

Volume 9 ■ Issue 3 ■ November 2015

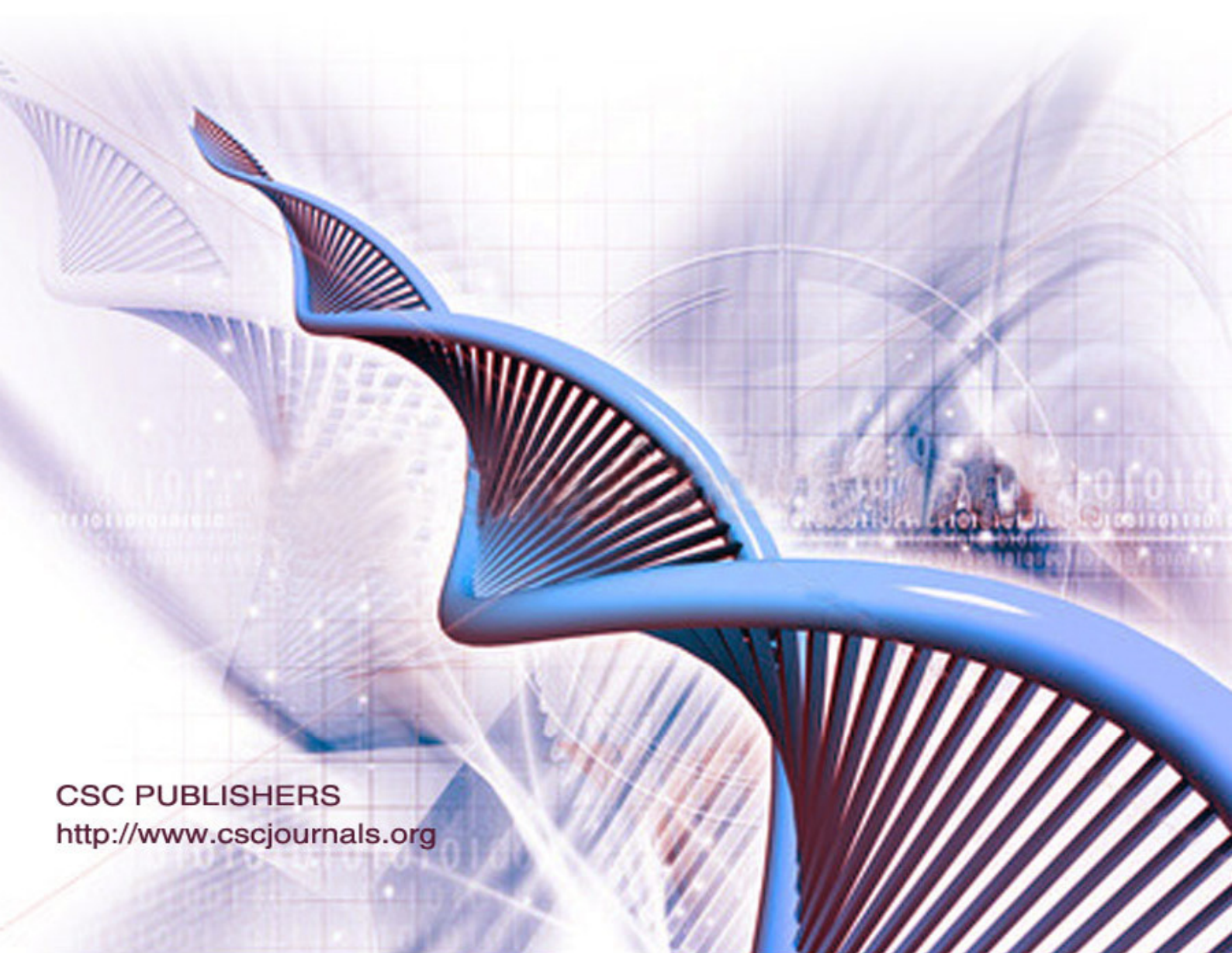
Editor-in-Chief
Professor João Manuel R. S. Tavares

INTERNATIONAL JOURNAL OF

BIOMETRICS AND BIOINFORMATICS (IJBB)

ISSN : 1985-2347

Publication Frequency: 6 Issues / Year



CSC PUBLISHERS
<http://www.cscjournals.org>

INTERNATIONAL JOURNAL OF BIOMETRICS AND BIOINFORMATICS (IJBB)

VOLUME 9, ISSUE 3, 2015

**EDITED BY
DR. NABEEL TAHIR**

ISSN (Online): 1985-2347

International Journal of Biometrics and Bioinformatics (IJBB) is published both in traditional paper form and in Internet. This journal is published at the website <http://www.cscjournals.org>, maintained by Computer Science Journals (CSC Journals), Malaysia.

IJBB Journal is a part of CSC Publishers

Computer Science Journals

<http://www.cscjournals.org>

INTERNATIONAL JOURNAL OF BIOMETRICS AND BIOINFORMATICS (IJBB)

Book: Volume 9, Issue 3, November 2015

Publishing Date: 30-11-2015

ISSN (Online): 1985-2347

This work is subjected to copyright. All rights are reserved whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication of parts thereof is permitted only under the provision of the copyright law 1965, in its current version, and permission of use must always be obtained from CSC Publishers.

IJBB Journal is a part of CSC Publishers

<http://www.cscjournals.org>

© IJBB Journal

Published in Malaysia

Typesetting: Camera-ready by author, data conversion by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers, 2015

EDITORIAL PREFACE

This is the *Third* Issue of Volume *Nine* of International Journal of Biometric and Bioinformatics (IJBB). The Journal is published bi-monthly, with papers being peer reviewed to high international standards. The International Journal of Biometric and Bioinformatics is not limited to a specific aspect of Biology but it is devoted to the publication of high quality papers on all division of Bio in general. IJBB intends to disseminate knowledge in the various disciplines of the Biometric field from theoretical, practical and analytical research to physical implications and theoretical or quantitative discussion intended for academic and industrial progress. In order to position IJBB as one of the good journal on Bio-sciences, a group of highly valuable scholars are serving on the editorial board. The International Editorial Board ensures that significant developments in Biometrics from around the world are reflected in the Journal. Some important topics covers by journal are Bio-grid, biomedical image processing (fusion), Computational structural biology, Molecular sequence analysis, Genetic algorithms etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with Volume 9, 2015, IJBB appears with more focused issues related to biometrics and bioinformatics studies. Besides normal publications, IJBB intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

The coverage of the journal includes all new theoretical and experimental findings in the fields of Biometrics which enhance the knowledge of scientist, industrials, researchers and all those persons who are coupled with Bioscience field. IJBB objective is to publish articles that are not only technically proficient but also contains information and ideas of fresh interest for International readership. IJBB aims to handle submissions courteously and promptly. IJBB objectives are to promote and extend the use of all methods in the principal disciplines of Bioscience.

IJBB editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJBB. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJBB provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members

International Journal of Biometric and Bioinformatics (IJBB)

EDITORIAL BOARD

EDITOR-in-CHIEF (EiC)

Professor João Manuel R. S. Tavares
University of Porto (Portugal)

ASSOCIATE EDITORS (AEiCs)

Assistant Professor. Yongjie Jessica Zhang
Mellon University
United States of America

Professor. Jimmy Thomas Efird
University of North Carolina
United States of America

Professor. H. Fai Poon
Sigma-Aldrich Inc
United States of America

Professor. Fadiel Ahmed
Tennessee State University
United States of America

Professor. Yu Xue
Huazhong University of Science and Technology
China

Associate Professor Chang-Tsun Li
University of Warwick
United Kingdom

Professor. Calvin Yu-Chian Chen
China Medical university
Taiwan

EDITORIAL BOARD MEMBERS (EBMs)

Assistant Professor. M. Emre Celebi
Louisiana State University
United States of America

Dr. Ganesan Pugalenth
Genome Institute of Singapore
Singapore

Dr. Vijayaraj Nagarajan
National Institutes of Health
United States of America

Dr. Wichian Sittiprapaporn
Mahasarakham University
Thailand

Dr. Paola Lecca
University of Trento
Italy

Associate Professor. Renato Natal Jorge
University of Porto
Portugal

Assistant Professor. Daniela Iacoviello
Sapienza University of Rome
Italy

Professor. Christos E. Constantinou
Stanford University School of Medicine
United States of America

Professor. Fiorella SGALLARI
University of Bologna
Italy

Professor. George Perry
University of Texas at San Antonio
United States of America

Assistant Professor. Giuseppe Placidi
Università dell'Aquila
Italy

Assistant Professor. Sae Hwang
University of Illinois
United States of America

Associate Professor Quan Wen
University of Electronic Science and Technology
China

Dr. Paula Moreira
University of Coimbra
Portugal

Dr. Riadh Hammami
Laval University
Canada

Dr Antonio Marco
University of Manchester
United Kingdom

Dr Peng Jiang
University of Iowa
United States of America

Dr Shunzhou Yu

General Motors Global R&D Center
United States of America

Dr Christopher Taylor

University of New Orleans
United States of America

Dr Horacio Pérez-Sánchez

University of Murcia
Spain

TABLE OF CONTENTS

Volume 9, Issue 3, November 2015

Pages

- 25 - 42 Analysis of Human Electrocardiogram for Biometric Recognition Using Analytic and AR
Modeling Extracted Parameters
Branislav Vuksanovic, Mustafa Alhamdi

Analysis of Human Electrocardiogram for Biometric Recognition Using Analytic and AR Modeling Extracted Parameters

Branislav Vuksanovic

*Faculty of technology/School of engineering
University of Portsmouth
Portsmouth, PO1 2UP, United Kingdom*

branislav.vuksanovic@port.ac.uk

Mustafa Alhamdi

*Faculty of technology/School of engineering
University of Portsmouth
Portsmouth, PO1 2UP, United Kingdom*

mustafa.alhamdi@port.ac.uk

Abstract

The electrocardiograph (ECG) contains cardiac features unique to each individual. By analyzing ECG, it should therefore be possible not only to detect the rate and consistency of heartbeats but to also extract other signal features in order to identify ECG records belonging to individual subjects. In this paper, a new approach for automatic analysis of single lead ECG for human recognition is proposed and evaluated. Eighteen temporal, amplitude, width and autoregressive (AR) model parameters are extracted from each ECG beat and classified in order to identify each individual. Proposed system uses pre-processing stage to decrease the effects of noise and other unwanted artifacts usually present in raw ECG data. Following pre-processing steps, ECG stream is partitioned into separate windows where each window includes single beat of ECG signal. Window estimation is based on the localization of the R peaks in the ECG stream that detected by Filter bank method for QRS complex detection. ECG features – temporal, amplitude and AR coefficients are then extracted and used as an input to K-nn and SVM classification algorithms in order to identify the individual subjects and beats. Signal pre-processing techniques, applied feature extraction methods and some intermediate and final classification results are presented in this paper.

Keywords: ECG Biometric, Filtering, QRS Detection, AR Model, Extraction and Classification.

1. INTRODUCTION

Biometric recognition provides authentication by identifying each individual based on the biological and physiological signal characteristics. A number of identification methods have been investigated in the last decades [1], using physical features such as finger prints, face images [2] and biological signal behavior such as electrocardiograph (ECG) [3]. Analysis of ECG signals as a biological tool for individual recognition has become an active research field in the past two decades [4][5]. Validity of using ECG as a biometric tool is supported by the fact that each ECG signal has certain unique features [6] which can be used to distinguish it from other ECG signals. Differences between ECG signals are usually caused by the variability of heart position and orientation relative to the ribs (the ribs being the reference clinically used to place the precordial electrodes on body), which are highly variable among different persons. Other differences can be related to body habitus [7], sex, age, length, and weight of the subjects.

The ECG signal classification is usually considered in the light of selection, extraction and classification of extracted features. High recognition rate has been achieved with the approach based on the autocorrelation (AC) in conjunction with discrete cosine transform (DCT) [3]. Proposed method does not require any waveform or fiducial point detections but AC and DCT can be computationally demanding and require long ECG records for each subject or individual to

identify them successfully. A method known as Pulse Active Width (PAW) is implemented on ECG for biometric authentication [8]. The results of this approach have indicated that PAW yields equivalent performance in terms of accuracy compared to conventional temporal and amplitude feature extraction methods. However, PAW is a complex signal processing technique which requires powerful digital signal processors to overcome the time delay.

The electrocardiogram is an emerging biometric modality that has seen about 13 years of development reported in peer-reviewed literature [9], and as such deserves a systematic review and discussion of the associated methods and findings. In particular, the categorization of methodologies in ECG based biometry relies on the feature extraction and classification schemes. Finally, comparative analysis is presented to estimate performance of this approach of ECG biometric system. The comparative study includes the cases where training and testing data sets come from the same and different partition of ECG signal measured on different days. Recently, cardiovascular signals have been studied for use in identity recognition problems using electrocardiography [10], [11] and [12].

A number of different techniques have been implemented in the past years to design ECG classification system. Mahmoodabadi [13] described an approach for ECG feature extraction which utilizes multi-resolution wavelet transform for ECG feature extraction. In this desired study, the ECG signals from modified lead II were chosen for processing. Experimental results showed that proposed approach for ECG feature extraction achieved accuracy of up to 98% to classify ECG signals. A mathematical morphology based on Kohonen self-organizing maps (SOM) and Learning Vector Quantization (LVQ) to design ECG feature extraction was proposed by Tadejko and Rakowski [14]. The main focus of this work is to evaluate the classification performance of an automatic classifier of the ECG designed to detect abnormal beats using new concept of feature extraction stage. The accuracy or correct classification rate of this method is 94%. An automatic extraction of both time interval and morphological features from ECG to classify ECGs into normal and arrhythmic was described by Alexak [15]. The method utilized the combination of artificial neural networks (ANN) and Linear Discriminant Analysis (LDA) techniques for feature extraction. The average accuracy of this method is about 85%. A new ECG pattern classification model based on a generic feature extraction method is presented in [16]. Correct classification rate is 88% for this method. Work described in [17] presents an approach to ECG signal filtering and classification system using HMM. Achieved accuracy of this method is up to 95%.

In this paper, a new approach is developed and evaluated for automatic analysis of single lead electrocardiogram (ECG) for human recognition and individual identification. This approach is depended on analytic (Amplitude, Time and Width) and modeling (AR) features extracted from the ECG beat. Obtained results indicate high level of accuracy and shorter processing needed to identify the individuals. Eighteen analytic and modeling features are extracted to identify individuals. K nearest neighbor (k-nn) classification and Support Vector Machine (SVM) algorithms are applied in order to classify those features and evaluate the proposed approach. ECG feature selection and extraction using AR modeling has recently been used [18] resulting in accurate classification of various arrhythmia and ventricular arrhythmia conditions.

The remainder of this paper is organized as follows. Section 2 gives a brief description of the techniques used in the pre-processing phase to clean ECG signals of noise and other artefacts. Section 3 provides a review of QRS detection method used in this work. Feature selection and extraction methods are discussed in Section 4 whilst Section 5 contains experimental results and discussion of those results. Conclusions are presented in Section 6.

2. PRE-PROCESSING

ECG signal is recorded by attaching electrodes to different places on the skin, such as chest, legs, arms and neck [19]. Collected ECG data usually contain noise components of low-frequency caused by driftline wonder and high frequency components caused by power line interferences [20]. The presence of noise will corrupt the signal and make the feature extraction and

classification process more difficult and less accurate. Many research papers have discussed the removal of noise and power line interference from the ECG signals. In [21] a non linear adaptive method to eliminate power line interference from the ECG signals is presented. The wavelet coefficient threshold based hyper shrinkage function was used in [22] to detrend the raw ECG signals. In [23] a simplified lattice based adaptive IIR notch filter has been suggested to remove power line interference. Digital FIR filters has been used for the power line noise reduction [24].

In this work, method proposed in [25] has been applied in order to decrease the effect of driftline wonder and preserve the information about the physiology of an individual's ECG. The raw ECG signals are first downsampled to the lowest sampling frequency of the ECG signals from the set (1000 Hz in this case) and then filtered through the 6th order high pass Butterworth filter. Figures 1a and 1b illustrate the performance of the applied Butterworth filter in the pre-processing stage of the proposed system. The output of the Butterworth filter is then passed through power line interferences filer. Here, another Butterworth filter – 10th order low pass type is used to remove high frequency components and noise. Figures 1c and 1d show the signals after the power line interference filtering stage.

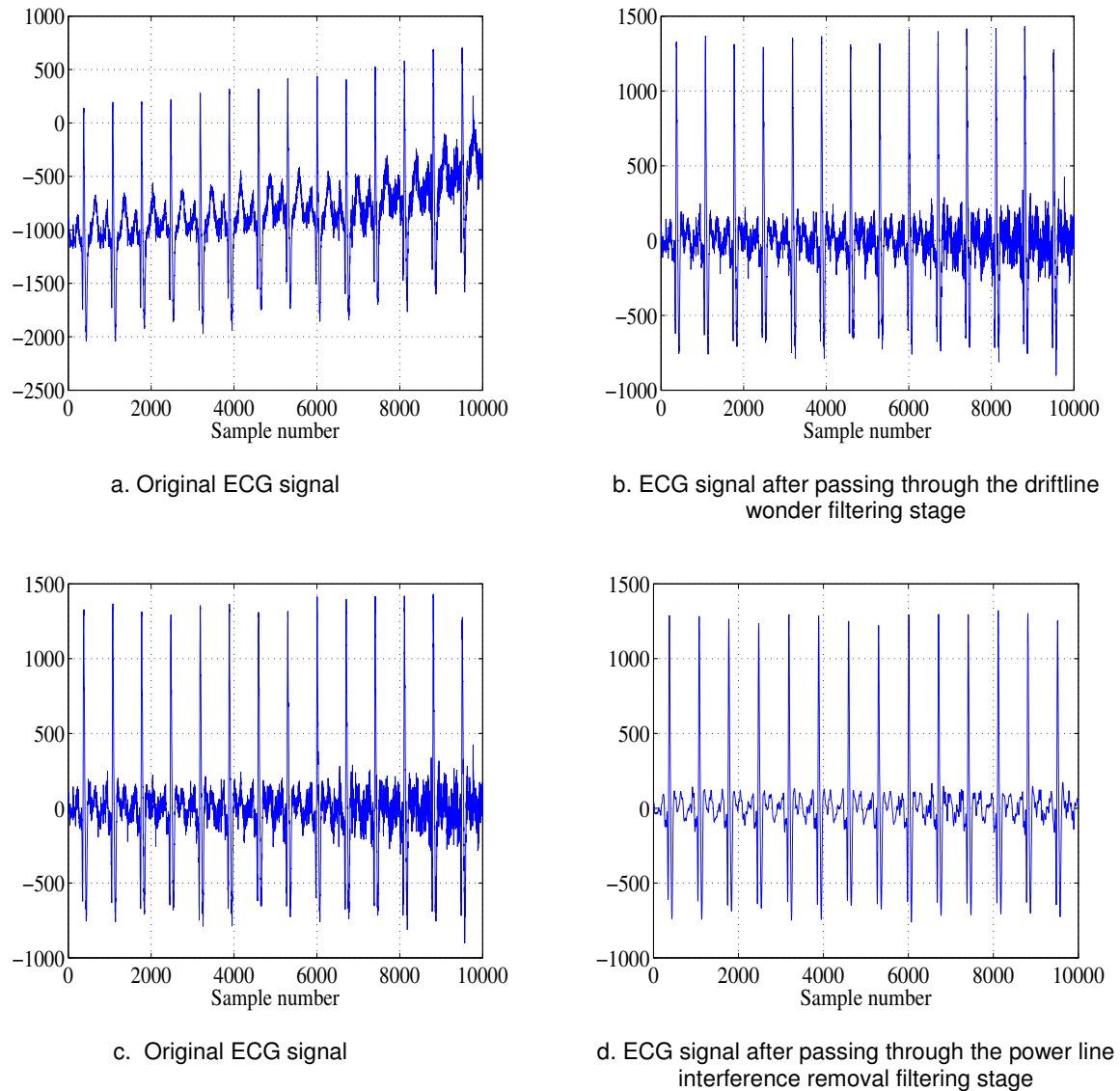


FIGURE 1: Driftline wonder and power line interference filtering stages.

3. QRS DETECTION PHASE

The QRS complex is the most important feature of the ECG signal. Without the accurate knowledge of the QRS complex location, P and T waves are hard to detect and distinguish from each other. Most of the QRS detection methods rely on filtering stage followed by averaging and thresholding [26] in order to distinguish between noise signal and the QRS complex. Threshold value is usually chosen according to the peak height or peak location of ECG signal [27]. A research paper has investigated the problem of accurate QRS complex detection depending on the applications of wavelet filter banks [28]. The short-time Fourier transform (STFT) was also employed after ECG filtering stage in order to detect QRS complex [29].

In this work, Filter bank method employs a bank of linear phase filters to decompose the ECG signal into subbands with uniform frequency bandwidths in order to account for the ECG signal energy distribution in the frequency domain [28]. Processing can be performed on each subband independently. The filterbank used to analyse ECG signals contains M analysis and M synthesis filters, each of length L and is used to produce the subband signals $U_l(z)$ by filtering the input signal $X(z)$:

$$U_l(z) = H_l(z)X(z) \quad l = 0, 1, \dots, M - 1 \quad (1)$$

where $H_l(z)$ represents the transfer function of each bandpass filter. The set of synthesis filters can then combine the processed subbands to result in a processed version of the input signal. Thus, a filter bank-based algorithm involves decomposing a signal into frequency subbands, processing these subbands according to the application at hand, and then sometimes reconstructing the processed subbands. Many scenarios deal with signals which contain specific energy distributions in the frequency domain. For example, with regard to the ECG, a significant proportion of the energy from the QRS complex extends to a frequency of 40 Hz, and even more if the Q, R, and S waves have very sharp morphologies. The P and T waves, in general, have a significant proportion of their energy only up to 10 Hz.

The effective bandwidth of $U_l(z)$ is $\frac{\pi}{M}$ and, thus, it can be downsampled to reduce the total rate. The downsampling process keeps one sample out of M samples. The downsampled signal $W_l(z)$ is:

$$W_l(z) = \frac{1}{M} \sum_{k=0}^{M-1} U_l\left(z^{\frac{1}{M}} W^k\right) \quad (2)$$

where $W = e^{-j\left(\frac{2\pi}{M}\right)}$. The subbands $U_l(z)$ and $W_l(z)$ are bandpassed versions of the input $X(z)$ and also $W_l(z)$ has a lower rate than $U_l(z)$. The filtering process can be efficiently done at $\frac{1}{M}$, the input rate by taking advantage of the downsampling. This process is referred to as the polyphase implementation and contributes to the computational efficiency of the filter bank-based algorithm. A variety of features, indicative of the QRS complex have been extracted from the signal. Figure 2 shows QRS detection of this method by combining the subbands of interest. For example, feature P_1 corresponding to the energy in those subbands 1, 2 and 3, can be computed as:

$$P_1 = \sum_{l=1}^3 |W_l(z)| \quad (3)$$

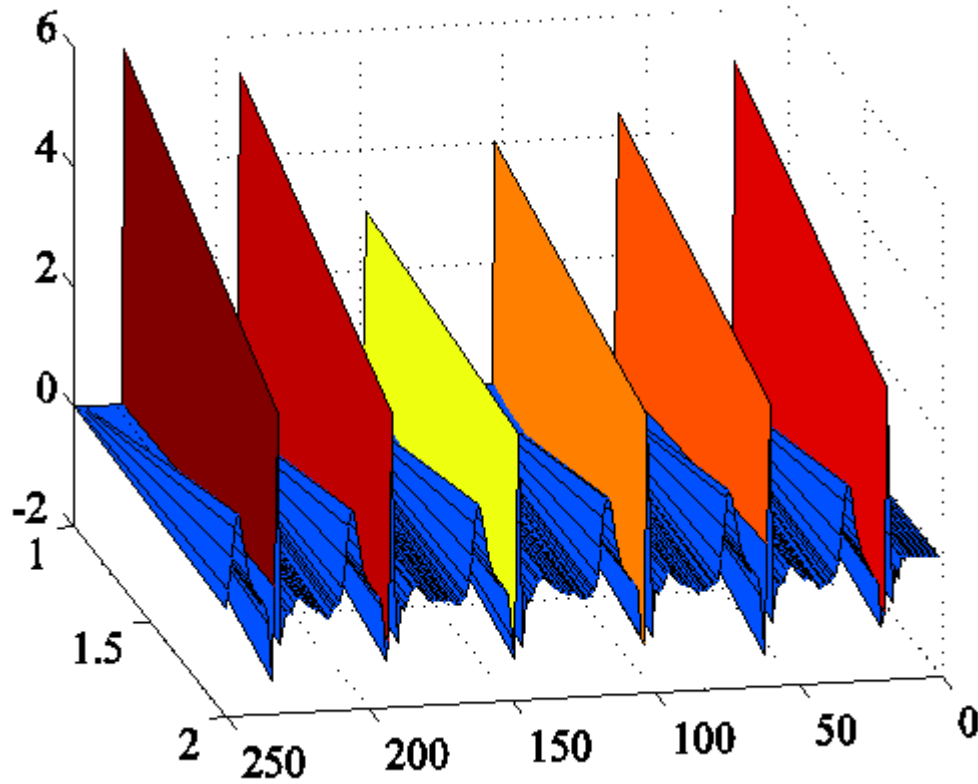


FIGURE 2: Filtered ECG signal with detected QRS complexes (*) using filter bank method.

In the ECG signals, P and T waves occur before and after the QRS complex respectively. The frequency components of QRS complex range between 10 Hz and 25 Hz. In this work, QRS detection plays an important stage as it is used to calculate RR interval of ECG stream, the main indicator for the window estimation process. Window estimation is applied in order to extract windows such that each window contains single beat of ECG windows. If $X(t)$ is an ECG stream of beats, window of single ECG beat is estimated based on the RR interval between any two consecutive R peaks “*” detected from the ECG stream and shafted by half of the RR interval.

$$W_B(t) = \sum_{i=1}^B X(t + 0.5RR_i) + X(t - 0.5RR_i) \quad (4)$$

where RR is the interval between two consecutive R beats, $W_B(t)$ is a window that contain single beat and B is the number of beats in the ECG stream. Figure 3 shows four different samples of window estimation process, each window is plotted in 3D view and contains only one beat.

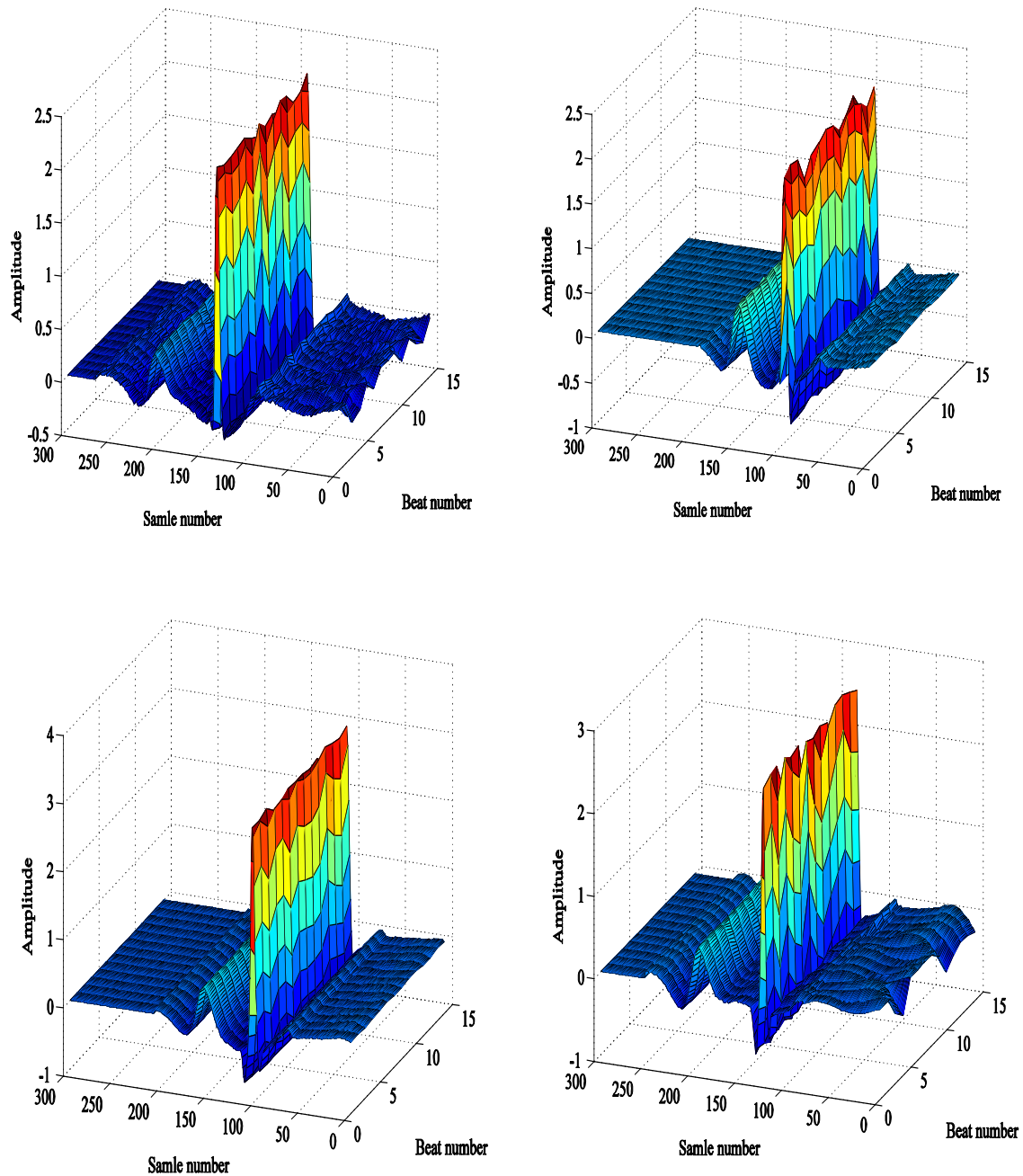


FIGURE 3: Window estimation for four subjects from MIT/BIH database.

4. METHODOLOGY

Biometrics-based human identification is essentially a subject and beat recognition problem which involves pre-processing, feature extraction and classification stages. ECG feature extraction plays a significant role in diagnosing most of the cardiac diseases. Feature extraction scheme proposed in this work uses amplitudes, time intervals, width and AR parameters of the ECG signal for subsequent analysis and identification. For the purpose of comparative study, the system proposed in this work follows the procedure of extracting eighteen features from each ECG beat listed in Table 1.

Analytic	Temporal	PQ	Time interval between P and Q waves.	PS	Time interval between P and S waves.
		PR	Time interval between P and R waves.	TS	Time interval between T and S waves.
		TQ	Time interval between T and Q waves.	TR	Time interval between T and R waves.
		PT	Time interval between P and T waves.		
	Amplitude	P	Amplitude height of P wave.	T	Amplitude height of T wave.
		S	Amplitude height of S wave.	R	Amplitude height of R wave.
		Q	Amplitude height of Q wave.		
	Width	QRS	Width in time of QRS.	T	Width in time of T wave.
		P	Width in time of P wave.		
Modeling	AR coefficients	a_1	First coefficient of AR model.	a_3	Third coefficient of AR model.
		a_2	Second coefficient of AR model.		

TABLE 1: List of extracted features.

4.1 Analytic Features

Analytic features are usually being extracted by analysing and localizing certain fiducial points from the ECG signals as shown in Figure 4. In order to extract analytic features of ECG window or single beat, procedure similar to [30] is followed. As the R peak is already detected using QRS detection, Q, S, P and T amplitudes are localized by finding local positive and negative points separately around R positions of each window. As example, the maximum positive to the left of ECG beat in Figure 4 is the P wave amplitude, and the maximum negative to the left is Q wave amplitude. Same procedure is followed to find T and S amplitudes on the right side. Previously proposed methods [10], [11] and [12] for ECG-based identity recognition use temporal attributes as well as amplitude distances between detected fiducial points. In this work temporal, amplitude and width distances between indicated points have been fitted into k-nn and SVM classification algorithms to find subject and beat classification rate. Temporal and width extracted parameters are calculated using fiducial points in Figure 4, e.g. the time interval between P_{start} and P_{end} is the P wave width. Similar method is used to estimate QRS complex width, the interval between Q_{start} and S_{end} and so on to estimate all the analytic features in Table 1.

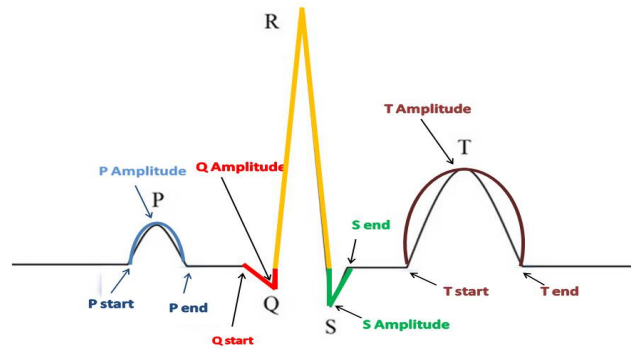


FIGURE 4: Fiducial points of ECG signal.

4.2 Modeling Features

In addition to analytic ECG features described in the previous section, proposed system uses modeling features obtained through the modeling of two or more successive ECG beats. To achieve this, a discrete form of an autoregressive (AR) signal model of order p , $AR(p)$, is applied [31] to model the measured ECG signal. The estimated coefficients are then added to a feature set used in the final, classification stage of the proposed system.

Using the proposed AR model, a sampled signal sequence $y(n)$ can be represented by the relationship:

$$y(n) = a_1y(n-1) + a_2y(n-2) + \dots + a_py(n-p) + \varepsilon(n) \quad (5)$$

where a_k ($k = 1, 2, \dots, p$) are the model coefficients used in the classification process, the $\varepsilon(n)$ is a white noise series and σ^2 is the innovation process with zero mean and variance. Accurate modeling of two successive beats from ECG signal can be achieved using a 3rd order AR model as described in [18]. Individual section of a pre-processed ECG signal in red and the corresponding AR(3) model in blue are shown in Figure 5. High accuracy of obtained AR models can be observed from this plot. a_1 , a_2 and a_3 AR model coefficients are extracted from each beat of ECG stream and fitted into k-nn and SVM with analytic features in the classification process.

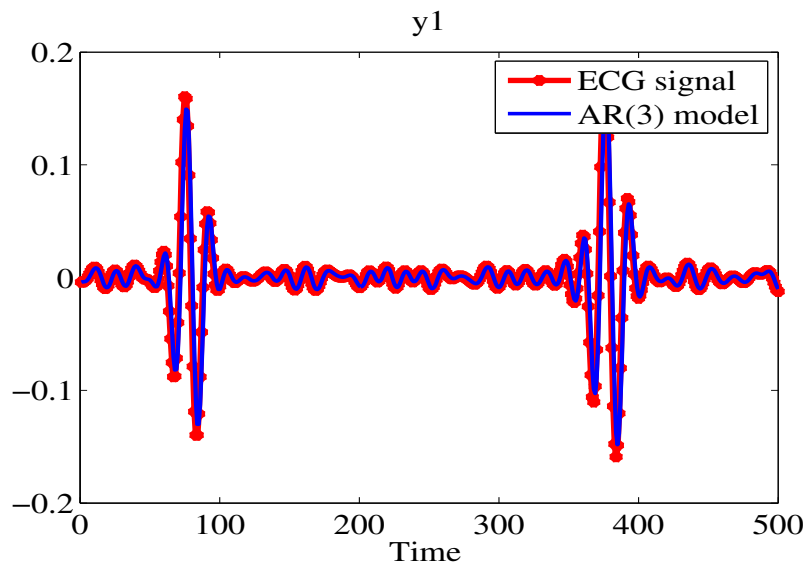
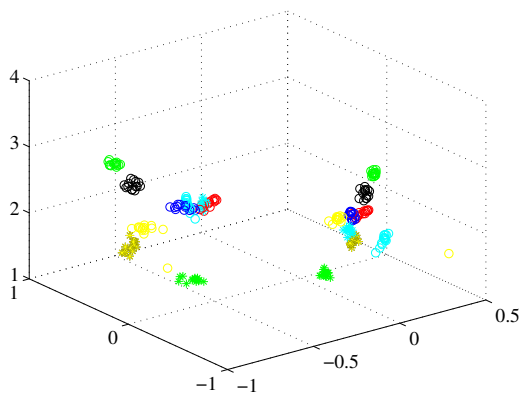


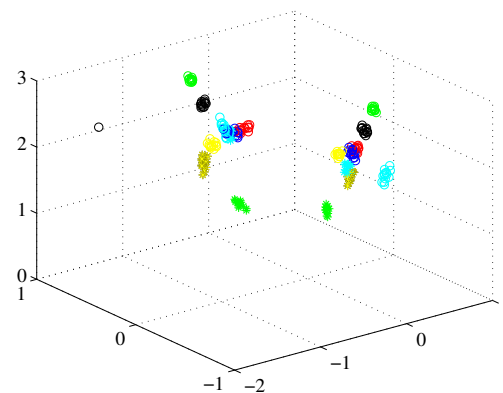
FIGURE 5: AR model and original ECG signal.

5. EXPERIMENT RESULTS

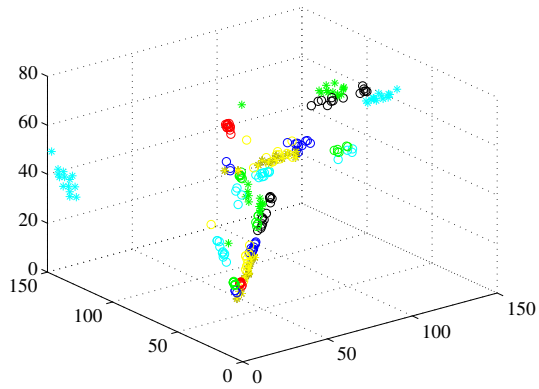
To evaluate the performance of the proposed methods, experiments are conducted using signals from three sets of publicly available ECG databases: PTB [32], MIT-BIH [33] and Milano [34]. The PTB database is provided by the National Metrology Institute of Germany and contains 549 records from 294 subjects. The MIT-BIH database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Laboratory between 1975 and 1979. Each record from this database consists of conventional 12 and 3 leads ECG. The duration of the recordings vary for each subject. Milano database consists of 23 recordings about 30 minutes long, from 23 young normal subjects. Recordings consist of the three orthogonal ECG leads. A concerted effort was made to ensure the maximum fidelity of recordings. A subset of 13 subjects was selected to test proposed method. The selected subjects did not exhibit significant arrhythmias. Some of ECG records held in the database contain significant amount of noise and other artifacts which can significantly reduce the valid heartbeat information. Most severely affected records were therefore not used in the performed tests. Since the database only contains one record for each subject, records were partitioned into two halves. The first half was then used as a “training set” and the second half as a “test set” for classification stage. The training and test sets should in the final evaluation of the proposed system be an ECG recordings made on different days for the same subject. The feature matrix for classification is formed of 18 extracted parameters, described in the earlier sections and 14 windows or beats for each subject. By using ECG stream of 8 seconds length with no overlapping, different number of beats are separated into windows where each window has one beat. Several different window lengths tested in this work show approximately the same classification performance as long as full multiple ECG beats are present in the extracted window. Figure 6 shows the scatter plot of extracted parameters including amplitude, time, width and AR coefficients for both parts (training and test) of each ECG subject in the database. Each color represents extracted parameters for individual ECG sample.



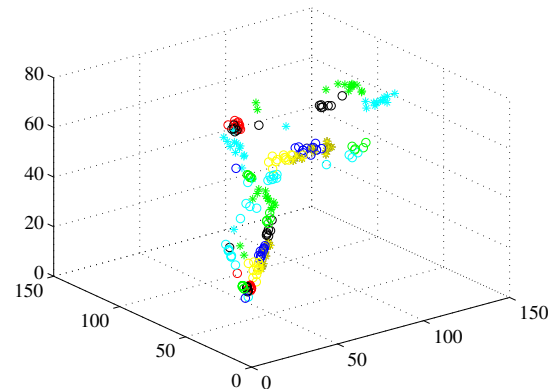
a) Amplitude (training)



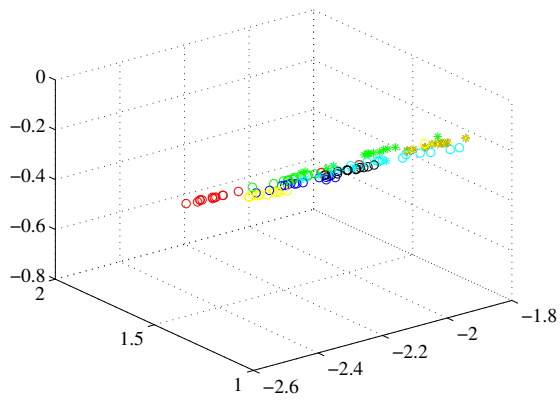
Amplitude (test)



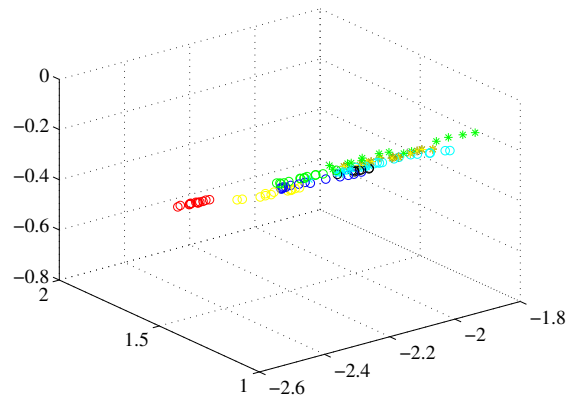
b) Time (training)



Time (test)



c) AR (training)



AR (test)

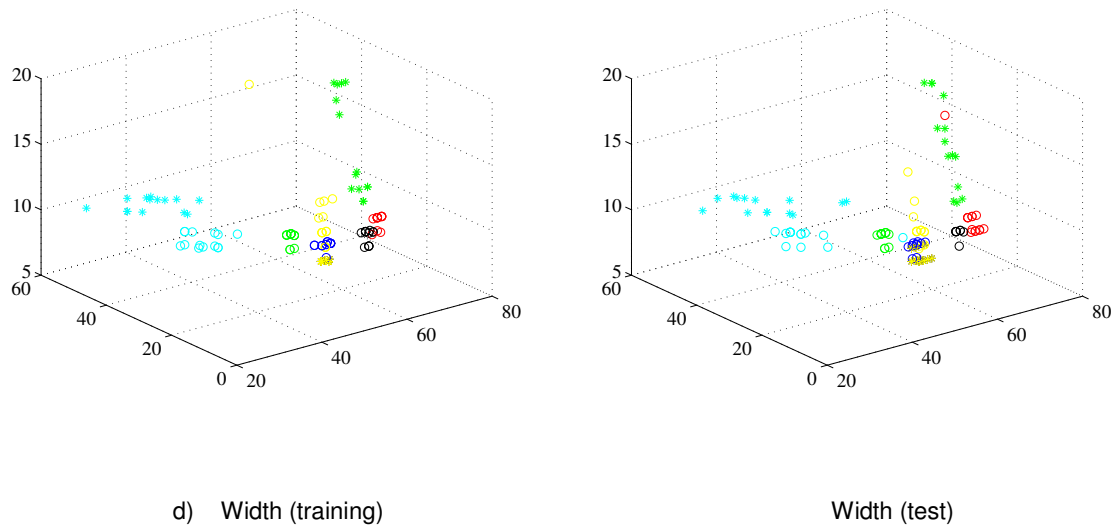


FIGURE 6: Extracted parameters in 3D feature space.

5.1 Individual Subject Recognition using k-nn

To evaluate the performance of the proposed system, k-nn classification algorithm has been applied using “knnclassify” function from classification toolbox in Matlab. The k-nearest neighbor is a statistical classification algorithms used for classifying objects based on closest training examples in the feature space. K-nn is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification. The k-nn is amongst the simplest of all machine learning algorithms: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors. The “knnclassify” function can be called using syntax: “C=knn(training, test, classes)”. In this case “C” shows the classification factors that express algorithm performance given in Table 2. “Training” and “test” sets are different 8 sec time intervals of each subject and parameter “classes” specifies the number of subjects included in this test. Both training and test sets represent the ECG recordings taken on different days on the same set of subjects. Training set includes 18 extracted parameters and 14 windows or beats for 13 subjects, where test set includes another 18 parameters from other 14 windows or beats for the same 13 subjects.

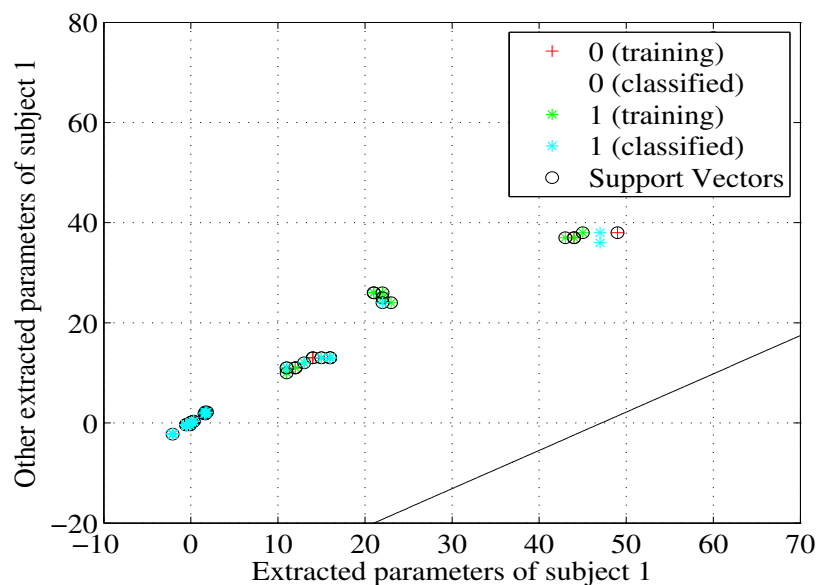
The obtained results presented in Table 2 show that 100% accuracy of subject classified can be obtained if all extracted features are used in classification stage. High identification accuracy comes from the matching parameters between the training and test sets of extracted features for the 13 subjects used in this experiment. A high percentage of negative productive values are obtained when smaller number of parameters is used in tests, e.g. amplitude and AR parameters only. Negative productive value indicates the performance of a diagnostic testing procedure. It is defined as the proportion of subjects with some bias compared to the average of other readings that are correctly classified. Table 2 gives the correct identification rate and illustrates comparison of experimental results for different sets of parameters used in training and test sets combination. Suggesting width parameters alone to classify 13 subjects illustrate very low correct rate (55.56%). It is important to note that amplitude and AR parameters can reach a high rate of classification (98.89%) although high percentage of negative test through the procedure is recorded. By using all extracted features (Amplitude, Time, Width and AR model) to identify individual, it is possible to reach 100 % accuracy of classified subjects with a reasonable negative test value.

Classification Type	Correct Rate	Negative Productive Value (NP)
Width	55.56%	33%
Amplitude	98.89%	86%
Amplitude and AR	95%	87%
Amplitude, Time, Width and AR model	100%	17%

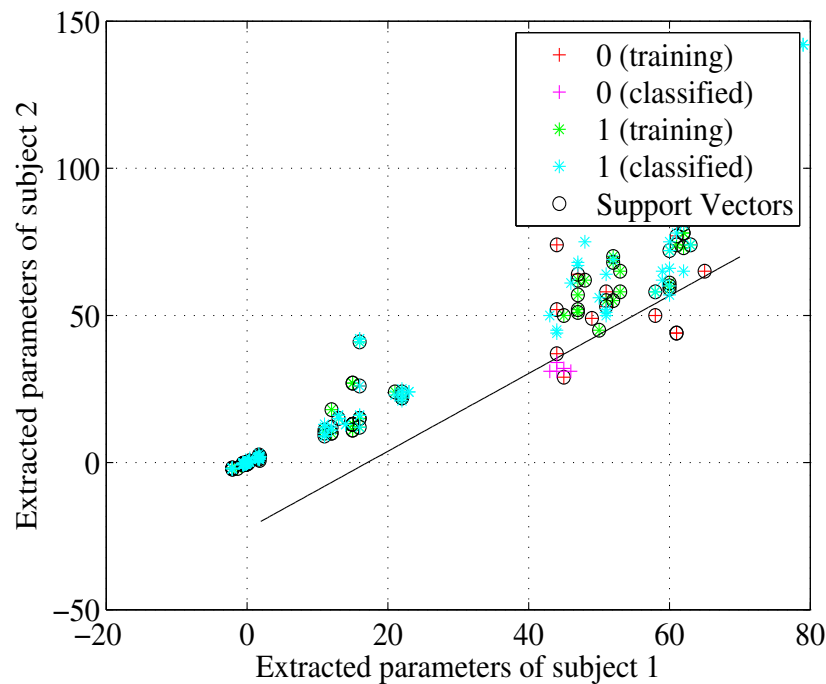
TABLE 2: Classification results of biometric system using k-nn.

5.2 Individual Subject Recognition using SVM

In addition to high accuracy classification between subjects, further work of Support Vector Machine (SVM) algorithm is applied to estimate classification rate for individual subjects. In order to perform individual recognition of subject, all extracted parameters for single subject are compared with extracted parameter of the same subject and with extracted parameter of another subject. Figure 7a illustrates no classification between extracted parameters of subject 1 and the extracted parameters of the same subject, as the SVM margin line is lower than all extracted parameters in free space dimension. While, Figure 7b shows noticeable classification rate between extracted parameters of subject 1 and the extracted parameters of subject 2, SVM margin line is located between the extracted parameters.



a. Classification results of extracted parameters for the same subject



b. Classification results of extracted parameters for two different subjects.

FIGURE 7: SVM classification results of individual subject recognition.

5.3 Subject Beat Recognition using k-nn and SVM

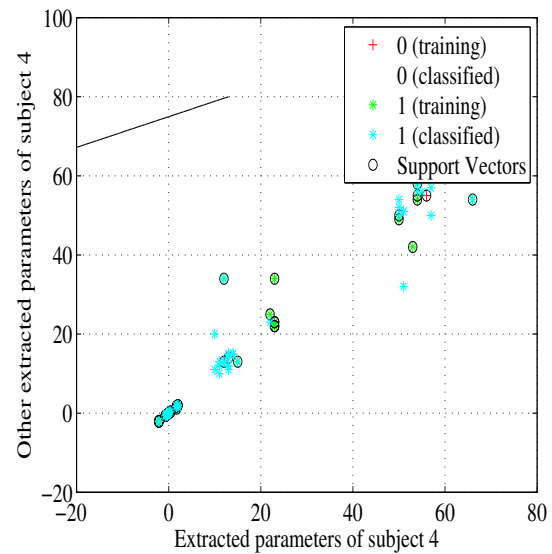
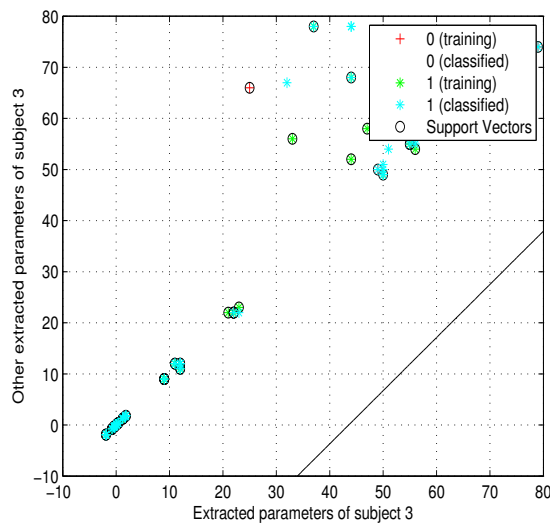
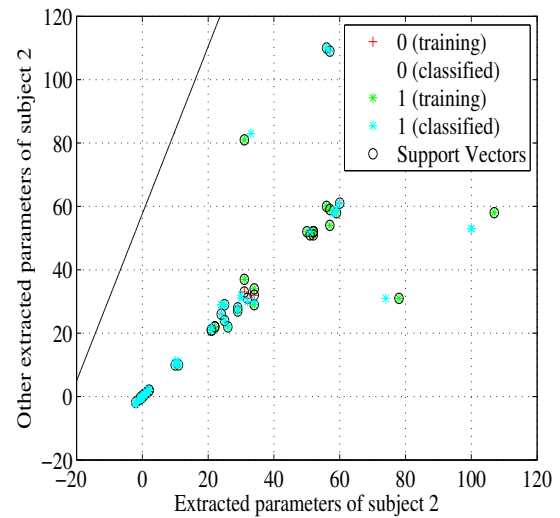
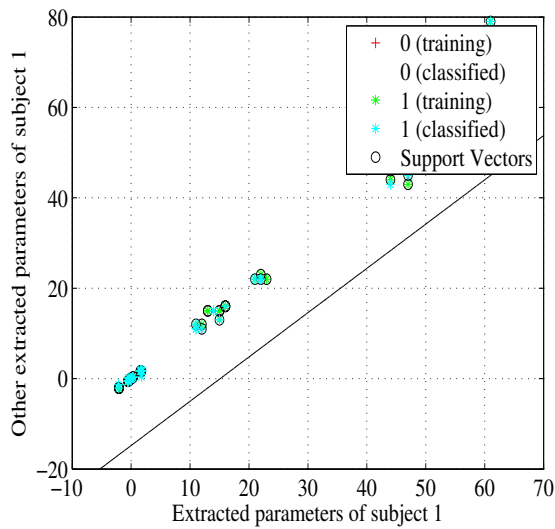
In addition to the high classification rate of individual subject recognition using k-nn and SVM, this work investigates the matching and similarity between the extracted parameters from the individual beats of the same and different subjects. Similarity between extracted parameters of two beats of the same subject proves the high classification rate of this approach. In this test, 14 beats from 5 different subjects have been classified using k-nn and SVM to show the matching between the extracted parameters of beats for the same and different subjects. The extracted parameters from the first seven beats of each subject are used as “test” in the k-nn and SVM classification algorithms, while the parameters extracted from the other seven beats are used as “training” set. Table 3 shows correct classification rate between extracted parameters from beats for the same subject using k-nn, while the classification rate is very low between the extracted parameters of different subjects.

Subjects	1	2	3	4	5
1	100 %	55 %	54%	58 %	44 %
2	55 %	100 %	52 %	56 %	54 %
3	54 %	52 %	90.5 %	48 %	59 %
4	58 %	56 %	48 %	100 %	57 %
5	44 %	54 %	59 %	57 %	91 %

TABLE 3: k-nn classification results of beats for the same and different subjects.

SVM classification algorithm is also used to evaluate the similarity and matching between the extracted parameters of beats for the same subject. Figure 8 shows the similarity and matching of five subjects to identify beat recognition. Results show high similarity and matching between test

and training data selected for this test and lead SVM margin line to locate out of classification area.



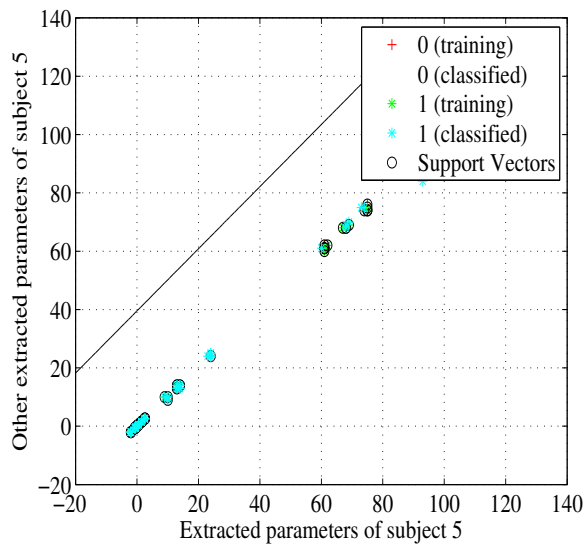


FIGURE 8: Classification results of beats for the same subject using k-nn and SVM.

In order to evaluate the goodness of using AR models to design ECG biometric algorithm, correct classification rate of this method has to be compared with the other correct classification rates of ECG biometric approaches mentioned in the introduction. Table 4 shows that the correct classification rate to identify humans based on their ECG signal using AR model can achieve up to 100 % classification accuracy. This high classification rate is achieved due to the high accuracy of modeling to the ECG signal using AR models, see Figure 5.

ECG Biometric method	Correct classification rate %
Wavelet transform	98
Kohonen self-organizing maps (SOM) and Learning Vector Quantization (LVQ)	94
Artificial Neural Networks (ANN) and Linear Discriminant Analysis (LDA)	85
Hidden Markov Model (HMM)	95
Autoregressive (AR) model	100

TABLE 4: Evaluation table of correct classification for different ECG biometric methods.

The ability to build accurate ECG biometric system using modeling techniques depends on the accuracy to model ECG signal. Autoregressive Moving Average (ARMA) model could be investigated as an extraction approach for more advanced methods of modeling to ECG signal. The extracted parameters from applied ARMA model could be used in the classification algorithm in order to achieve high correct rate of classification.

6. CONCLUSIONS

A biometric system for automatic analysis of a single lead electrocardiogram (ECG) for human identification is presented in this work. The first stage of this system consists of a band-pass Butterworth filter used to remove noise and other artifacts present in the raw ECG signal. The R peaks of an ECG stream are then localized by filtering a signal through the bank of filters and analysing each filtered signal component individually. The individual heart beats of an ECG record are aligned along the R peak position and truncated by a window size based on the heart rate variability (HRV) of individuals' ECG. The temporal, amplitude, width and AR parameters extracted from each ECG signal are used to identify individual ECGs. An approach of modeling two or more successive ECG beats, using a discrete form of an autoregressive (AR) signal model of order p is applied to extract AR modeling features, $AR(p)$. A subset of 13 subjects was selected in this work in order to test proposed system. All the extracted parameters for each subject are fitted into k-nn and SVM classification algorithm to classify and eventually identify "test" parts of each signal. 100% identification accuracy for 13 different ECG recordings has been achieved in this project using k-nn and SVM classification algorithms, when each ECG signal is split into "train" and "test" section. Results presented in this work indicate that the ECG based system might lead to an accurate biometric identification in some practical applications.

7. REFERENCES

- [1] Y. L. Singh and P. Gupta (2008), ECG to Individual Identification, Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on, pp. 1 - 8.
- [2] S.-H. Lin (2000), An Introduction to Face Recognition Technology, Informing Science special Issue on Multimedia Informing Technologies , vol. 3, no. 1, pp. 1-7.
- [3] Y. Wang, F. Agraftioti and D. Hatzinak (2008), Analysis of Human Electrocardiogram, Hindawi Publishing Corporation , EURASIP Journal on Advances in Signal Processing, pp. 11.
- [4] T. Shen and W. Tompkins (2002), One-lead ECG for identity verification, in: Proceedings of the 24th Annual International Conference of IEEE EMBS, pp. 62-63.
- [5] K. Jain, A. Ross and S. Prabhakar (2004), An introduction to biometric recognition, IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 1, pp. 4-20.
- [6] R. Hoekema and G. J. H. Uijen (2001), Geometrical aspects of the interindividual variability of multilead ECG recordings, IEEE Transactions on Biomedical Engineering, vol. 48, no. 5, pp. 551–559.
- [7] I. Eisenstein, J. Edelstein and R. Sarma (1982), The electrocardiogram in obesity, J. Electrocardiol., vol. 15, no. 2, pp. 115–118.
- [8] S. Safie, J. Soraghan and L. Petropoulakis (2011), ECG biometric authentication using Pulse Active Width (PAW), Biometric Measurements and Systems for Security and Medical Applications (BIOMS), 2011 IEEE Workshop on, pp. 1 - 6.
- [9] I. Odinaka, A. Kaplan, J. O'Sullivan and L. Po-Hsiang (2012), ECG Biometric Recognition: A Comparative Analysis, IEEE Transactions on Information Forensics and Security, vol. 7, no. 6, pp. 1812 - 1824.

- [10] L. Biel, O. Pettersson, L. Philipson and P. Wide (2001), ECG analysis: A new approach in human identifications, IEEE Trans. Instrum. Meas., vol. 50, no. 33, pp. 808–812.
- [11] M. Kyoso and A. Uchiyama (2001), Development of an ECG identification system, in Proc. 23rd Ann. EMBS Int. Conf.
- [12] J. Irvine, B. Wiederhold, L. Gavshon and S. Israel (2001), Heart rate variability: A new biometric for human identification, Proc. Int. Conf. Artificial Intelligence (ICAI 2001).
- [13] S. Mahmoodabadi, A. Ahmadi and M. Abolhas (2005), ECG Feature Extraction using Daubechies Wavelets, Proceedings of the fifth IASTED International conference on Visualization, Imaging and Image Processing, pp. 343-348.
- [14] P. Tadejko and W. Rakowski (2007), Mathematical Morphology Based ECG Feature Extraction for the Purpose of Heartbeat Classification, 6th International Conference on Computer Information Systems and Industrial Management Applications, CISIM '07, pp. 322-327.
- [15] C. Alexakis, H. O. Nyongesa, R. Saatchi and N. D. Har (2007), Feature Extraction and Classification of Electrocardiogram (ECG) Signals Related to Hypoglycaemia, Conference on computers in Cardiology, pp. 53.
- [16] P. Hee-Soo and W. Soo-Min (2009), ECG Pattern Classification Based on Generic Feature Extraction, Proceedings of the 3rd WSEAS Int. Conf. on circuits, systems, signal and telecommunication, (CISST'09) ISSN:, vol. 21.
- [17] W. Liang, Y. Zhang and J. Tan (2009), A Novel Approach to ECG Classification Based upon Two-Layered HMMs in Body Sensor Networks, Multidisciplinary Digital Publishing Institute, vol. 14, pp. 5994-6011,.
- [18] B. Vuksanovic and M. Alhamdi (2013), ECG Based System for Arrhythmia Detection and Patient Identification, ITI2013, pp. 1-7.
- [19] D. Gari, A. Francisco and M. Patrick (2006), Biomedical Engineering. :Advanced Methods And Tools for ECG Data Analysis, Artech House, Inc.
- [20] I. Duskalov (1998), Developments in ECG acquisition, preprocessing, parameter measurement, and recording, IEEE Engineering in Medicine and Biology Magazine, vol. 7, no. 2, pp. 50 - 58.
- [21] A. Ziarani and A. Konrad (2010), Non linear Adaptive method of elimination of power line interference in ECG signals, IEEE Transactions on Biomedical Eng, vol. 49, no. 6, pp. 540-544.
- [22] S. Pooranchandra and N. Kumaravel (2010), A novel method for elimination of power line frequency in ECG signal using hyper shrinkage functions, Digital Signal Processing, vol. 18, no. 2, pp. 116-126.

- [23] S. Dhillon and S. Chakrabarti (2001), Power Line Interference removal From Electrocardiogram Using A Simplified Lattice Based Adaptive IIR Notch Filter, Proceedings of the 23rd Annual EMBS International conference, Istanbul, Turkey.
- [24] M. Chavan, R. Aggarwala and M. Uplane (2008), Interference reduction in ECG using digital FIR filters based on Rectangular window, WSEAS Transactions on Signal Processing, vol. 4, no. 5, pp. 340-49.
- [25] S. K. Jagtap (2012), The impact of digital filtering to ECG analysis: Butterworth filter application, Communication, Information & Computing Technology (ICCICT), 2012 International Conference on, pp. 1-6.
- [26] H.-Y. Zhou and K.-M. Hou (2008), Embedded real-time QRS detection algorithm for pervasive cardiac care system, 9th International Conference on Signal Processing, pp. 2150 - 2153.
- [27] J. F. Moraes (2002), QRS complex detection algorithm using electrocardiogram leads, Computers in Cardiology, pp. 205 - 208.
- [28] G. Kokturk (1998), A real-time simulated QRS detection system constructed using wavelet filtering technique, IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, pp. 281 - 284.
- [29] N. Uchaipichat and S. Inban (2010), Development of QRS Detection using Short-time Fourier Transform based Technique, IJCA Journal.
- [30] S. A. Israel, J. M. Irvine and A. Cheng (2005), ECG to identify individuals, Pattern Recognition, pp. 133-142.
- [31] C. Anderson, E. Stolz and S. Shamsunder (1995), Discriminating Mental Tasks Using EEG Represented by AR Models, Engineering in Medicine and Biology Society, IEEE 17th Annual Conference , vol. 2.
- [32] A. Goldberger, L. Amaral, L. Glass and Hausdorff (2000), Components of a New Research Resource for Complex Physiologic Signals., PhysioBank, PhysioToolkit, and PhysioNet:, vol. 23, no. 21, pp. 101.
- [33] R. Mark (1990), The MIT-BIH Arrhythmia Database on CD-ROM and software for use with it, IEEE in Computers in Cardiology 1990, Proceedings.
- [34] D. Milano (1992), Politecnico Biosignals Archives on CD-ROM, (Copyright (C) Politecnico. 1992).

INSTRUCTIONS TO CONTRIBUTORS

The *International Journal of Biometric and Bioinformatics (IJBB)* brings together both of these aspects of biology and creates a platform for exploration and progress of these, relatively new disciplines by facilitating the exchange of information in the fields of computational molecular biology and post-genome bioinformatics and the role of statistics and mathematics in the biological sciences. Bioinformatics and Biometrics are expected to have a substantial impact on the scientific, engineering and economic development of the world. Together they are a comprehensive application of mathematics, statistics, science and computer science with an aim to understand living systems.

We invite specialists, researchers and scientists from the fields of biology, computer science, mathematics, statistics, physics and such related sciences to share their understanding and contributions towards scientific applications that set scientific or policy objectives, motivate method development and demonstrate the operation of new methods in the fields of Biometrics and Bioinformatics.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJBB.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with Volume 9, 2015, IJBB will appear with more focused issues related to biometrics and bioinformatics studies. Besides normal publications, IJBB intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

LIST OF TOPICS

The realm of International Journal of Biometrics and Bioinformatics (IJBB) extends, but not limited, to the following:

- Bio-grid
- Bioinformatic databases
- Biomedical image processing (registration)
- Biomedical modelling and computer simulation
- Computational intelligence
- Computational structural biology
- DNA assembly, clustering, and mapping
- Fuzzy logic
- Gene identification and annotation
- Hidden Markov models
- Molecular evolution and phylogeny
- Molecular sequence analysis
- Bio-ontology and data mining
- Biomedical image processing (fusion)
- Biomedical image processing (segmentation)
- Computational genomics
- Computational proteomics
- Data visualisation
- E-health
- Gene expression and microarrays
- Genetic algorithms
- High performance computing
- Molecular modelling and simulation
- Neural networks

CALL FOR PAPERS

Volume: 10 - Issue: 1

i. Paper Submission: January 31, 2016

ii. Author Notification: February 28, 2016

iii. Issue Publication: March 2016

CONTACT INFORMATION

Computer Science Journals Sdn Bhd

B-5-8 Plaza Mont Kiara, Mont Kiara

50480, Kuala Lumpur, MALAYSIA

Phone: 006 03 6204 5627

Fax: 006 03 6204 5628

Email: cscpress@cscjournals.org

CSC PUBLISHERS © 2015
COMPUTER SCIENCE JOURNALS SDN BHD
B-5-8 PLAZA MONT KIARA
MONT KIARA
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6204 5627

FAX: 006 03 6204 5628

EMAIL: cscpress@cscjournals.org