

Fiducial feature reduction analysis for electrocardiogram (ECG) based biometric recognition

M. M. Tantawi · K. Revett · A. Salem · M. F. Tolba

Received: 19 April 2012 / Revised: 27 June 2012 / Accepted: 27 June 2012 /
Published online: 20 July 2012
© Springer Science+Business Media, LLC 2012

Abstract Although the electrocardiogram (ECG) has been a reliable diagnostic tool for decades, its deployment in the context of biometrics is relatively recent. Its robustness to falsification, the evidence it carries about aliveness and its rich feature space has rendered the deployment of ECG based biometrics an interesting prospect. The rich feature space contains fiducial based information such as characteristic peaks which reflect the underlying physiological properties of the heart. The principal goal of this study is to quantitatively evaluate the information content of the fiducial based feature set in terms of their effect on subject and heart beat classification accuracy (ECG data acquired from the PhysioNet ECG repository). To this end, a comprehensive set of fiducial based features was extracted from a collection of ECG records. This feature set was subsequently reduced using a variety of feature extraction/selection methods such as principle component analysis (PCA), linear discriminant analysis (LDA), information-gain ratio (IGR), and rough sets (in conjunction with the PASH algorithm). The performance of the reduced feature set was examined and the results evaluated with respect to the full feature set in terms of the overall classification accuracy and false (acceptance/rejection) ratios (FAR/FRR). The results of this study indicate that the PASH algorithm, deployed within the context of rough sets, reduced the dimensionality of the feature space maximally, while maintaining maximal classification accuracy.

M. M. Tantawi (✉) · A. Salem · M. F. Tolba
Faculty of Computer and Information Sciences, Ain Shams University,
Cairo, Egypt
e-mail: manalmt@hotmail.com

K. Revett
Faculty of Informatics and Computer Science, The British University in Egypt,
El Sherouk City, Egypt

Keywords Biometrics · Cognitive biometrics · Electrocardiogram (ECG) · Information gain ratio (IGR) · Linear discriminant analysis (LDA) · Parameterized averaged support heuristics (PASH) algorithm · Principal component analysis (PCA), Rough sets

1 Introduction

Biometrics is a scientific approach to person identification and/or verification based on physiological or behavioral characteristics (motoric or cognitive) of individuals (Revett 2008). Biometric systems have become integrated into the fabric of everyday life—deployed where and whenever secure access to a trusted instrument is required. Physiological biometrics considers anatomical traits that are unique to an individual such as: face, fingerprint and iris, while behavioral biometrics deals with functional traits such as: signature, gait and keystroke (Revett 2008). An alternative and burgeoning branch of biometrics which has been gaining momentum over the past decade is the deployment of biological signals (biosignals) such as the electroencephalogram (EEG) and electrocardiogram (ECG) as biometric traits. The potential benefit of deploying a biosignal based approach to biometric recognition (the basis of *cognitive biometrics*) is that they are potentially more reliable relative to *behavioural* based approaches (i.e. no one can exactly reproduce the same signature or keystroke typing pattern all the time). Lastly, biosignals are difficult to spoof or falsify and liveness is typically not an issue! (Revett 2008; Sufi et al. 2010).

The ECG has been primarily deployed for medical diagnostic purposes. The application of ECG to biometrics began with work by Forsen and colleagues in 1977 (Forsen et al. 1977), which resulted in one of the first publications presenting an implementation of ECG based biometrics on a small cohort of subjects. As with the current approach, Forsen utilized a series of landmarks contained within the ECG that were reproducible and prominent. Visual inspection of a heart beat within an ECG trace reveals three prominent excursions from baseline (see Fig. 2a). These excursions are termed waves—and are labeled P, QRS and T waves, which occur in this temporal order (Sufi et al. 2010). These wave complexes form landmarks that appear in every heart (though it is normal for infrequent variations in some aspects of the waves to occur), which may be deployed as points of references for other more subtle components within an ECG. These landmarks are measured with respect to their latency and amplitude—forming fiducial marks within the ECG heart beat trace. Clearly, one would expect that the within variance is typically much smaller than between variance with respect to the values of fiducial marks, in order to qualify as a potential biometric. The validity of using ECG as a biometric trait is supported by the fact that the physiological and geometric differences of the heart in different individuals show certain uniqueness in their ECG signals (Sufi et al. 2010; Wang et al. 2008). The principal goal of this study is to evaluate just how reproducible and consistent the fiducial components of an ECG are for person authentication. Further, one would like to determine the relative information content of individual fiducial marks in terms of their uniqueness and hence their importance as a biometric feature. This paper will utilize standard biometric quality measures such as the false acceptance/rejection rate (FAR/FRR) to evaluate the feature set elements.

Furthermore, the generalizability of the resulting feature set/classifier combination will be evaluated as a more comprehensive measure of the reliability of the results.

In our previous work (Tantawi et al. 2012), a set of 28 fiducial features (a ‘superset’), representing the vast majority of fiducial features reported in the literature, were generated and used for classification purposes. This feature set consists of time intervals, amplitudes, and angles between fiducial points, all of which can be detected from a typical ECG heartbeat signal. The superset was utilised for both training and testing purposes. This study evaluates the effect of reducing this superset of features and answering four “can” questions: 1) Can the reduced set improve or at least preserve the accuracy achieved by the superset? 2) Can the reduced set provide a stable system? 3) Can the reduced set preserve the achieved FAR & FRR by the superset or improve them? 4) Can the reduced set preserve generalizability? To approach these questions objectively, feature extraction and feature selection techniques were evaluated. For feature extraction, principle component analysis (PCA) and linear discriminant analysis (LDA) methods were utilised. Feature selection was implemented via information gain ratio (IGR) and rough sets based approaches. The rough set approach includes applying the PASH algorithm (Zhang and Yao 2004). It should be noted that this is the first published application of rough sets and PASH on ECG data in the context of biometrics.

The remainder of this paper is organized as follows. Section 2 provides a review of related works. The proposed methods are discussed in Section 3. A detailed description for the experiments and their results is given in Section 4. Section 5 provides results discussion and analysis. Finally, the conclusion is provided in Section 6.

2 Related work

The ECG was first proposed as a biometric by Forsen and colleagues in 1977 (Forsen et al. 1977). Since then, a variety of papers have been published which demonstrate the feasibility of deploying ECG based biometrics (Forsen et al. 1977; Wang et al. 2008; Tantawi et al. 2012; Biel et al. 2001; Shen 2005; Shen et al. 2002; Gahi et al. 2008; Singla and Sharma 2010; Singh and Gupta 2006, 2009; Israel et al. 2005; Irvine and Israel 2009; Guennoun et al. 2009; Wan and Yao 2008; Wao and Wan 2010; Chan et al. 2005, 2008; Fatemian and Hatzinakos 2009; Plataniotis et al. 2006; Agrafioti and Hatzinakos 2008; Khalil and Sufi 2008; Sufi et al. 2008; Li and Narayanan 2010; Coutinho et al. 2010; Ghofrani and Bostani 2010; Odinaka et al. 2010; Tawfik et al. 2010; Ting and Salleh 2010; Venkatesh and Jayaraman 2010; Mai et al. 2011; Ye et al. 2010; Safie et al. 2011a, b). These systems rely on either fiducial features (Forsen et al. 1977; Wang et al. 2008; Tantawi et al. 2012), (Biel et al. 2001; Shen 2005; Shen et al. 2002; Gahi et al. 2008; Singla and Sharma 2010; Singh and Gupta 2006, 2009; Israel et al. 2005; Irvine and Israel 2009; Guennoun et al. 2009) or non fiducial features (Wan and Yao 2008; Wao and Wan 2010; Chan et al. 2005, 2008; Fatemian and Hatzinakos 2009; Plataniotis et al. 2006; Agrafioti and Hatzinakos 2008; Khalil and Sufi 2008; Sufi et al. 2008; Li and Narayanan 2010; Coutinho et al. 2010; Ghofrani and Bostani 2010; Odinaka et al. 2010; Tawfik et al. 2010; Ting and Salleh 2010; Venkatesh and Jayaraman 2010; Mai et al. 2011; Ye et al. 2010; Safie et al. 2011a, b) such as coefficients of wavelets (Wan and Yao 2008; Wao and Wan 2010; Chan et al. 2005, 2008; Fatemian and Hatzinakos 2009);

(Ye et al. 2010), discrete cosine transform (Wang et al. 2008; Plataniotis et al. 2006; Agrafioti and Hatzinakos 2008; Tawfik et al. 2010) and polynomials (Khalil and Sufi 2008; Sufi et al. 2008). Fiducial and non fiducial approaches have their relative benefits and downsides. Generally speaking, extracting fiducial marks is prone to error, which in turn reduces the overall classification accuracy. On the other hand, holistic approaches tend to require more sophisticated algorithms for performing classification as it relies on derived features contained within the data in order to achieve the same classification accuracy as a fiducial based approach. The balance is generally tipped in favour of fiducial based approaches. Regardless of the high level feature extraction methodology deployed (fiducial versus non-fiducial), most published reports yield subject identification accuracy well above 95 %, attesting to the inherent uniqueness contained within ECG feature space, irrespective of the various processing pipelines adopted (Tantawi et al. 2012; Agrafioti et al. 2011). In this work, reduction of fiducial features is considered. Hence, a brief survey of papers deploying different approaches for fiducial features reduction can be provided as follows:

Biel and colleagues employed a set of 30 features extracted by specialized equipment, which was subsequently reduced to 21 by correlation matrix analysis (Biel et al. 2001). The resulting feature set was tested on a dataset of 20 subjects, yielding 95–100 % subject identification accuracy. Shen and colleagues approach were able to distinguish all 20 subjects using 7 features extracted from the QRS wave, utilizing a template matching and decision based neural network (DBNN) for classification (Shen et al. 2002). However, in later work (Shen 2005), Shen et al. increased the feature set to 17 features, which were extracted from the entire heartbeat, and were able to distinguish 168 subjects with 95 % accuracy. An information gain based feature selection algorithm has been utilized by Gahi et al. (2008) to reduce a set of 24 features to 9, 100 % accuracy was achieved on a dataset of 16 subjects using a Mahalanobis criterion for classification. Israel et al. (2005) and Wang et al. (2008) utilized the Wilk's lambda method to reduce fiducial features and deployed linear discriminant analysis (LDA) for the purpose of classification. Israel et al. (2005) utilized a set of 12 features for their classification approach. Experiments on a dataset of 29 subjects yielded 100 % subject identification accuracy and 81 % heartbeat recognition accuracy and importantly, demonstrated that the system was invariant to different anxiety states. On the other hand, Wang et al. (2008) reduced a set of temporal and amplitude features (9 temporal and 6 amplitude), which were derived from the PhysioNet ECG data repository (experiments were carried on PTB and MIT_BIH datasets). The system achieved 100 % subject identification accuracy and 92.4 % and 94.8 % heartbeat recognition accuracy for PTB and MIT_BIH datasets respectively.

These studies indicate that there are unique features contained within an ECG that are characteristic of an individual. in the context of the ECG as a potential biometric, one can make the following observations based on the results of these studies: 1) features were reduced by the application of one or more statistical measures, without providing comparative data; 2) except for Wang et al. (2008), Israel et al. (2005) all of these systems were trained and tested using heartbeats from the same recording session for each subject which does not provide evidence of stability; 3) none of these studies provided any reliable quantitative measure of the ability to reject the heartbeats of any system attacker (false acceptance rate (FAR)

data), and 4) except for Wang et al. (2008) these studies did not address the issue of generalization to other datasets. This paper attempts to address these issues, utilising different feature reduction methods and a publically accessible ECG data repository (the PhysioNet suite of ECG datasets).

3 Methodology

The ECG based authentication system reported in this paper follows the same architecture proposed in Tantawi et al. (2012) but with the addition of a feature reduction stage before the classification stage. Thus, our system encompasses 5 stages as follows: 1) preprocessing; 2) fiducial points detection; 3) feature extraction and normalization; 4) feature reduction; 5) classification (subject identification). Figure 1 depicts the general block diagram for our system.

3.1 Preprocessing & fiducial point detection

A preprocessing stage is needed to reduce noise and to eliminate baseline wandering. A Butterworth filter of second order with cutoff frequencies of 0.5 and 40 Hz is applied for that purpose. After data preprocessing, the detection of fiducial points is a preliminary and crucial stage that includes detecting the peak and the end points of each of the three complexes QRS, P and T, see Fig. 2a. Details about the detection algorithms employed in this paper for detection of fiducial points can be found in Tantawi et al. (2012).

3.2 Feature extraction & normalization

After automated detection of the fiducial points, a set of features was extracted from each heartbeat. In this work, a set of 28 features were utilized, which represents the majority of features utilized in the literature. As shown in Fig. 2b, these features encompass 19 temporal features (distances between fiducial points), 6 amplitude features and 3 angle features. The temporal features are normalized using the full heartbeat duration to avoid heart rate variability effects. Note that the amplitude features were normalized by the R amplitude to avoid signal attenuation, the temporal and angle features were used as raw features. All features were mapped onto the range [0..1] in order to minimize magnitude effects on the classification schemes.

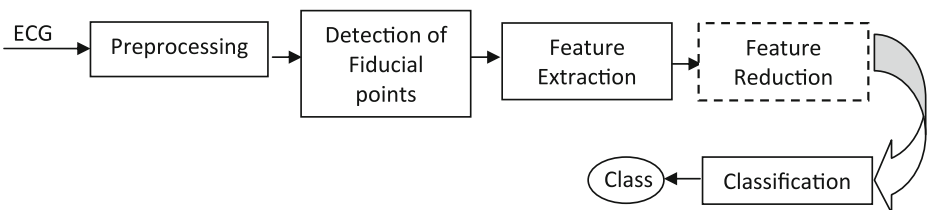


Fig. 1 The processing pipeline schematic deployed in this study. Full details are presented in the text

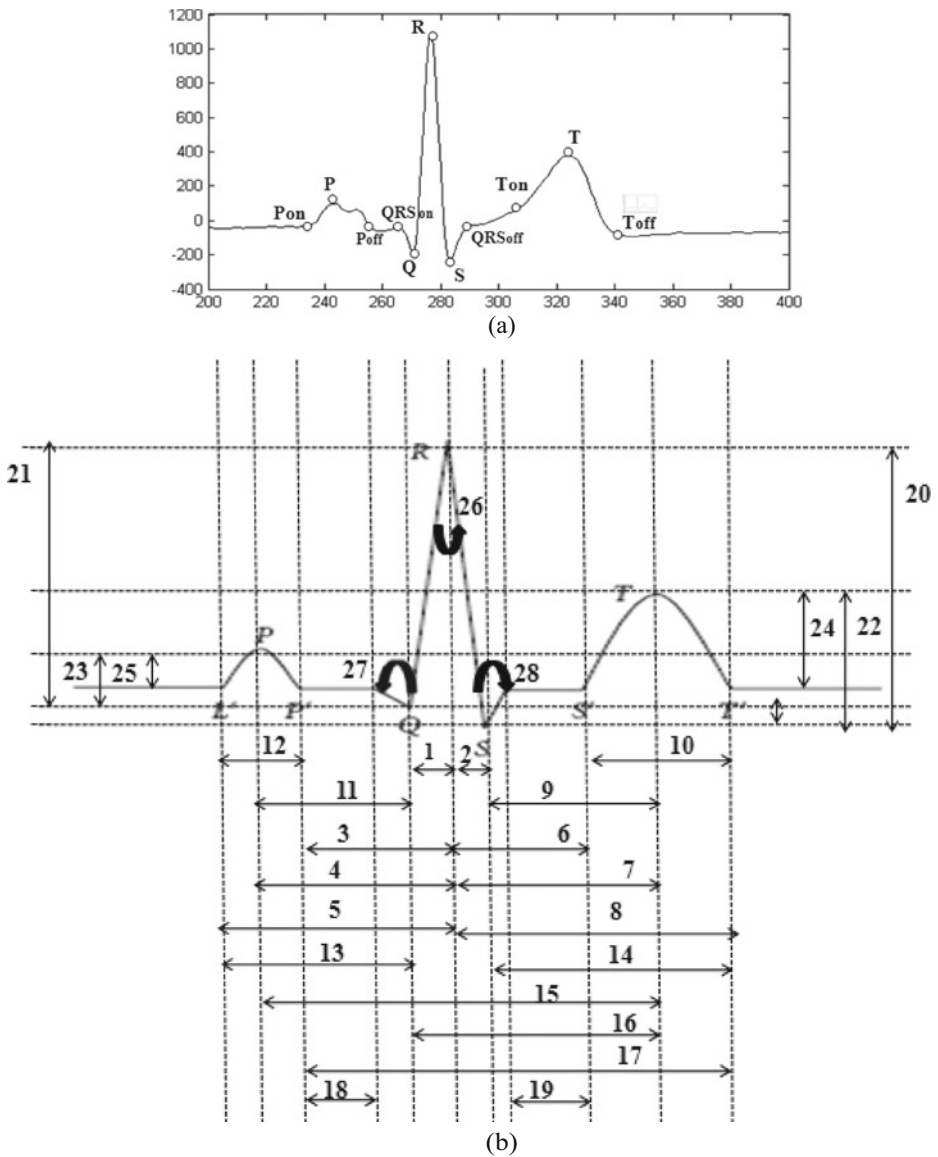


Fig. 2 (a) Representative locations and fiducial labels for the major landmarks in a heartbeat, while (b) presents details of the 28 fiducial features used in this work, depicted with respect to an idealized ECG trace of a single heart beat

3.3 Feature reduction

The goal of feature reduction is to reduce the computational load on the data processing stages, which in turn generally improves the performance of the classifier by removing irrelevant information that may confuse the classifier. There are two basic feature reduction approaches: 1) feature extraction and 2) feature selection.

Patterns are hard to find in data of high dimension. Hence, feature extraction methods transform the data (or the extracted features from the data) from their original space of dimension n to a new subspace of dimension m where $m < n$, while preserving most of the useful information (the target concept is preserved). On the other hand, feature selection methods remove superfluous and redundant features (typically by either feature ranking or subset selection) and keep the remaining features intact (Zhang and Yao 2004; John et al. 1994).

3.3.1 Feature extraction methods

In this paper, two common feature extraction methods were utilized: principle component analysis (PCA) and linear discriminant analysis (LDA) methods. The next subsections provide the needed details about the two methods.

A. Principle Component Analysis (PCA)

PCA is a prevalent statistical method for reducing the dimensionality via feature extraction, which preserves most of the information context of a feature space. Thus, it has found successfully application in pattern recognition and data compression. PCA can identify the orthogonal basis functions of a subspace, which contains most of the data variance. The orthogonal axes (Principle Components) are the first m *eigenvectors* corresponding to the first m largest eigenvalues of the covariance matrix of the data where $m \ll n$ (Smith 2002). The covariance matrix of the data is computed as follows:

$$Cov = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N (X_i - \bar{X})(X_j - \bar{X}) \tag{1}$$

where N is the number of training vectors, X_i is the i th training vector and \bar{X} is the mean of training vectors $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$. Finally, all what is needed to transform any input data vector of n dimension to the new subspace of m dimension is to multiply this vector with the matrix U of m chosen principle components.

$$Y_i = U^T(X_i - \bar{X}) \quad i = 1 \dots N \tag{2}$$

B. Linear Discriminant Analysis (LDA)

LDA is another feature extraction approach that employs supervised learning to find a set of m feature basis vectors V in such a way that the ratio of between-class and within-class scatters of the training dataset is maximized (Balakrishnama and Ganapathiraju 1998). The maximization is equivalent to solve the following eigenvalue problem

$$V = \operatorname{argmax}_V \frac{|V^T S_b V|}{|V^T S_w V|}, \quad V = \{v_1, v_2, \dots, v_m\}, \tag{3}$$

where S_b and S_w are between-class and within-class scatter matrices. The m vectors of V are eigenvectors corresponding to the first m largest eigenvalues. For a dataset

X of N examples belong to C classes, the number of examples belong to each class is n_i and x_{ij} is the example number j for class i , S_b and S_w can be computed as follows

$$S_b = \frac{1}{N} \sum_{i=1}^C n_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T \quad (4)$$

$$S_w = \frac{1}{N} \sum_{i=1}^C \sum_{j=1}^{n_i} n_i (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)^T \quad (5)$$

Where \bar{x}_i is the mean of training examples of class i and \bar{x} is the mean of the whole dataset. Finally, transforming any training vector x_{ij} is done simply by multiplying it with the V matrix as follows

$$y_{ij} = V^T x_{ij} \quad (6)$$

3.3.2 Feature selection methods

The simplest method for feature selection is to generate all the possible subsets of n features, keeping the optimal subset, defined in terms of the resulting classification accuracy. However, this kind of search is only feasible for small n , the value of which is partially determined by the complexity of the classification task. An alternative approach is to perform a random search (i.e. genetic algorithms) and try to find a sub optimal subset. Typically, because this approach is probabilistic, the results are typically not stable or necessarily optimal. Another alternative and commonly used approach utilizes some form of heuristic search (Zhang and Yao 2004; John et al. 1994), where a heuristic function (similarity measure) is utilized and the features are chosen in such a way that maximizes this heuristic function. The heuristic selection can be sequential forward (i.e. agglomerative) where one begins with an empty set then features are added one by one, or one can begin with a set that contains features that are known indispensable (if exists) then continue adding features one by one. Finally, a backward elimination approach (hierarchical) can be deployed, where you begin with the full set of features, then irrelevant features are removed one by one according to the heuristic function. The heuristic search in the feature space can be stepwise where one can add (remove) a feature that was removed (added) before during the search process or one can adopt a greedy approach without feedback (John et al. 1994). In this paper, two heuristic functions or measures for feature selection efficacy were deployed, based on very different measures: 1) information gain based measure and 2) a rough sets based measure. The next subsection provides a brief overview of each of the methods.

A. Information Gain Ratio (IGR) measure

The IGR (Mitchel 1997) is a statistical measure that has been utilized in decision trees learning algorithms to provide a basis for selecting amongst candidate attributes at each step while growing the tree. IGR depends on the concept of entropy, which in this context characterizes the impurity of an arbitrary collection of examples S . If the target attribute can take on c different values, then the entropy (Mitchel 1997) of S relative to this c -wise classification is defined as,

$$\text{Entropy}(S) = \sum_{i=1}^c -p_i \log_2 p_i, \quad (7)$$

where p_i is the proportion of S belonging to class i . Using this formulation (7), one can simply define IGR of an attribute as the expected reduction in entropy caused by partitioning the examples according to this attribute. More precisely, the information gain, $Gain(S, A)$ of an attribute A (Mitchel 1997), relative to a collection of examples S , is defined as

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v), \tag{8}$$

where $Values(A)$ is the set of all possible values for attribute A , and S_v is the subset of S for which attribute A has value v (i.e., $S_v = \{s \in S | A(s) = v\}$). Thus, the features are ranked according to their IGR. The selection algorithm begins with an empty set F of best features and then proceeds to add features from the ranked set of features until the classification accuracy begins to drop or it reaches a specific selected value (Mitchel 1997).

B. Rough Sets based selection

Rough sets (Pawlak 1982) is a mathematical approach which was developed as a method for extracting information from data collections through a set of reducts and decision rules. It has attracted a great deal of interest recently in field of data mining, feature selection and decision rule generation. In this paper, we apply a forward selection approach. In the beginning we have an empty set of features R . we use a heuristic function to measure the significance of the unselected features. Each time the feature of maximum significance is added to generate the next candidate feature subset until the stop criterion is reached. The heuristic function commonly used in rough sets based approaches is the positive region which can be defined as the set of object that be certainly classified to the classes (partitions of universe according to decision attribute D) using attributes of set R and is denoted by the symbol $POS_R D$. Thus, the most significant feature is defined as the one that can cause the maximum increase in the cardinality of the positive region achieved by the selected features until the current step. In this paper, the parameterized averaged support heuristic (PASH) algorithm was utilised (Zhang and Yao 2004). This algorithm gives the opportunity to include predictive instances that may produce predictive rules, which hold true with high probability but are not necessarily 100 %. Moreover, the heuristic function proposed by this algorithm has the advantage of choosing not only the feature ‘ a ’ that provides a larger positive region but also the one that provides maximal average support for the most significant rules *for all* decision classes. Thus the heuristic function is defined as the multiplication of both properties as follows:

$$F(R, a) = |POS_{R+\{a\}}(D)| \times Q(R, a), \tag{9}$$

$$Q(R, a) = \frac{1}{C} \sum_{i=1}^C S(R, a, d_i), \tag{10}$$

$$S(R, a, d_i) = Max\ Size \left(\frac{POS_{R+\{a\}}(D = d_i)}{IND(R + \{a\})} \right), \tag{11}$$

where C is the number of classes, D is the decision attribute and has the domain $\{d_1, d_2, \dots, d_c\}$ and $IND(R + \{a\})$ is the indiscernibility (equivalence) relation, if $(x, x') \in$

IND($R+\{a\}$) then x and x' are indiscernable by set of features $R+\{a\}$ (Zhang and Yao 2004).

3.4 Classification (subject identification)

After the feature sets have been acquired, the reduced set of features of a heartbeat is fed into the classifier in order to perform the classification task (i.e. associate a heartbeat to a particular subject ECG record). Due to its superiority with the full set of features in our previous work (Tantawi et al. 2012), Radial Basis Functions (RBF) neural network was employed here also as a classifier. RBF is distinguished by its rapid training, generality and simplicity. It has been proved that RBF networks, with enough hidden neurons, are also universal approximators. The RBF network is based on the simple idea that an arbitrary function $y(x)$ can be approximated as the linear superposition of a set of localized basis functions $\varphi(x)$ (Haykin 1999). The RBF is composed of three different layers: the input layer in which the number of nodes is equal to the dimension of input vector. In the hidden layer, the input vector is transformed by a radial basis activation function (Gaussian function):

$$\varphi(x, c_j) = \exp\left(\frac{-1}{2\sigma^2} \|x - c_j\|^2\right) \quad (12)$$

Where $\| \cdot \|$ denotes the Euclidean distance between the input data sample vector x and the center c_j of Gaussian function of the j th hidden node; finally the outer layer with a linear activation function, the k th output is computed by equation

$$F_k(x) = \sum_{j=1}^m w_{kj} \varphi(x, c_j) \quad (13)$$

w_{kj} represents a weight synapse associates with the j th hidden unit and the k th output unit with m hidden units (Haykin 1999). The orthogonal least square algorithm (Chen and Chng 1996) is applied to choose the centers, which is a very crucial issue in RBF training due to its significant impact on the network performance. This algorithm was chosen for its efficiency and because there are very few parameters to be set or randomly initialized.

4 Experimental work

In this paper, the results of a set of six experiments are presented. The first five considered different methods for feature reduction, while the last one examined the generalization capability of the resulting systems. To evaluate the performance of the feature reduction methods, the following Physionet databases were utilized:

- 1) PTB database (Oeff et al. 1995): contains ECG records from 51 normal subjects. In this paper these records were divided into 2 sets of subjects: 1) PTB_1, includes 13 subjects with more than one record, some of which were acquired several years apart and; 2) PTB_2, which contains single ECG records from 38 subjects, recorded over 1.5–2 min.
- 2) MIT BIH Normal Sinus Rhythm database (The MIT-BIH Normal Sinus Rhythm Database, <http://www.physionet.org/physiobank/database/nsrdb/>): contains 18

ECG (2 leads) recordings from different subjects. The subjects included in the database did not exhibit significant arrhythmias. The data was acquired at 128 Hz, and recorded over a single period of 24 h continuously. The vast majority of the ECG records contained significant amounts of noise (as judged by baseline drift and other quality measures). Moreover, the records of 4 subjects include artifacts that reduce the valid heart beat information in them. Thus, this database was deployed as a source of impostors only, along with 6 subjects from the MIT BIH long term database (The MIT_BIH Long Term Database, <http://www.physionet.org/physiobank/database/ltdb/>.) which has the same level of signal quality.

- 3) Fantasia database (The Fantasia Database, <http://www.physionet.org/physiobank/database/fantasia/>): it contains records for 40 subjects, one record for each subject, acquired over a 2 h interval, sampled at 250Hz.

Training and testing using different records is recommended to provide the test results more reliability and robustness. However, PTB_1 is the only dataset in Physionet with more than one record for each subject. Hence, it was employed for training purposes, although it encompasses the lowest number of subjects comparing with the other PhysioNet datasets. For comparison, all experiments in this paper followed the same training & testing strategy proposed in Tantawi et al. (2012) and can be summarized in the following steps:

1. The PTB_1 dataset was partitioned in three sets, one set was utilised for training and two sets for testing. The training set contains 13 records, one for each subject. The set labelled ‘Test set 1’ contained 14 records (each of them contains 100 beats) which were recorded on the same day as the training records, but at different sessions, while the set labelled ‘Test set 2’ contains nine records (acquired from six subjects) recorded at various time intervals (several months to years) after the training records.
2. 10 beats (≈ 8 s) were randomly chosen from each of the training records.
3. The full set of 28 features was extracted and normalized for each heartbeat.
4. The features are reduced according to the considered feature reduction method.
5. The feature sets were fed into an RBF classifier for training and testing purposes (classification).
6. The classification task utilised the data as indicated in step1, from which subject identification (SI) and heartbeat recognition (HR) accuracies were measured. SI accuracy is defined as the percentage of subjects correctly identified by the system. While, the HR accuracy is the percentage of heartbeats correctly recognized for each subject. A subject was considered correctly identified if more than half of his\her beats were correctly classified to him\her and a heartbeat is recognized by majority voting of the classifier outputs.
7. Calculating the false acceptance rate (FAR) and the false rejection rate (FRR) is typically acquired by adjusting one or more acceptance thresholds. The acceptance thresholds are varied, and for each value, the FAR and FRR are computed. In this work, two thresholds were employed, termed Θ_1 and Θ_2 . Θ_1 was designed to modulate heart beat (HB) classification while Θ_2 modulates subject identification (SI). Note that this threshold (Θ_2) represents the minimum number of correctly classified beats out of 100 testing beats. The FRR is computed based on Test set 1 only because Test set 2 has only records for 6

subjects. The set of imposters for computing FAR (approximately 100 subject ECG trails) were gathered from three different databases: including 38 subjects of PTB_2, 24 subjects of MIT_BIH and 40 subjects of Fantasia.

4.1 Temporal and amplitude features separately

A question that still requires empirical evidence to answer is the information content of the three classes of fiducial based features: amplitude, magnitude, and angle features (see Fig. 2b) of same importance. To address this important question, experiments were conducted that examined the quantitative importance of classes of features, measured in terms of the resulting classification accuracy when utilizing a particular subset of the feature space. Note that throughout this work; the classification task was implemented using a Radial basis Function (RBF) neural network, trained as per Tantawi et al. (2012). There are a total of 19 temporal features and 9 amplitude features (including 3 angle features). Note that because of the paucity of angle features, they were incorporated into the amplitude features throughout this work. The RBF neural network was adjusted to have a spread (width Gaussian of functions) of 1 and the sum squared error (SSE) of 3 (these values were found empirically). The results for both test sets are presented in Table 1. The results indicate that both feature classes provide roughly equal accuracy across all tests. Though, this result must be interpreted in terms of the disparity between the cardinality of the respective feature classes (19 for temporal features and 9 for both amplitude and angle). Further, amplitude features can separate different classes of training data better than temporal features, as indicated by the data presented in Fig. 3. With respect to the FAR & FRR measures, the thresholds were adjusted (step 7) linearly, and the resulting FAR/FRR were calculated. Even with individual parameter tuning, the results indicate that each class alone (temporal or amplitude/angle) was not sufficient to provide the necessary classification accuracy for deployment as a biometric. For the remaining experiments, the full feature set (28) was utilised, and a variety of feature reduction techniques were applied in order to determine which reduced feature set provided the best classification accuracy.

4.2 Feature reduction methods

In this study, a series of feature extraction and feature selection approaches were utilised (PCA, LDA, IGR, and Rough Sets) in order to establish which features were most relevant for the purposes of subject classification. The efficacy of each approach

Table 1 Summary of the subject identification (SI) and average heart beat classification (HR) results obtained utilising temporal or amplitude features only

	Test set 1		Test set 2	
	SI accuracy	HR accuracy	SI accuracy	HR accuracy
Temporal features	100 %	85.6 ± 9.5 %	66.67 % (4\6)	63.1 ± 15.5 %
Amplitude features	100 %	84 ± 15 %	66.67 % (4\6)	58.5 ± 17 %

For HR, $m \pm \text{std.}$ is provided where m is the mean for all subjects and std. is the standard deviation around mean

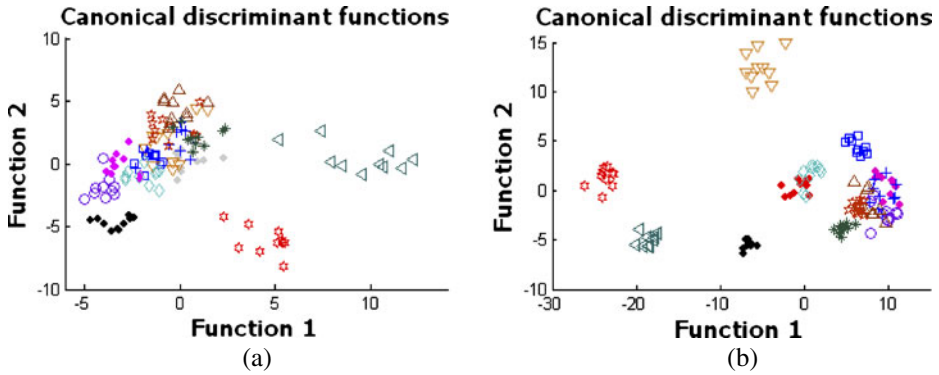


Fig. 3 Presents all-class scatter plots for **(a)** temporal features **(b)** amplitude/angle features. Each class is represented by a different colored geometric shape

was measured in terms of SI and HR, along with measures of FAR and FRR—typical biometric quality indicators.

4.2.1 Principle component analysis (PCA) based method

The PCA algorithm was applied to the full feature set and the first m principle components were chosen. The optimum value of m is one that provides the best classification accuracy. After transforming the feature vector of each heartbeat from its original space to the new m -dimension space, the vectors are fed into the RBF classifier with a spread of 1 and an SSE of 5 for training (values were found empirically). Figure 4a presents the HR accuracy for both test sets achieved using the RBF classifier with different values of m beginning with $m = 6$ the value that achieves 100 % SI accuracy for test set 1. It is clear that the HR accuracy become stable when $m \geq 10$ especially for Test set 1. The thresholds Θ_1 and Θ_2 were readjusted in order to measure FAR and FRR (step 7). Figure 4b presents the dependence of FRR and FAR with Θ_2 when Θ_1 is 0.5. One important result is that when $\Theta_2 = 70$, the FRR is 0 %, yielding an FAR will be 11.7 %. While, in case of

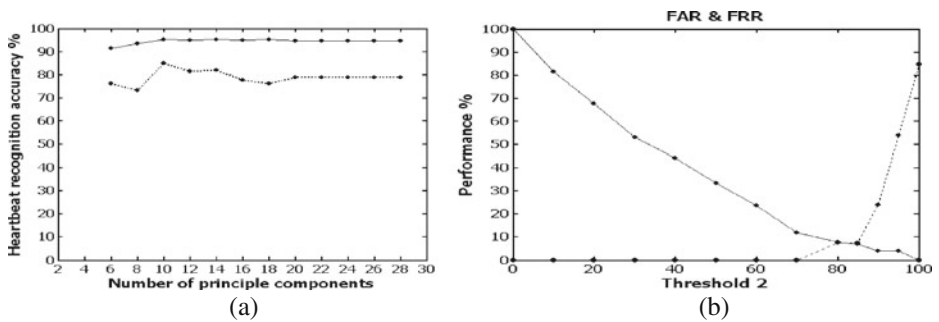


Fig. 4 **(a)** The HR accuracy achieved by Test set 1 (*upper solid line*) and Test set2 (*lower dashed line*) as a function of m . **(b)** Depicts FAR (*solid line*) & FRR (*dashed line*) as a function of changing Θ_2 when Θ_1 was held fixed at 0.5

Table 2 Summary of the best results achieved by PCA before and after threshold adjustments, for HR, $m \pm \text{std.}$ is provided where m is the mean for all subjects and std. is the standard deviation around mean

	Test set 1		Test set 2	
	SI accuracy	HR accuracy	SI accuracy	HR accuracy
$m = 10$	100 %	95 ± 4 %	83.3 % (5\6)	85 ± 9 %
$m = 10$	93.4 % (12\13)	94.5 ± 3.5 %	66.67 % (4\6)	78 ± 5.6 %
$\Theta_1 = 0.5, \Theta_2 = 85$ %				

$\Theta_2 = 85$ % the FAR was 6.8 % and FRR was 7.6 % (which means 1 of the 13 subjects may be rejected when he tries to log in and asked to try again). Table 2 presents the best results achieved by PCA before and after systematic threshold adjustment.

4.2.2 Linear discriminant analysis (LDA) based method

The LDA was applied to the full feature set and the first n components were selected. The optimum value of n is the one that provides the best classification accuracy. After transforming the feature vector of each heartbeat from its original space to the new n -dimension space, the vectors are fed into the RBF classifier with spread 1 and SSE 2 for training. These values were found empirically during the experiment. Figure 5 shows the HB accuracies for both test sets achieved by the RBF classifier with different values of n . It is clear that using LDA poor results for test set 2, with only the records of 3 subjects out of 6 recognized, with an average HB accuracy of 80 %. The thresholds Θ_1 and Θ_2 were readjusted to measure FAR & FRR (step 7). The best results for FAR & FRR were achieved when $n = 10$. Figure 5 presents the variation of FRR and FAR with Θ_2 when Θ_1 is 0.4. When $\Theta_2 = 70$, the FAR was 10.7 % when FRR was 0 %. While, in case of $\Theta_2 = 85$ % the FAR becomes 6.8 % and FRR will be 7.6 % (which means 1 of the 13 subjects may be rejected when s/he tries to log in). Table 3 shows the best results achieved by LDA before and after thresholds readjustment.

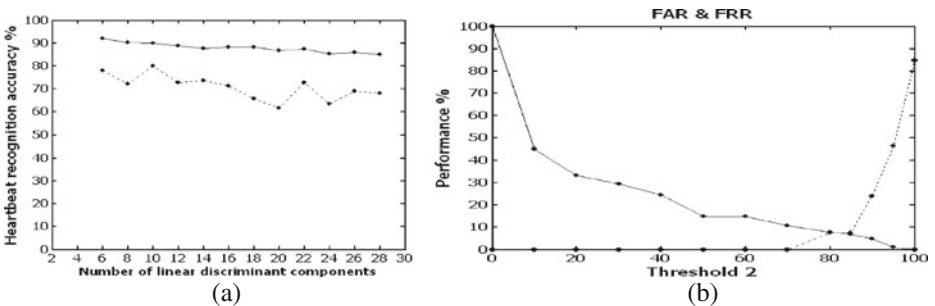


Fig. 5 (a) The HR accuracy achieved by Test set 1 (upper solid line) and Test set2 (lower dashed line) as a function of n . (b) Depicts FAR (solid line) & FRR (dashed line) as a function of changing Θ_2 when Θ_1 was held fixed at 0.4

Table 3 Presents the best results achieved by LDA before and after thresholds readjustment

	Test set 1		Test set 2	
	SI accuracy	HR accuracy	SI accuracy	HR accuracy
n = 6	100 %	92 ± 8 %	50 % (3\6)	77.78 ± 13 %
n = 10	100 %	89.9 ± 10 %	50 % (3\6)	80 ± 14 %
n = 10	93.4 %	94.6 ± 4.5 %	33.34 % (2\6)	76.5 ± 10 %

$\Theta_1 = 0.4, \Theta = 85 \%$

For HR, $m \pm \text{std.}$ is provided where m is the mean for all subjects and std. is the standard deviation around mean

4.2.3 Information gain ratio (IGR) based method

The 28 features were arranged in descending order according to the information gain property. Starting from an empty set, features were added one by one until the best accuracy was achieved by the RBF classifier with a spread of 1 and SSE of 4 (these values were found empirically for each added feature during the experiment). Figure 6a summarises the results of the RBF classifier with respect to varying the number of features (the IGR acceptance threshold). The best results for both test sets were achieved with only 14 features (50 % reduction), in terms of classification accuracy. However, this set of 14 features did not yield the best results in terms of FAR/FRR (step 7). In order to provide acceptable FAR/FRR values (in addition to overall classification accuracy), a set of 23 features (Table 4 and see Table 5 for a listing of the attribute labels read according to Fig. 2) was required. Figure 6b presents the values for FRR and FAR as a function of Θ_2 when Θ_1 was fixed at 0.55 (this value yielded the optimal FAR/FRR results overall). When $\Theta_2 = 85 \%$, FRR was 0 % and was 8.8 %. It is clear that any attempt to more improvement in the FAR will be at the expense of FRR which will significantly increase. Table 6 summarises the best results achieved by IGR before and after thresholds readjustment.

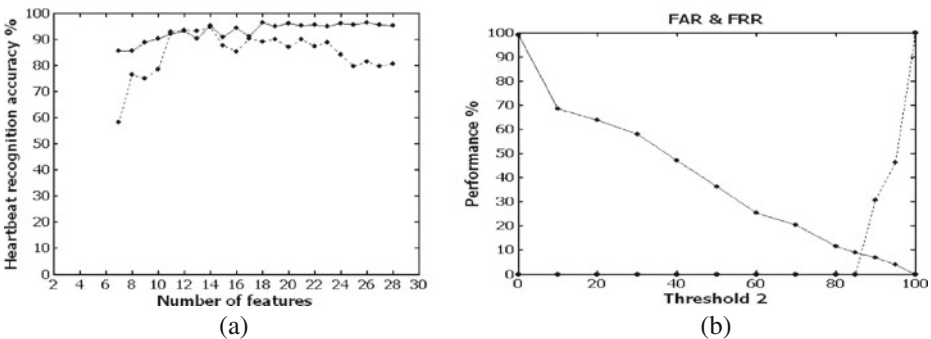


Fig. 6 (a) The HB accuracy achieved by Test set 1 (upper solid line) and Test set2 (lower dashed line) as a function of number of features selected. (b) Depicts FAR (solid line) & FRR (dashed line) as a function of changing Θ_2 when Θ_1 was held fixed at 0.55

Table 4 Ranked list of selected features with respect to classification accuracy and FAR/FRR selected by IGR (first row) and RS (second row) based methods

Feature selection method	Index of selected features
IG (23 features)	28 23 24 25 3 5 13 4 17 18 26 12 27 15 21 11 2 14 20 6 1 8 16
RS (21 features)	28 24 23 17 20 5 25 4 3 13 12 14 18 16 15 8 6 19 11 2 26

The bold elements highlight the features unique to the approach

Table 5 List of indexes and names of selected features by IGR and RS based methods, (m, n) means m features are selected out of n. indexes are ranged in order 1–19, 20–25 and 26–28 for temporal, amplitude and angle features respectively (shown in Fig. 2b)

Selected features	IGR (23 features)	RS (21 features)
Temporal	(15/19) 1) RQ 6) RS' 14) ST' 2) RS 8) RT' 15) PT 3) RP' 11) PQ 16) QT 4) RP 12) P'L' 17) P'T' 5) RL' 13) QL' 18) P'O'	(15/19) 2) RS 8) RT' 15) PT 3) RP' 11) PQ 16) QT 4) RP 12) P'L' 17) P'T' 5) RL' 13) QL' 18) P'O' 6) RS' 14) ST' 19) F'S'
Amplitude	(5/6) 20) RS 23) PQ 25) PL' 21) RQ 24) TT'	(4/6) 20) RS 24) TT' 23) PQ 25) PL'
Angle	(3/3) 26) < R 28) < S 27) < Q	(2/3) 26) < R 28) < S

Missing numbers represent features unselected by each algorithm

Table 6 Summary of the best results (classification) achieved by the IGR based method before and after threshold readjustment

	Test set 1		Test set 2	
	SI accuracy	HR accuracy	SI accuracy	HR accuracy
14 features	100 %	95.2 ± 3.3 %	83.3 % (5\6)	94.7 ± 3 %
23 features	100 %	95 ± 3 %	83.3 % (5\6)	88.8 ± 11 %
23 features	100 %	91.5 ± 6 %	66.67 % (4\6)	80.8 ± 8 %

⊖1 = 0.55, ⊖2 = 85 %

For HR, m ± std. is provided where m is the mean for all subjects and std. is the standard deviation around mean

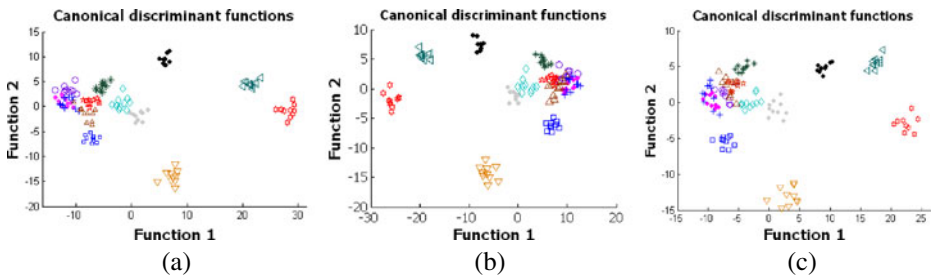


Fig. 7 Presents the All-class scatter plot (a) full set of features (b) IGR feature set (c) RS feature set. Each class is represented by a different colored geometric shape

4.2.4 Rough sets (RS) based method

The PASH algorithm was applied to the super set of features. The features are added one by one according to the criteria discussed in Section 3.3.2.B. The algorithm stopped after a set of 21 features (Table 4 and see Table 5 for a listing of the attribute labels read according to Fig. 2) were selected (yielding a 25 % reduction in the feature space). The all-class scatter plot in Fig. 7 provides evidence that the RS (utilizing the PASH algorithm) provides better separation for classes than the IGR based approach. The 21 features from each heartbeat of the training set were fed into a RBF classifier with spread 1 and SSE 3 (these values were found empirically during the experiment) from the training task. These results are summarised in Table 7. The thresholds Θ_1 and Θ_2 are readjusted to measure FAR & FRR (step 7). Figure 8 shows the variation of FRR and FAR with Θ_2 when Θ_1 is 0.57. When $\Theta_2 = 70$, the FRR was 0 %, but the FAR was 15 %. However, in case of $\Theta_2 = 85$ % the FAR becomes only 4.9 % and FRR will be 7.6 % (which means 1 of the 13 subjects may be rejected when he tries to log in and asked to try again).

4.3 Generalization to other datasets

Generalization is a very crucial issue, since it provides evidence on the ability of the system to preserve its performance when it is trained and tested using other datasets without any change in the structure of the algorithms, network structure, or parameters. Two datasets were deployed in this part of the study: the PTB_2 and Fantasia, both of which are somewhat noisier than the rest. The parameters and thresholds (Θ_1 and Θ_2) were fixed to their optimum values according to the

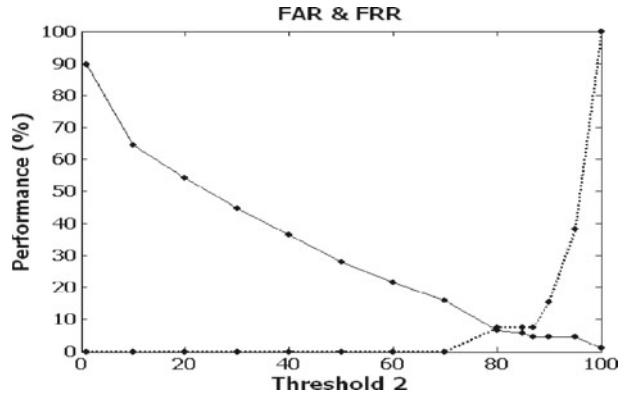
Table 7 Summary of the best results achieved by RS based method before and after threshold readjustment

	Test set 1		Test set 2	
	SI accuracy	HR accuracy	SI accuracy	HR accuracy
21 features	100 % (13\13)	96 ± 2 %	83.3 % (5\6)	93.2 ± 4 %
21 features	93.4 % (12\13)	95 ± 4 %	83.3 % (5\6)	84 ± 5.6 %

$\Theta_1 = 0.57, \Theta_2 = 85$ %

For HR, $m \pm \text{std.}$ is provided where m is the mean for all subjects and std. is the standard deviation around mean

Fig. 8 Depicts FAR (solid line) & FRR (dashed line) as a function of changing Θ_2 when Θ_1 was held fixed at 0.57



considered feature reduction method (values achieved using PTB_1 dataset) and then the system was trained using each of the considered datasets for generalization (PTB_2 or Fantasia). Since there is only one record for each subject in the two datasets, the system was trained and tested using beats extracted from the same record (10 beats were selected randomly, and 100 beats from the remaining set were selected for testing). The average HR accuracy and FRR and FAR were computed and are summarised in Table 8. The set of imposter for the FAR test in this experiment encompasses subjects of PTB_1, MIT_BIH and Fantasia (PTB_2), if the PTB_2 (Fantasia) was using for training.

5 Analysis & discussion

This work examined the information content of a set of 28 fiducial features. The fiducial features that were extracted fall into three general categories: temporal, amplitude, and angle (see Fig. 2 for details). The vast majority of the 28 features were temporal (19), with 6 amplitude and 3 angle features. To the author's knowledge, no systematic study has been published reporting the relative contribution to classification accuracy of these three features classes, in terms of specific contributions at the class level and individual features within class levels, in terms of classification accuracy, stability (Test set 1 and Test set 2) and FAR/FRR. In this study, the classification accuracy of amplitude and angle features was investigated together,

Table 8 Presents the generalization results for PTB_2 and Fantasia datasets with each of the considered feature reduction methods, for HR, m \pm std. is provided where m is the mean for all subjects and std. is the standard deviation around mean

	PTB_2			Fantasia		
	HR %	FRR %	FAR %	HR %	FRR %	FAR %
PCA	96.1 \pm 3 %	2.5 %	13.7 %	86.75 \pm 6.5 %	20 %	8 %
LDA	98 \pm 1.8 %	2.5 %	10.4 %	89.2 \pm 10 %	5 %	12 %
IG	97 \pm 3 %	5 %	10.4 %	90 \pm 8 %	10 %	9.3 %
RS	97 \pm 2 %	2.5 %	6.5 %	89.9 \pm 6 %	7.5 %	4 %

because of the small number of angle features (3 of the 28 features). Therefore, our analysis compared the information content of temporal and non-temporal features with respect to their classification accuracy.

As an initial investigative step, the classification accuracy of temporal and amplitude/angle features was analysed. The classification accuracy at the subject level was 100 % for both feature classes, and a slight reduction in HR (−1.6 %) accuracy for the amplitude feature vector. This is a small difference and should be interpreted in light of the significant difference in the cardinality of each feature set: 19 for temporal and 9 for amplitude+angle. In addition, the amplitude and angle features were examined from a scatter plot analysis, which depicts within and between class separations. The amplitude+angle features provided more robust separation between classes than the temporal features (see Fig. 3). Further, three of the amplitude and angle features occupy the first three most significant places in both IGR and RS ranked feature lists (Table 4). These results provide evidence that the amplitude and angle features be more reliable indicators of subject individuality than temporal features. This interpretation must be qualified with the possibility of error in the temporal feature extraction process. The amplitude and angle features in most cases rely on detecting the peaks (P, R and T) and valleys (Q and S) of the three complex waves of an ECG heartbeat which are usually detected with high accuracy due to their sharpness. On the other hand, temporal features rely strongly on the beginning and ending points of the three complex waves. Such points are very challenging and they are usually subject to some small margin of error which may not be consistent across all features, and hence may not be controlled for by a suitable translation factor. This should be controlled for by performing a test on an artificial ECG time series with known temporal features (part of future work).

The level of redundancy with the feature set was examined using a set of four feature alteration strategies. As a control, the reduced feature vectors were processed and examined in the same manner as full feature set was utilized. The same basic measures of SI, HR, and FAR/FRR were acquired after performing these dimensionality altering steps sequentially. These include: two that implement dimension reduction (LDA and PCA) and two that implement feature selection (IGR and RS). It should be noted that PCA and LDA are typically employed to reduce dimension of non fiducial features, these familiar approaches were utilised for the first time in the context of fiducial based ECG feature classification. Moreover, this work also introduced the PASH algorithm into the rough sets rule generation strategy—the first such application within the biosignal domain generally. Our comparison between the dimension reduction and feature selection approaches was based on the accuracies achieved, not on the level of dimensionality reduction due to the different mechanism of each.

According to the results summarized in Fig. 9, the feature selection methods performed better than the feature extraction methods (LDA and PCA) for all metrics recorded in this study. In this paper, we applied rough sets (RS) for the first time in the literature for reducing fiducial features for biometric use. Our RS approach (with the PASH algorithm for rule selection) generated the best overall results in terms of SI, HR, FAR/FRR, and generalizability. The generalizability was tested using datasets of round 40 subjects and this gives evidence on the scalability of our RS based system whose parameters and thresholds were adjusted according to PTB_1 dataset (13 subjects only). Moreover, RS approach along with

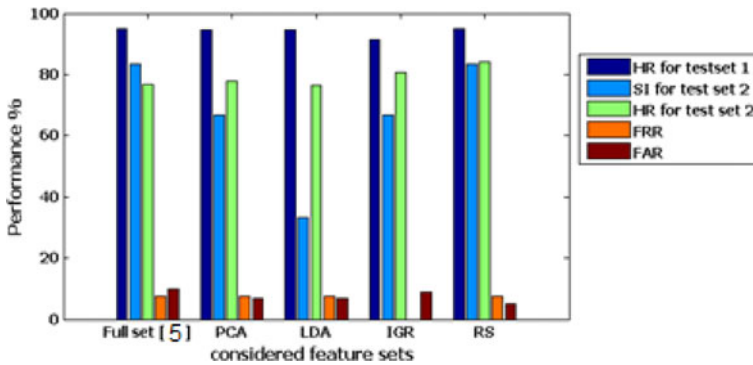


Fig. 9 A graphical summary of heart beat (HR) recognition, subject identification (SI), FAR, and FRR results obtained from each of the four feature reduction methods

RBF classifier provides comparable results with other reduction methods like Wilk's lambda or non fiducial based methods for feature extraction like wavelets, Discrete cosine transform (DCT) and polynomial coefficients along with Euclidean distance criteria for classification or more sophisticated classifiers like neural networks (MLP, SVM and RBF). The experimental results indicate that if the SI accuracy is the only relevant metric, as in most of the existing systems, any of this dimensionality altering approaches yields 100 % accuracy using only 6–8 features (25 % of full set). However, SI accuracy alone may not be a sufficient metric when extrapolating to large population of subjects to be identified—a more robust system may be required. In this study, to achieve high HR rates, stability, and generalizability, required roughly half of the feature set. Finally, 75 % of the full set was needed to achieve the best (and acceptable levels) FAR & FRR measures. This is can be explained by the fact that classifying a heartbeat is usually done by majority voting which means the heartbeat is labeled with the class corresponding to the classifier output of maximum value regardless of the magnitude of the output. However, in order to obtain acceptable results for the FAR & FRR, the maximum value should exceed some threshold (Θ_1) and the number of classified beats to a subject should exceed another threshold (Θ_2) in order to be considered identified. Thus, more features are needed to pump more power to the winner classifier output and increase HR accuracy of each person.

All the experiments presented in this work were evaluated using the physionet datasets. These datasets includes many subjects (men and women of different ages) with a variable length, single ECG record. However, except for the subjects of PTB_1, each subject has only one record which makes it difficult to test the stability. Further, physionet datasets are noisy, especially the fantasia dataset, which present significant and possibly unrealistic challenges to the development of an ECG based biometric system from this data. Lastly, the detection of fiducial points has a significant impact on the accuracy of the extracted features. A subject may be rejected just because his\her ECG record is noisy and the extracted features are not accurate enough. For example, in the stability test, there is one subject that has one record belong to Test set2 that was not recognized by the full set or any of the reduced feature sets presented in this work, resulting in maximum SI accuracy for

Test set 2 of 83.3 %. This record was recognized only when the amplitude features alone were used. Thus, one explanation may be given that may that the extracted temporal features were not accurate enough and confused the classifier. This issue may be addressed by utilizing non-fiducial features in conjunction with fiducial based approaches. One could envision switching from one to the other, depending on the noise level within the data, if one feature set appears to be more suitable for handling noisy data.

6 Conclusion

In summary, the major conclusions of this work indicate that dimensionality reduction enhances classification performance, and that rough sets was superior in most respects to PCA, LDA, and IGR based approaches. The results of this work indicate that at least 25 % of the typical fiducial based feature space can be eliminated, yielding a feature set with approximately 21 features. The amplitude features tend to be more informative and appear at the upper end when features are ranked in descending order of information content. It should be noted in this regard that amplitude (or amplitude + angle features) alone, as with temporal features, do not provide the level of SI, HR, and FAR/FRR required by a biometric system. Their combination is required in order to maximize overall classification metrics. The PASH algorithm was introduced in this work in connection with ECG based feature reduction. From our previous results, the PASH algorithm enhanced the overall performance of the rough sets approach across all measured metrics, relative to using a high pass feature filter (see Tantawi et al. 2012 for more details). The results from this study are of course only applicable to fiducial based features, and we are engaged currently with expanding the feature space to include a variety of non-fiducial features, providing a comprehensive set of features extracted using a linear based approach. Further, in order to overcome the drawbacks previously mentioned on physionet datasets, we are collecting subject data of our own, which includes test/re-test acquisition over six time points (t_0 , 1 month, 2 months, 6 months, and 12 months) for large number of subjects as much as possible. It is hoped that having such a dataset will provide the means to more thoroughly evaluate the temporal stability and scalability of the classification scheme developed in this work.

References

- Agrafioti, F., & Hatzinakos, D. (2008). ECG based recognition using second order statistics. In *Sixth annual conference on communication networks and services research (CNSR)* (pp. 82–85). Halifax, Canada.
- Agrafioti, F., Gao, J., Hatzinakos, D. (2011). Heart Biometrics: Theory, Methods and Applications, In Biometrics: Book 3, J. Yang, Eds., Intech.
- Balakrishnama, S., & Ganapathiraju, A. (1998). *Linear discriminant analysis-a brief tutorial*. Institute for Signal and Information Processing.
- Biel, L., Petersson, O., Philipson, L. (2001). ECG analysis: a new approach in human identification. *IEEE Transactions on Instrumentation and Measurement*, 50(1), 808–812.
- Chan, A., Hamdy, M.H., Badre, A., Badee, V. (2005). Wavelet distance measure for person identification using electrocardiograms. *IEEE Transactions on Instrumentation and Measurement*, 37(2), 248–253.

- Chan, A., Hamdy, M., Badre, A., Badee, V. (2008). Person identification using electrocardiograms. *IEEE Transactions on Instrumentation and Measurement*, 57(2), 248–253.
- Chen, S., & Chng, E. (1996). Regularized orthogonal least squares algorithm for constructing radial basis function networks. *International Journal of Control*, 64(5), 829–837.
- Coutinho, D., Fred, A., Figueiredo, M. (2010). One-lead ECG-based personal identification using Ziv-Merhav cross parsing. *20th int. conf. on pattern recognition* (pp. 3858–3861).
- Fatemian, S., & Hatzinakos, D. (2009). A new ECG feature extractor for biometric recognition. *Proceedings of the 16th international conference on digital signal processing* (pp. 323–328). IEEE Press Piscataway, NJ, USA.
- Forsen, G., Nelson, M., Staron, R. (1977). Personal attributes authentication techniques. In A.F.B. Griffin (Ed.), *Rome Air Development Center report RADC-TR-77-1033*. RADC, New York.
- Gahi, Y., Lamrani, A., Zoglat, A., Guennoun, M., Kapralos, B., El-Khatib, K. (2008). *Biometric identification system based on electrocardiogram data*, *New Technologies, Mobility and Security NTMS '08* (pp. 1–5). Tangier, Morocco.
- Ghofrani, N., & Bostani, R. (2010). Reliable features for an ECG-based biometric system. *17th Iranian conference of biomedical engineering* (pp. 1–5).
- Guennoun, M., Abbad, N., Talom, J., Rahman, M., & El-Khatib, K. (2009). Continuous authentication by electrocardiogram data. *Science and Technology for Humanity (TIC-STH), 2009 IEEE Toronto international conference* (pp. 40–42).
- Haykin, S. (1999). *Neural networks: A comprehensive Foundation* (2nd ed.). Prentice Hall.
- Irvine, J.M., & Israel, S.A. (2009). *A sequential procedure for individual identity verification using ECG*. Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing, Vol. 2009, pp. 1-13.
- Israel, S.A., Irvine, J.M., Cheng, A., Wiederhold, M.D., Wiederhold, K. (2005). ECG to identify individuals. *Pattern Recognition*, 38(1), 133–142.
- John, G., Kohavi, R., Pfleger, K. (1994). Irrelevant features and the subset selection problem. *Proceedings of 11th international conference on machine learning* (pp. 121–129).
- Khalil, I., & Sufi, F. (2008). Legendre polynomials based biometric authentication using QRS complex of ECG. *4th int. conference on intelligent sensors, sensor networks and information processing, ISSNIP*. Melbourne.
- Li, M., & Narayanan, S. (2010). Robust ECG biometrics by fusing temporal and cepstral information. *20th international conference on pattern recognition* (pp. 1326–1329).
- Mai, V., Khalil, I., & Meli, C. (2011). ECG biometric using multilayer perceptron and radial basis function neural networks. *33rd annual international conference of the IEEE EMBS* (pp. 2745–2748). Boston, Massachusetts USA.
- Mitchel, T. (1997). *Machine learning* (2nd ed.). McGraw-Hill, New York, US.
- Odinaka, I., Lai, P.-H., Kaplan, A., O'Sullivan, J., Sirevaag, E., Kristjansson, S., et al. (2010). Ecg biometrics: A robust short-time frequency analysis. *IEEE international workshop on information forensics and security* (pp. 1–6).
- Oeff, M., Koch, H., Boussejot, R., Kreiseler, D. (1995). The PTB Diagnostic ECG Database, National Metrology Institute of Germany. <http://www.physionet.org/physiobank/database/ptbdb/>.
- Pawlak, Z. (1982). Rough sets. *International Journal of Computer and Information Sciences*, 11, 341–356.
- Plataniotis, K., Hatzinakos, D., Lee, J. (2006). ECG biometric recognition without fiducial detection. In *Proceedings of biometrics symposiums (BSYM '06)* (pp. 1–6). Baltimore, MD, USA.
- Revett, K. (2008). *Behavioral biometrics*. A remote access approach. Wiley, ISBN: 978-0-470-51883-0.
- Safie, S., Soraghan, J., Petropoulakis, L. (2011a). Electrocardiogram (ECG) Biometric Authentication Using Pulse Active Ratio (PAR). In *Peer-review for IEEE transaction of information forensics and security* (Vol. 6, No. 4, pp. 1315–1322).
- Safie, S., Soraghan, J., Petropoulakis, L. (2011b). ECG based biometric for doubly secure authentication. *19th European signal processing conference (EUSIPCO)* (pp. 2274–2278). Barcelona, Spain.
- Shen, T.W. (2005). *Biometric identity verification based on electrocardiogram (ECG)*. Ph.D. Thesis, University of Wisconsin, Madison.
- Shen, T., Tompkins, W., Hu, Y. (2002). One-lead ECG for identity verification. *Proc. 2nd Joint EMBS/BMES Conf.* (pp. 62–63).
- Singh, Y.N., & Gupta, P. (2006). *ECG to individual identification*. Prabhu Goel Security Center of IIT, Kanpur.

- Singh, Y.N., & Gupta, P. (2009). Biometrics method for human identification using electrocardiogram. In M. Tistarelli, & M.S. Nixon (Eds.), *ICB 2009, LNCS* (Vol. 5558, pp. 1270–1279).
- Singla, S., & Sharma, A. (2010). ECG based biometrics verification system using LabVIEW. *Songklanakarin International Journal of Science and Technology*, 32(1), 241–246.
- Smith, L. (2002). *A tutorial on principal components analysis*. http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf.
- Sufi, F., Khalil, I., & Habib, I. (2008). *Polynomial distance measurement for ECG based biometric authentication*. Security and Communication Networks, Wiley Interscience. doi:10.1002/sec.76.
- Sufi, F., Khalil, I., Hu, J. (2010). *ECG based authentication*. Handbook of Information and Communication Security, Springer.
- Tantawi, M., Revett, K., Tolba, M.F., Salem, A. (2012). On the applicability of the physionet electrocardiogram (ECG) repository as a source of test cases for ECG based biometrics. *International Journal of Cognitive Biometrics. Inderscience*, 1(1), 66–97.
- Tawfik, M., Selim, H., Kamal, T. (2010). Human identification using time normalized QT signal and the QRS complex of the ECG. *7th international symposium on communication systems networks and digital signal processing* (pp. 755–759).
- Ting, C.M., & Salleh, S. (2010). ECG based personal identification using extended kalman filter. *10th international conference on information sciences signal processing and their applications* (pp. 774–777).
- Venkatesh, N., & Jayaraman, S. (2010). Human electrocardiogram for biometrics using DTW and FLDA. *20th international conference on pattern recognition (ICPR)* (pp. 3838–3841).
- Wan, Y., & Yao, J. (2008). A neural network to identify human subjects with electrocardiogram signals. *Proceedings of the world congress on engineering and computer science 2008, WCECS 2008*. San Francisco, USA.
- Wang, Y., Agrafioti, F., Hatzinakos, D., Plataniotis, K. (2008). Analysis of human electrocardiogram for biometric recognition. *EURASIP Journal on Advances in Signal Processing*, 1, 1–6.
- Wao, J., & Wan, Y. (2010). Improving computing efficiency of a wavelet method using ECG as a biometric modality. *International Journal of Computer and Network Security*, 2(1), 15–20.
- Ye, C., Coimbra, M., Kumar, B. (2010). Investigation of human identification using two-lead electrocardiogram (ECG) signals. *4th int. conf. on biometrics: Theory applications and systems* (pp. 1–8).
- Zhang, M., & Yao, J. (2004). A rough sets based approach to feature selection. *Proc. IEEE annual meeting of fuzzy information NAFIP* (pp. 434–439).