Characterisation of electrocardiogram signals based on blind source separation

M. I. Owis   A.-B. M. Youssef   Y. M. Kadah

Biomedical Engineering Department, Cairo University, Giza, Egypt

Abstract—*Blind source separation assumes that the acquired signal is composed of a weighted sum of a number of basic components corresponding to a number of limited sources. This work poses the problem of ECG signal diagnosis in the form of a blind source separation problem. In particular, a large number of ECG signals undergo two of the most commonly used blind source separation techniques, namely, principal component analysis (PCA) and independent component analysis (ICA), so that the basic components underlying this complex signal can be identified. Given that such techniques are sensitive to signal shift, a simple transformation is used that computes the magnitude of the Fourier transformation of ECG signals. This allows the phase components corresponding to such shifts to be removed. Using the magnitude of the projection of a given ECG signal onto these basic components as features, it was shown that accurate arrhythmia detection and classification were possible. The proposed strategies were applied to a large number of independent 3 s intervals of ECG signals consisting of 320 training samples and 160 test samples from the MIT-BIH database. The samples equally represent five different ECG signal types, including normal, ventricular couplet, ventricular tachycardia, ventricular bigeminy and ventricular fibrillation. The intervals analysed were windowed using either a rectangular or a Hamming window. The methods demonstrated a detection rate of sensitivity 98% at specificity of 100% using nearest neighbour classification of features from ICA and a rectangular window. Lower classification rates were obtained using the same classifier with features from either PCA or ICA and a rectangular window. The results demonstrate the potential of the new method for clinical use.*

Keywords—*Principal component analysis, Independent component analysis, Arrhythmia detection, ECG, Statistical classifiers*

1 Introduction

THE EARLY detection of ECG arrhythmia is important for timely management of the patient. It relies on the ability to detect variations in ECG morphology within a small time aperture to detect the presence of any such abnormalities and to identify their type. The current techniques of arrhythmia detection rely on direct (through correlation measurement) or indirect (through quantitative features) comparison between the current ECG signal and samples from a database containing the arrhythmia types of interest. Encouraging results have been obtained using autocorrelation function (GUILLÉN et al., 1989), frequency-domain features (MINAMI et al., 1999), time-frequency analysis (AFONSO and TOMPKINS, 1995) and wavelet transform (KHADRA et al., 1997).

The blind source separation problem has received an increasing amount of interest, in the past ten years, in the area of signal analysis. This problem assumes the acquired signal to be composed of a weighted sum of a number of basic

components corresponding to a number of limited sources. The solution to this problem consists of a set of basic components that correspond to an optimised basis set for the particular problem at hand. The advantages of using such components, rather than any other arbitrary choice for the basis functions, include the ability to reduce the number of features significantly, in addition to separating the components that correspond to noise. This makes the use of such techniques desirable for biomedical signals (OUDA et al., 2001).

Blind source separation techniques have been applied in several aspects of ECG signal processing. Examples of such applications include separating fetal and maternal ECG signals (ZARZOSO et al., 1997; DE LATHAUWER et al., 2000), analysis of the ST segment for ischaemia detection (STAMKOPOULOS et al., 1998) and identification of humans using ECG (BIEL et al., 2001). They have been also used in related areas to diagnose peripheral vascular disease from multiple blood flow measurements (PANERAI et al., 1988). In all these techniques, the measurements were either acquired simultaneously or gated to a certain reference point in the signal (e.g. the R point). Although the blind source separation-based techniques were shown to be successful in the above applications, their use in ECG signal classification has not been addressed in the literature.

In this paper, we pose the problem of ECG signal classification as a blind source separation problem. A large database of

ECG signal samples was utilised to compute the basic components of the ECG signals using principal component analysis (PCA) and independent component analysis (ICA). The signals were preprocessed to obtain the magnitude of their Fourier transformation, to reduce the number of components resulting from different shifts of the same signal. The ECG signal window at hand was projected onto such components, and the projection magnitudes were considered as signal features. Feature vectors from all signals in the training database were collected and used to define the feature space of the problem. Subsequent features from the test set were classified to the closest type within that feature space using statistical classification techniques. The implementation details and the results are presented and discussed to assess the performance of the new technique and its practicality for clinical use.

2 Methods

2.1 Shift invariance transformation

The detection of ECG arrhythmia type relies on observing changes in the ECG signal characteristics as computed from a short window of the signal. This window is generally taken as a moving window that covers a number of seconds of signal, starting from the current sample and moving back in time. This results in unknown phase shift to the ECG signal within the sample window. Moreover, even if the signal window is synchronised to the onset of an R point, heart rate variability prevents the direct comparison of windows obtained from different patients or at different times for the same patient.

Given the sensitivity of PCA/ICA-based techniques to the presence of such practical conditions, their direct use in the analysis of ECG signals has been limited. To overcome this limitation, we propose to apply a simple transformation whereby the magnitude of the Fourier transformation of the signal is used instead of the time-domain signal. As the relative delay between samples is manifested as linear phase in the frequency domain, the proposed transformation yields the same result for all circular shifts of a given signal. Given the periodic nature of the ECG signal, windows with different starting points generally approximate circular shifts of the same signal. As a result, such windows provide similar outputs after this transformation. This alleviates the need for a reference point and allows flexible choice of sample windows.

In our implementation, two types of window were used to select 3 s intervals from the ECG signals, namely, rectangular and Hamming. Although the use of tapered windows such as the Hamming window is the standard in Fourier transformation procedures because of their better side lobe suppression characteristics, the use of a rectangular window in this application was justified for two reasons. First, the Fourier shift theorem holds only approximately when a non-rectangular window is applied (BRACEWELL, 1984). Secondly, the application of the discrete Fourier transformation using a rectangular window amounts to an orthogonal transformation that preserves 2-norm distance between signals in both the time and frequency domains (GOLUB and VAN LOAN, 1993). This also keeps noise within individual frequency bins uncorrelated (PAPOULIS, 1991). Therefore the comparison between the results of the two types of windowing function is of interest in this application.

The length of the window functions used was taken as the number of samples within the selected 3 s interval (i.e. 750 points for ventricular fibrillation samples and 1080 points in all other ECG signal types). As ECG signals are real, the magnitude of the Fourier transformation of the sample windows of the training (or design) database is symmetric. Therefore half of the points in the Fourier transformed data were considered redundant and subsequently excluded from further analysis.

2.2 Principal component analysis (PCA)

Principal component analysis is analogous to Fourier analysis in that the data are described in terms of the coefficients of a predetermined orthogonal set. Rather than using complex exponentials, the orthogonal set in PCA is determined adaptively based on the analysed data set. In particular, PCA derives the directions of a set of orthogonal vectors that point in the direction of the highest variance of the data. The principal components are calculated as the eigenvectors of the covariance matrix of the data (GERBRANDS, 1981). The eigenvalues denote the variances corresponding to these eigenvectors. Hence, PCA is an efficient technique for dimensionality reduction in multivariate statistical analysis.

Given a data matrix X of size $m \times n$, composed of m n-point sample windows, let a centred matrix Z be computed as $Z = (X - E\{X\})$, where $E\{X\}$ is the matrix of mean vectors. Then, PCA is defined as (GERBRANDS, 1981)

$$Y = B^T Z \quad \text{and} \quad K_x = B \Lambda B^T \tag{1}$$

where Λ is a diagonal matrix with eigenvalues of the covariance matrix K_x on the diagonal (with eigenvalues ranked such that $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$), and the columns of B are the corresponding eigenvectors. The output of the PCA transform is uncorrelated vectors. The covariance matrix of the output Y is $K_y = E\{YY^T\} = \Lambda$. Owing to the orthogonality of the matrix B, (1) can be rewritten as $Z = BY$, where the matrices are $m \times n$, $m \times m$ and $m \times n$, respectively. In general, principal component analysis can be used for dimensionality reduction by truncation of the signal components corresponding to the smallest eigenvalues. This can be described as $Z' = B'Y'$, where the matrices are $m \times n$, $m \times q$ and $q \times n$, respectively, with $q < n$. In this case, the selection of the number of eigenvalues allows the inclusion of as much of the variability of the original data as needed. The efficiency of this approximation is estimated by the ratio of the chosen variance to the total system variance, as follows:

$$E = \frac{\sum_{i=1}^{q} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \tag{2}$$

For feature extraction, each centred sample is represented by its projections on the q principal components, computed as the inner product between the centred sample and each of the computed eigenvectors.

2.3 Independent component analysis (ICA)

Independent component analysis (ICA) is a more general form of PCA, whereby higher-order statistics are used in addition to second-order moments, which PCA relies on, to determine the basis vectors (HYVÄRINEN et al., 2001). Let v_i, $i = 1, \ldots, m$, be the measured signals, and let s_j, $j = 1, \ldots, r$, be the signals from independent components (ICs) with zero mean and unit variance. The basic problem in ICA is to estimate the mixing matrix A and the matrix of realisations of the independent components S, such that the matrix of measured signals $V = A \cdot S$. The major constraint of this problem is for r to be less than m. In most cases, r is assumed known, and often $r = m$.

The basic algorithms for computing the independent components rely on measuring the non-Gaussianity of the different vectors within the whitened subspace of signals of interest. The most common method for this purpose is the use of the fourth central moment or kurtosis (HYVÄRINEN and OJA, 1997). The value of kurtosis is zero for Gaussian random vectors and it assumes non-zero values for other distributions. Therefore the iterative maximisation of the kurtosis enables the non-Gaussian components (or independent components) corresponding to the

true underlying sources to be estimated. In this work, a fast fixed-point algorithm was used to perform this task (HYVÄRINEN *et al.*, 2001; HYVÄRINEN and OJA, 1997).

The matrix of data vectors V is first whitened using PCA. The whitened data matrix X is defined by $X = M \cdot V = M \cdot A \cdot S = B \cdot S$, where M is the whitening matrix. The problem of finding an arbitrary full-rank matrix A is reduced to the simpler problem of finding an orthogonal matrix. Subsequently, this matrix can be used to compute the independent components as $S = B^T \cdot X$. In other words, we are looking for an orthogonal matrix W^T such that the matrix $W^T X$ is composed of good estimates of the independent components. To estimate all independent components, an orthogonalising projection is added to the iteration to remove the previously estimated components. This is implemented using a Gram–Schmidt procedure, whereby the previously estimated components compose an orthonormal basis of a subspace of the same dimension as the number of components (GOLUB and VAN LOAN, 1984). This basis set is used to remove all components within this subspace from all subsequent independent components. This step is necessary to satisfy the constraint of uncorrelatedness of independent components, given the orthogonality of independent components in the whitened space (HYVÄRINEN *et al.*, 2001).

For feature-extraction purposes, each centred sample is represented by its projections on q independent components in a similar fashion to the PCA. Nevertheless, there is no direct way to order the independent components based on their contribution, unlike the case of principal components, as a result of the whitening operation.

To identify those independent components that provide the best discrimination between pathologies, we sort the independent components based on a two-step approach. In the first step, a table of p-values of the standard two-sample t-test, between all possible pairs of classes within the five classes of interest, is constructed based on the training data samples and a particular independent component. This results in a table containing ten distinct values. The second step involves the calculation of the number of p-values in this table above 5% (i.e. not statistically significant at the 5% level), which we will call here the discrimination index (DI). The independent components are then sorted according to this number. This ensures that the selected number of independent components will contain the ones that are most discriminating.

In this paper, the number of independent components is selected based on the discrimination index, whereby the first group contains those independent components with one or fewer p-values more than 5%, the second group contains those with two or fewer, and so on. The reason for this approach is to eliminate any bias in the classification outcome as a result of the arbitrary selection of independent components within the group having the same discrimination index.

2.4 Signal classification

The feature vectors obtained from PCA and ICA obtained from both the rectangular and Hamming windows are used to classify the different arrhythmia types using three types of statistical classifier, namely, minimum distance classifier, Bayes minimum-error classifier and voting k-nearest neighbour (KADAH *et al.*, 1996). The available data set was divided into training (design) and test subsets. The results of this classification are expressed in terms of the sensitivity and specificity of the outcomes. The sensitivity is computed as the conditional probability of detecting an abnormal rhythm when there is, in fact, an arrhythmia. On the other hand, the specificity is the probability of correct detection of the normal rhythm.

The minimum distance classifier assumes the classes to be similar in distribution and linearly separable. Hence, the decision hyperplanes are allocated halfway between the centres of clusters of different classes. Test samples are classified by each being assigned to the class that has the nearest mean vector to its feature vector. In a more general form, the Bayes minimum-error decision rule classifies an observation (i.e. a test sample) to the class that has the highest *a posteriori* probability among the five classes. The data set is assumed to have a Gaussian conditional density function, and the *a priori* probabilities are assumed to be equal for the five types. On the other hand, the voting k-nearest neighbour (kNN) technique is non-parametric and assigns a test sample to the class of the majority of its k closest neighbours. Here, we consider the nearest-neighbour classifier (i.e. $k = 1$). The proposed classifiers are utilised either to classify the two-class problem simply to detect abnormal rhythm (i.e. normal against abnormal), which we will call the detection problem, or to classify the five different classes of signal, which we call the classification problem.

3 Results

The proposed feature estimation techniques were implemented and applied to a large number of ECG signals obtained from the MIT-BIH arrhythmia database. The data set used for this work consisted of five different types, including normal (NR), ventricular couplet (VC), ventricular tachycardia (VT), ventricular bigeminy (VB) and ventricular fibrillation (VF). Each type was represented by 64 independent 3 s long signals in the training (design) data set. The testing data set consisted of another 32 independent signals of the same length from each type. The sampling rate of the VF signals was 250 samples s^{-1} and the other signals were sampled at 360 samples s^{-1}. The signals were Fourier transformed and reduced to include only the components corresponding to positive frequencies, described above.

To obtain the same size for all signals, only the first 375 frequency components were used in the subsequent analysis. This was achieved by simply truncating the number of frequency components beyond 375 when the sampling rate was 360 samples s^{-1}. Given that the sampling window length in time was the same, the frequency bin size was the same between the different sampling rates, and the above procedure was correct.

The PCA and ICA procedures were applied to compute a total of 320 principal and independent components, as the available data matrix was 320×375, given the number of training vectors and the length of the feature vector. The results of using both rectangular and Hamming windows were computed. Fig. 1 demonstrates the averaged 375-point spectrum for each of the five ECG signal types using a rectangular window. In Figs 2 and 3, the first 20 principal components and 20 sample independent components are illustrated.

Table 1 demonstrates the energy-packing property of the PCA, whereby a small percentage of all principal components account for most of the energy for both the rectangular and Hamming windows. The results for both the detection and classification problems using PCA, when rectangular and Hamming windows were used, are shown in Tables 2–7, using different numbers of principal components to demonstrate the efficiency of the technique. The detection and classification results using ICA for both rectangular and Hamming windows are shown in Tables 8–13. rhe number of independent components used is a direct function of the discrimination index (DI) described above.

The Bayes minimum-error classifier was computed up to the value of components that allowed the covariance matrix to have a stable inverse. As a result, its results contained one entry in the case of ICA results. As can be observed from the results, the
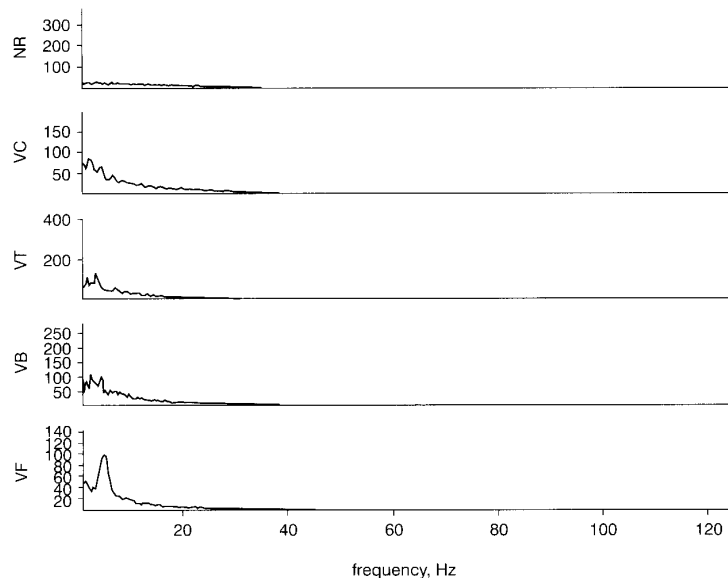
Fig. 1. *Plot of average positive half of spectrum of each of five ECG signal classes considered in this work, using a rectangular window*

nearest-neighbour classifier seemed to provide the best results, followed by the minimum-distance classifier. On the other hand, the results of the Bayes minimum-error classifier were rather poor and suggest that the distribution of the clusters may not be Gaussian as this technique assumes.

The first observation on the results is that the accuracy of the results improves as the number of principal/independent components increases, up to a certain value. After that, the accuracy may slightly deteriorate for some cases (see for example Tables 6, 7, 10 and 13). This indicates that there is indeed a part of the signal that contributes random noise to the classification problem. This is particularly evident in the case of PCA, where the additional principal components above a certain limit are associated with very small eigenvalues. This means that these components do not contribute much to the signal and can be ignored in the detection/classification process. This is apparent from the fact that the results seem to saturate beyond

a certain number of components. Hence, these results suggest the value of using either PCA or ICA for noise suppression and dimensionality reduction prior to classification with any other technique.

When we compare the results of using rectangular against Hamming windows, several observations can be made. First, the energy compactness of the Hamming window appears to be better than that of the rectangular window. Also, the detection results are generally considered better with the Hamming window than with the rectangular window. Nevertheless, the best detection result was obtained from the nearest-neighbour classification of ICA features using a rectangular window. Moreover, the classification results demonstrate a superior performance for the rectangular window. Table 14 shows a summary of the performance comparison of both windows. These results seem to support our hypothesis about the advantages of using the rectangular window.
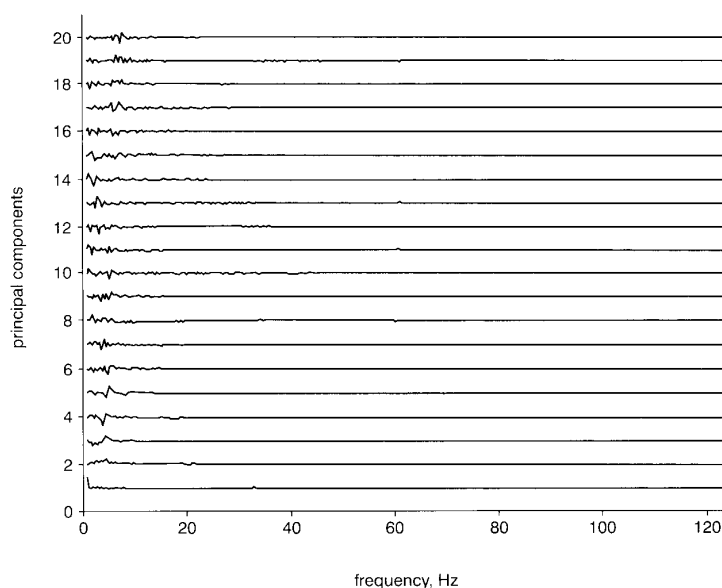

Fig. 2. *Plot of first 20 principal components obtained from all signals in training data set*
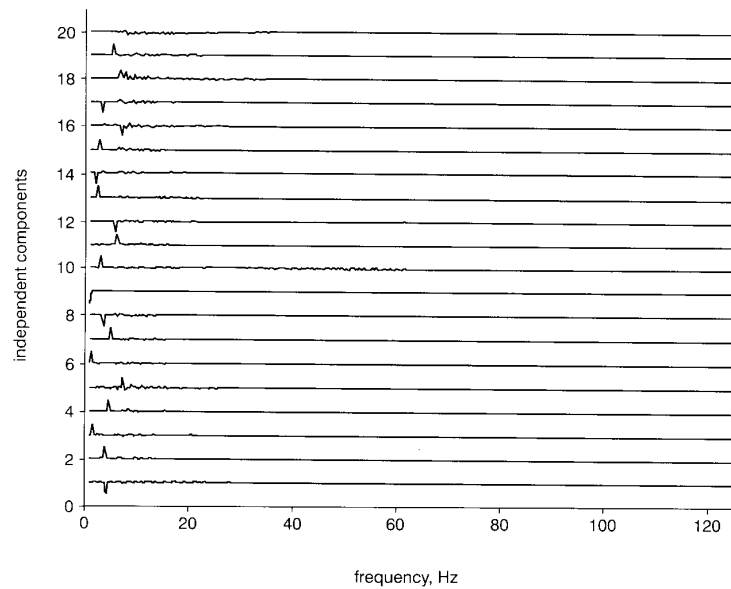
frequency, Hz

Fig. 3. *Plot of 20 sample independent components obtained from all signals in training data set*

Table 1   Efficiency of reducing dimensionality using PCA

q	320	100	80	60	40	30	20	15	10	5
Efficiency (rectangular)	100	99.97	99.92	99.7	99.1	98.2	96.0	93.3	88.3	78.4
Efficiency (Hamming)	100	99.98	99.95	99.8	99.4	98.7	96.9	95.2	91.8	83.4

Table 2   Minimum-distance classifier results for detection problem (PCA)

q	Rectangular		Hamming	
	specificity, %	sensitivity, %	specificity, %	sensitivity, %
5	68.8	95.3	65.6	96.1
10	71.9	94.5	75.0	96.1
15–320	71.9	94.3	78.1	96.1

Table 3   Bayes minimum-error classifier results for detection problem (PCA)

q	Rectangular		Hamming	
	specificity, %	sensitivity, %	specificity, %	sensitivity, %
5	65.6	91.4	59.4	93.8
10	56.3	96.9	56.3	100
15	40.6	98.4	46.9	100
20	46.9	99.2	34.4	100
30	34.4	100	40.6	100
40	31.3	100	34.4	99.2

4 Discussion

The results of PCA and ICA appear to become closer as the number of components included in the feature vector increases. This is a direct result of the fact that both PCA and ICA provide a complete basis set of vectors to describe the space of ECG signals. The differences between the two techniques are only apparent in the specific directions of each of these vectors. These differences tend to make the spanned subspaces obtained using the two techniques rather different when a small number of vectors are used. In the extreme case, when all vectors (in this case 375) are used, we can find an orthogonal transformation to transform the feature vector based on PCA to that based on ICA. Given that the classification techniques used here rely primarily on the Euclidean distance in assigning class membership, and as orthogonal transformations preserve Euclidean distance, it is not surprising to see that the classification results match in this special case and to realise the convergence of the two sets of results to the same solution.

The three classifiers implemented in this work appear to provide substantially different receiver operating characteristics, demonstrating the compromise between detection rates

Table 4   Nearest-neighbour classifier results for detection problem (PCA)

q	Rectangular		Hamming	
	specificity, %	sensitivity, %	specificity, %	sensitivity, %
5	59.4	94.5	65.6	97.7
10	75.0	93.0	75.0	96.9
15	90.6	96.1	87.5	97.7
20	93.8	96.1	93.8	96.9
30	96.9	96.1	90.7	97.7
40	96.9	96.9	90.7	96.1
60–320	96.9	97.7	90.7	97.7

Table 5 Minimum-distance classifier sensitivity (%) results for classification problem (PCA)

q	Rectangular					Hamming				
	NR	VC	VT	VB	VF	NR	VC	VT	VB	VF
5	68.8	68.8	53.1	21.9	78.1	65.6	62.5	56.3	6.25	78.1
10	71.9	71.9	56.3	25.0	81.3	75	62.5	59.4	6.25	81.3
15–320	71.9	75.0	56.3	25.0	81.3	78.1	62.5	59.4	9.4	81.3

Table 6 Bayes minimum-error classifier sensitivity (%) results for classification problem (PCA)

q	Rectangular					Hamming				
	NR	VC	VT	VB	VF	NR	VC	VT	VB	VF
5	65.6	78.1	62.5	46.9	71.9	59.4	68.8	65.6	71.9	68.8
10	56.3	65.6	71.9	68.8	78.1	56.3	68.8	84.4	81.3	84.4
15	40.6	65.6	81.3	81.3	75.0	46.9	62.5	87.5	75.0	87.5
20	46.9	65.6	84.4	62.5	87.5	34.4	59.4	87.5	87.5	93.8
30	34.4	53.1	81.3	68.8	90.6	40.6	50.0	87.5	84.4	100
40	31.3	40.6	84.4	71.9	87.5	34.4	56.3	81.3	81.3	90.7

Table 7 Nearest-neighbour classifier sensitivity (%) results for classification problem (PCA)

q	Rectangular					Hamming				
	NR	VC	VT	VB	VF	NR	VC	VT	VB	VF
5	59.4	53.1	65.6	56.3	71.9	65.6	59.4	68.8	84.4	75.0
10	75.0	71.9	71.9	71.9	78.1	75.0	59.4	68.8	84.4	84.4
15	90.6	75.0	68.8	84.4	81.3	87.5	59.4	75.0	84.4	90.6
20	93.8	78.1	68.8	84.4	87.5	93.8	62.5	84.4	87.5	90.6
30	96.9	75.0	71.9	84.4	87.5	90.6	71.9	75.0	93.8	90.6
40	96.9	71.9	75.0	87.5	87.5	90.6	68.8	75.0	90.6	87.5
60	96.9	68.8	71.9	84.4	90.6	90.6	68.8	75.0	90.6	90.6
80	96.9	68.8	71.9	81.3	87.5	90.6	71.9	75.0	90.6	90.6
100–320	96.9	68.8	71.9	81.3	90.6	90.6	71.9	75.0	90.6	90.6

Table 8 Minimum-distance classifier results for detection problem (ICA)

DI	Rectangular			Hamming		
	q	specificity, %	sensitivity, %	q	specificity, %	sensitivity, %
1	13	71.9	85.2	35	75.0	78.9
2	99	84.4	80.5	138	68.8	85.2
3	162	75.0	96.9	203	87.5	84.4
4	219	78.1	96.9	240	75.0	96.1
5–10	249–320	78.1	96.1	253–320	81.3	96.1

and false-alarm rates. The optimisation of this classification is beyond the scope of this paper. Nevertheless, the results of these classifiers provide a general conclusion about the classification accuracy and the upper limits in the sensitivity and specificity values obtainable using the proposed features.

The signal window length for this analysis can be arbitrarily chosen, provided it is less than 10 s. This is to satisfy the ANSI/AAMI EC13-1992 standard, which requires alarms for abnormal ECG signals to be activated within 10 s of their onset. Although increasing the window to the maximum possible size is desirable to obtain a better resolution in the frequency domain, this selection was made to ensure signal stationarity within the analysis window (WANG et al., 1998). The use of two sampling rate variations of the number of points within this duration was not found to be crucial, as long as the ECG signal was sufficiently sampled.

Table 9 Bayes minimum-error classifier results for detection problem (ICA)

DI	Rectangular			Hamming		
	q	specificity, %	sensitivity, %	q	specificity, %	sensitivity, %
1	13	53.1	93.0	35	46.9	98.4

Table 10 Nearest-neighbour classifier results for detection problem (ICA)

DI		Rectangular			Hamming	
	q	specificity, %	sensitivity, %	q	specificity, %	sensitivity, %
1	13	71.9	95.3	35	65.6	95.3
2	99	90.6	96.9	138	87.5	97.7
3	162	96.9	97.7	203	96.9	98.4
4	219	100	97.7	240	96.9	98.4
5	249	100	98.4	253	93.8	97.7
6–10	265–320	100	97.7	276–320	93.8	96.9

Table 11 Minimum-distance classifier sensitivity (%) results for classification problem (ICA)

DI			Rectangular						Hamming			
	q	NR	VC	VT	VB	VF	q	NR	VC	VT	VB	VF
1	13	71.9	68.8	59.4	37.5	53.1	35	75.0	43.8	40.6	25.0	46.9
2	99	84.4	62.5	62.5	53.1	46.9	138	68.8	46.9	59.4	53.1	56.3
3	162	75.0	78.1	53.1	21.9	84.4	203	87.5	53.1	65.6	56.3	50.0
4	219	75.0	75.0	56.3	18.8	84.4	240	75.0	59.4	59.4	6.3	84.4
5–10	249–320	78.1	78.1	56.3	21.9	87.5	253–320	81.3	59.4	59.4	9.4	81.3

Table 12 Bayes minimum-error classifier sensitivity (%) results for classification problem (ICA)

DI			Rectangular						Hamming			
	q	NR	VC	VT	VB	VF	q	NR	VC	VT	VB	VF
1	13	53.1	65.6	68.8	59.4	65.6	35	46.9	62.5	68.8	78.1	46.9

Table 13 Nearest-neighbour classifier sensitivity (%) results for classification problem (ICA)

DI			Rectangular						Hamming			
	q	NR	VC	VT	VB	VF	q	NR	VC	VT	VB	VF
1	13	71.9	59.4	59.4	59.4	65.6	35	65.6	46.9	46.9	62.5	75.0
2	99	90.6	56.3	65.6	71.9	87.5	138	87.5	62.5	68.8	68.8	84.4
3	162	96.9	68.8	68.8	71.9	87.5	203	96.9	62.5	75.0	90.6	96.9
4	219	100	68.8	68.8	84.4	87.5	240	93.8	68.8	75.0	93.8	93.8
5	249	100	68.8	71.9	81.3	90.6	253	93.8	65.6	68.8	90.6	90.6
6–10	265–320	100	65.6	71.9	81.3	90.6	276–320	93.8	71.9	75.0	90.6	90.6

Table 14 Summary of best performance of windowing functions in detection and classification problems using PCA and ICA as features

Classifier results		Detection		Classification	
		PCA	ICA	PCA	ICA
Minimum-distance	specificity	Hamming	Hamming		
	sensitivity	Hamming	rectangular	rectangular	rectangular
Bayes	specificity	rectangular	rectangular		
	sensitivity	Hamming	Hamming	rectangular	rectangular
Nearest-neighbour	specificity	rectangular	rectangular		
	sensitivity	Hamming	rectangular	rectangular	rectangular

It should be noted that the proposed methods classify ECG samples based on the presence and strength of the frequency components and ignoring their relative delays. This logic is particularly justified in those cases in which the heart rate becomes higher, as in tachycardia and fibrillation, where the frequency components become significantly different. Even though the intuitive justification in the other arrhythmia types may not be obvious, the results demonstrate the presence of significant differences between these types using only the magnitude information. The question of whether the inclusion of phase information would provide better discrimination remains for future work.

The results show that the use of PCA demonstrates a significant degree of energy compaction of the training samples. This result is not surprising and is similar to that of the previous work by SILIPO et al. (1995) and LAGUNA et al. (1999). However, the proposed transformation cannot directly be used for such applications as data compression for ECG signals, as the phase part of the signals was ignored. A more useful transformation for this application would be the discrete cosine transformation, whereby a $(2N$-$1)$-point symmetric signal is composed using the N-point ECG sample window. The analysis of such an implementation is beyond the scope of this paper.

5 Conclusions

Two blind source separation techniques were used to derive ECG signal features for arrhythmia detection and classification. A large database of ECG signals was used to compute a set of basic signal components that compose any ECG signal, using PCA and ICA. A set of features was obtained by projection of a given ECG signal onto the subspace of those basic signals, which were subsequently used for arrhythmia detection and classification using conventional statistical methods. The results indicate the value of such features for practical use in clinical settings.

References

AFONSO, V., and TOMPKINS, W. (1995): 'Detecting ventricular fibrillation, selecting the appropriate time-frequency analysis tool for the application', *IEEE Eng. Med. Biol. Mag.*, **15**, pp. 152–159
BIEL, L., PETTERSSON, O., PHILIPSON, L., and WIDE, P. (2001): 'ECG analysis: a new approach in human identification', *IEEE Trans. Instrum. Meas.*, **50**, pp. 808–812
BRACEWELL, R. (1986): 'The Fourier transform and its applications' (McGraw Hill, New York, 1986)
DE LATHAUWER, L., DE MOOR, B., and VANDEWALLE, J. (2000): 'Fetal electrocardiogram extraction by blind source subspace separation', *IEEE Trans. Biomed. Eng.* **47**, pp. 567–572
GERBRANDS, J. (1981): 'On the relationships between SVD, KLT, and PCA', *Pattern Recog.*, **14**, pp. 375–381
GOLUB, G., and VAN LOAN, C. (1993): 'Matrix computations' (The Johns Hopkins University Press, Baltimore, 1993)
GUILLÉN, S., ARREDONDO, M., MARTIN, G., and CORRAL, J. (1989): 'Ventricular fibrillation detection by autocorrelation function peak analysis', *J. Electrocardiol.*, **22**, pp. 253–262
HYVÄRINEN, A., and OJA, E. (1997): 'A fast fixed-point algorithm for independent component analysis', *Neural Comput.*, **9**, pp. 1483–1492
HYVÄRINEN, A., KARHUNEN, J., and OJA, E. (2001): 'Independent component analysis' (John Wiley & Sons, New York, 2001)
KADAH, Y., FARAG, A., ZURADA, J., BADAWI, A., and YOUSSEF, A. (1996): 'Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images', *IEEE Trans. Med. Imag.*, **15**, pp. 466–478
KHADRA, L., AL-FAHOUM, A., and AL-NASHASH, H. (1997): 'Detection of life-threatening cardiac arrhythmias using the wavelet transformation', *Med. Biol. Eng. Comput.*, **35**, pp. 626–632
LAGUNA, P., MOODY, G., GARCIA, J., GOLDBERGER, A., and MARK, R. (1999): 'Analysis of the ST-T complex of the electrocardiogram using the Karhunen-Loève transform: Adaptive monitoring and alternans detection', *Med. Biol. Eng. Comput.*, **37**, pp. 175–189
MINAMI, K., NAKAJIMA, H., and TOYOSHIMA, T. (1999): 'Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network', *IEEE Trans. Biomed. Eng.*, **46**, pp. 179–185
OUDA, B., TAWFIK, B., YOUSSEF, A., and KADAH, Y. (2001): 'Adaptive denoising technique for robust analysis of functional magnetic resonance imaging data'. IEEE Engineering in Medicine and Biology Conference, Istanbul, Turkey
PANERAI, R., FERRIERA, A., and BRUM, O. (1988): 'Principal component analysis of multiple noninvasive blood flow derived signals', *IEEE Trans. Biomed. Eng.*, **35**, pp. 533–538
PAPOULIS, A. (1991): 'Probability, random variables and stochastic processes' (WCB McGraw-Hill, New York, 1991)
SILIPO, R., LAGUNA, P., NARCHESI, C., and MARK, R. G. (1995): 'ST-T segment change recognition using artificial neural networks and principal component analysis', *Computers in Cardiology*, pp. 213–216
STAMKOPOULOS, T., DIAMANTARAS, K., MAGLAVERAS, N., and STRINTZIS, M. (1998): 'ECG analysis using nonlinear PCA neural networks for ischemia detection', *IEEE Trans. Signal Process.*, **46**, pp. 3058–3067
WANG, F., SAGAWA, K., and INOOKA, H. (1998): 'Time domain heart rate variability index for assessment of dynamic stress', *Computers in Cardiology*, pp. 97–100
ZARZOSO, V., NANDI, A., and BACHARAKIS, B. (1997): 'Maternal and foetal ECG separation using blind source separation methods', *IMA J. Math. Appl. Med. Biol.*, **14**, pp. 207–225

Authors' biographies

MOHAMED OWIS received his BSc and MSc in Biomedical Engineering from Cairo University in 1992 and 1996, respectively. He received his MSc degree in Computer Science from Old Dominion University in 1999. He received his PhD degree in Biomedical Engineering in January 2002. He is currently an Assistant Professor at the Biomedical Engineering Department, Cairo University. His research interests include computer networks, biomedical signal analysis and bioinformatics.

ABOU-BAKR YOUSSEF received his BSc and MSc in Electrical Engineering in 1974 and 1978, respectively. He received his medical MBBH from the medical school at Cairo University in 1980. He also received his PhD in Biomedical Engineering from Cairo University in 1982. He received his MD degree from Heidelberg University in 1989. He chaired the Biomedical Engineering Department at Cairo University between 1995 and 2001. He is currently a professor at the same department.

YASSER KADAH received his BSc and MSc in Biomedical Engineering from Cairo University in 1989 and 1992, respectively. He received his PhD in Biomedical Engineering from the University of Minnesota—Twin Cities in 1997. He is currently an Assistant Professor of Biomedical Engineering at Cairo University. His research interests include MRI, ultrasound imaging, and applied signal processing for biomedical problems.

Medical & Biological Engineering & Computing

Incorporating CELLULAR ENGINEERING

Contents   Volume 40   No. 5   September 2002