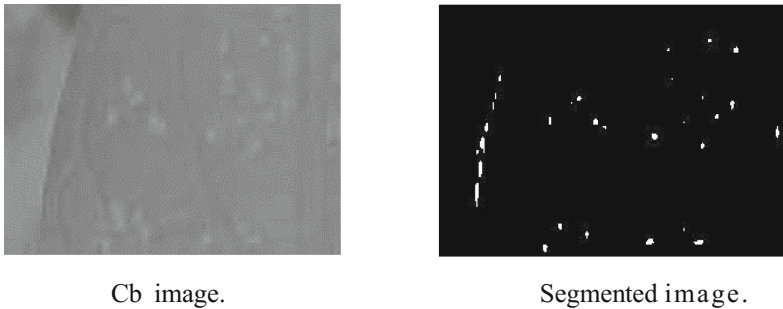
Fig. 6. Segmented image

The noise signal is removed using MATLAB function called erosion. The undesirable white pixels evacuated utilizing Erosion technique and the dilation operation for organizing component for probing and extending the shapes contained in the input image (Fig. 7).

Gray scale image   Eroded image

Fig. 7. Noise removed image

4 Simulink Model

Acquired live image data from an image capturing device. Then color data was converted between color spaces. The different shading spaces exist since they display shading data in ways that make certain estimations more advantageous or in light of the fact that they give an approach to distinguish hues that is more instinctive. For example, the RGB shading space characterizes a shading as the rates of red, green, and blue tones combined. Other shading models depict colors by their hue (green), saturation (dark green), and luminance, or intensity. Here the image taken from ROBO K20 Webcam was directly taken in YCbCr and it is converted into RGB using color space conversion. Latter image datatype conversion was done, where the floating-point datatypes are changeover is done, at the same time input is scaled up, and each block output the range either 0 or 1. While changing over, the output information also scales to dynamic scope of the required information type (Fig. 8).
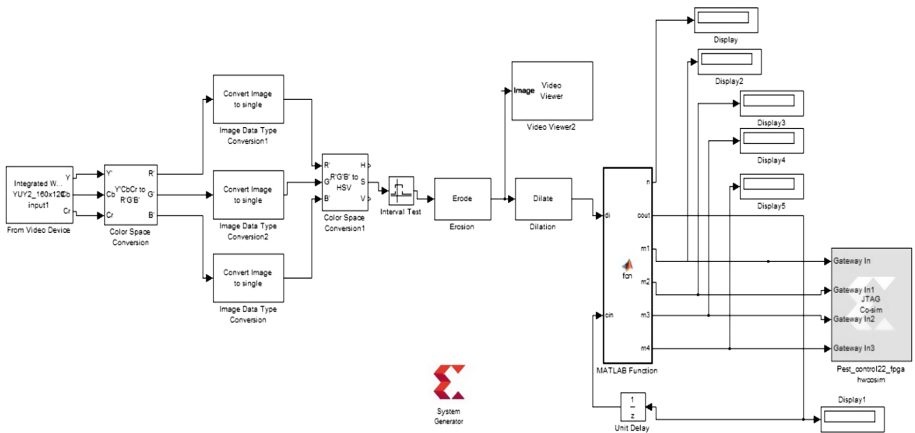
Fig. 8. Complete simulink model

Erosion was carried out, as erosion is one of two key operations in morphological picture handling from which all other structuring of elements and other operations are based. It was initially characterized for binary images, later being reached out to grayscale

pictures, and in this way to finish lattices. Then morphological dilation was performed on an intensity or binary image. Then, video and numeric values were displayed.

5　Results

Table below showcase the results for various Real-time image taken by the camera, for Cb converted input image and Segmented image (Table 1).

Table 1.　Results

Image	Cb MAX	Cb MIN	S MAX	S MIN
	127	102	0.62	0.10
	131	10	0.84	0.18
	131	110	0.52	0.05
	123	81	0.68	0.11
	124	100	0.37	0.04

	133	94	0.92	0.06
	130	107	0.54	0.03
	133	101	0.33	0.02
	124	103	0.58	0.03
	124	100	0.37	0.04
	123	96	0.41	0.06
	124	92	0.41	0.03
	130	104	0.37	0.04

6 Conclusion

In proposed work, designed and developed a system that monitor the whitefly action. This is done in order to help the farmers and improve agribusiness to lower economic damage and to avoid the harms from the pesticides used by farmers in bear hands which leads to problem to the farmer as well as the consumer. The algorithm has been developed using VHDL to find the white flies. The image of the white flies was captured through the camera and processed to find the number of white flies present on the leaf and counted. This technique need less processing and cost effective for agribusiness. As future scope a robot can be built to sprinkle pesticides based upon the count of the white flies present on the leaf, whenever required, to improve agribusiness.

References

1. Abdullah, N.E., Rahim, A.A., Hashim, H., Kamal, M.M.: Classification of rubber tree leaf diseases using multilayer perceptron neural network. In: 5th Student Conference on Research and Development 2007, SCOReD 2007, pp. 1–6. IEEE (2007)
2. Birthal, P.S., Chand, R., Joshi, P.K., Saxena, R., Rajkhowa, P., Khan, M., Khan, M.A., Chaudhary, K.R.: Formal versus informal: efficiency, inclusiveness, and financing of dairy value chains in India. International Food Policy Research Institute (2016)
3. Coakley, S.M., Scherm, H., Chakraborty, S.: Climate change and plant disease management. Annu. Rev. Phytopathol. **37**, 399–426 (1999)
4. Ferroni, M., Zhou, Y.: The private sector and India's agricultural transformation. Glob. J. Emerg. Mark. Econ. 0974910117716406 (2017)
5. Gilbertson, R.L., Batuman, O.: Emerging viral and other diseases of processing tomatoes: biology, diagnosis and management. Acta Hortic. **971**, 35–48 (2013)
6. Jones, D.R.: Plant viruses transmitted by thrips. Eur. J. Plant Pathol. **113**, 119–157 (2005)
7. Lindow, S., Webb, R.: Quantification of foliar plant disease symptoms by microcomputer-digitized video image analysis. Phytopathology **73**, 520–524 (1983)
8. Navas-Castillo, J., Fiallo-Olivé, E., Sánchez-Campos, S.: Emerging virus diseases transmitted by whiteflies. Annu. Rev. Phytopathol. **49**, 219–248 (2011)
9. Patil, S.B., Bodhe, S.K.: Leaf disease severity measurement using image processing. Int. J. Eng. Technol. **3**, 297–301 (2011)
10. Raut, S., Fulsunge, A.: Review on fruit disease detection using image processing techniques. Int. J. Innov. Emerg. Res. Eng. **4**(4) (2017). e-ISSN 2394 - 3343. p-ISSN 2394 – 5494
11. Sena Jr., D.G., Pinto, F.A.C., Queiroz, D.M., Viana, P.A.: Fall armyworm damaged maize plant identification using digital images. Biosyst. Eng. **85**, 449–454 (2003). https://doi.org/10.1016/S1537-5110(03)00098-9
12. Senthilrajan, A.: Pest control in paddy using segmentation in image processing. Eng. Sci. Int. Res. J. ISSN. 2320–4338
13. Shakoor, N., Lee, S., Mockler, T.C.: High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field. Curr. Opin. Plant Biol. **38**, 184–192 (2017)
14. Story, D., Kacira, M., Kubota, C., Akoglu, A., An, L.: Lettuce calcium deficiency detection with machine vision computed plant features in controlled environments. Comput. Electron. Agric. **74**, 238–243 (2010)

Correction to: An IOT Based Smart Shopping Cart for Smart Shopping

Srinidhi Karjol 🆔, Anusha K. Holla 🆔, C. B. Abhilash 🆔,
P. V. Amrutha, and Y. V. Manohar

Correction to:
**Chapter "An IOT Based Smart Shopping Cart for Smart
Shopping" in: T. N. Nagabhushan et al. (Eds.):**
Cognitive Computing and Information Processing, **CCIS 801,**
https://doi.org/10.1007/978-981-10-9059-2_33

In the originally published chapter 33 the names of two authors who contributed to the work were not included. P. V. Amrutha and Y. V. Manohar have been added as co-authors of the paper.

Author Index