UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 12-06
Statistics and Econometrics Series (03)
February 2012

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

DISCRIMINANT ANALYSIS OF MULTIVARIATE TIME SERIES USING WAVELETS

Elizabeth Ann Maharaj[1] and Andrés M. Alonso[2]

Abstract

In analyzing ECG data, the main aim is to differentiate between the signal patterns of those of healthy subjects and those of individuals with specific heart conditions. We propose an approach for classifying multivariate ECG signals based on discriminant and wavelet analyzes. For this purpose we use multiple-scale wavelet variances and wavelet correlations to distinguish between the patterns of multivariate ECG signals based on the variability of the individual components of each ECG signal and the relationships between every pair of these components. Using the results of other ECG classification studies in the literature as references, we demonstrate that our approach applied to 12-lead ECG signals from a particular database, displays quite favourable performance. We also demonstrate with real and synthetic ECG data that our approach to classifying multivariate time series out performs other well-known approaches for classifying multivariate time series. In simulation studies using multivariate time series that have patterns that are different from that of the ECG signals, we also demonstrate very favourably performance of this approach when compared to these other approaches.

Keywords: Time series; Wavelet Variances; Wavelet Correlations; Discriminant Analysis.

JEL Classification: C38; C22.

[1] Department of Econometrics and Business Statistics, Monash University (Madrid), 900 Dandenong Road, Caulfield East, VIC 3145, Australia. E-mail: ann.maharaj@monash.edu
[2] Departamento de Estadística. Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), España. E-mail: andres.alonso@uc3m.es

1. Introduction

The classification of multivariate ECG signals is an important task in biomedical science. In order to determine whether or not an individual has a specific heart condition, up to ten electrodes are attached to the individual and the ECG signals obtained. Large databases exist for large numbers of patients and controls, and the task for researchers is to develop automated methods of classification of such time series. Many authors have proposed methods to this end with varying degrees of classification accuracy. In particular, Al-Naima and Al-Timemy [1], De Chazel and Reilly [2] , Heden et al. ([3] and Bozzola et al. [4] have proposed methods that been applied to 12-lead ECG signals.

In our study of 12-lead ECG signals , we are particularly interested in differentiating between patterns of ECG signals of individuals with the heart condition, myocardial infarction, and of those without this condition. To this end, Al-Naima and Al-Timemy [1] used Discrete Fourier Transform (DFT) coefficients and Discrete Wavelet Transform (DWT) coefficients as the discriminating features with neural network classifiers. Using 12-lead ECG data with a training set of 45 records (26 controls and 19 with myocardial infarction) and a test set of 20 records (12 controls and 8 with myocardial infarction), the sensitivities achieved for the test set were between 80% to 90%, while the specificity was 90%.

De Chazel and Reilly [2] used linear and quadratic discriminants with five features sets, including DWT coefficients, standard cardiology features and time domain features. Using a database of 500 12-lead ECG signals from 345 patients with different cardiac diseases, and from 155 controls, and multiple runs of ten-fold cross-validation, they obtained sensitivities of 69-73%, 78-83% and 25-38% when they classified anterior, inferior or combined myocardial infarction, while the specificity of their best classifier was 90%.

Heden et al. [3] conducted a study using 12-lead ECG records of a group of 1120 individuals with acute myocardial infarction and a control group consisting of 10,352 cases. They used six time domain measurements from each of the 12 leads as inputs in to artificial neural networks. Using a three-fold cross-validation procedure, their method achieved sensitivities of 46.2%-65.9% and specificities of 86.3-95.4%.

Bozzola et al. [4] extracted a set of 8 time domain parameters from each of the 12-lead ECG records and input these into a hybrid neuro-fuzzy system for the classification of myocardial infarction. They used a training set of ECG records of 179 controls and 404 with myocardial infarction, and a test set of ECG records of 60 controls and 135 with myocardial infarction. Their method achieved test set sensitivities of 72% , 80-88% and 52-60% when they classified anterior, inferior or combined myocardial infarction and specificities of 92-93%.

In all of the techniques employed by the above-mentioned authors, the components of the 12-lead ECG signal are treated as if they were independent of each other. In practice, each 12-lead ECG signal can be regarded as a 12-component multivariate time series. An important consideration in the analysis of multivariate time series is the relationship between the individual components of each series. Hence, given the absence of this consideration in ECG classification literature, we are motivated to examine the inclusion of these interrelationship features to assess whether they provide useful information, and thus lead to perhaps more accurate classification results. In this paper, we therefore propose the discriminant analysis of the multivariate time series, namely, the 12-lead ECG signals, based on wavelet features of variances and correlations. In using these features, our goal is to distinguish between the patterns of multivariate ECG signals based on the variability of the individual components of each ECG signal, and the relationships between these components. Taking into account the relationships between every pair of components is a novel approach, and to our knowledge has not been considered before in multiple-lead ECG classification. Furthermore, an advantage of using wavelet features is that the time series do not necessarily have to be mean or variance stationary as in deed most ECG data are not. Using data from a publicly available ECG database, we show that discriminant analysis based on the wavelet features of variance and correlations achieves exceptional accuracy rates in differentiating between the 12-lead ECG records of healthy individuals and those with myocardial infarction. Using simulated data we also demonstrate that our approach of classifying multivariate time series outperforms other well- known methods

In Section 2, we will describe the tools required for our proposed method and the method itself. In Section 3, we apply this method to the ECG data and evaluate its performance using the hold-out-one cross-validation technique. In Section 4 we evaluate the performance of this approached using synthetic ECG data when compared to the other approaches. In Section 5, we carry out other studies with simulated data to evaluate the performance of this method on multivariate time series that have different patterns to that of conventional ECG signals. We conclude in Section 6.

2. Methods

2.1. Discriminant Analysis

In practice, a time series is known to belong to one of g groups. The task is to classify the time series into one of these g groups in an optimal manner. Assumptions are made concerning the Gaussian probability density function of the different groups. In linear discriminant analysis, it is assumed that the groups have equal covariance matrices and differ only in their means. While in quadratic discriminant analysis, the covariance matrices of the groups are not assumed to be equal. There are several ways to evaluate the performance of a discriminant analysis procedure. One method is to split the sample into training and hold-out samples and evaluate the error rate associated with the hold-out set which was not used in deriving the classification rule. Another method is to use the hold-out-one technique of cross-validation which is particularly useful if the samples sizes are not very large. This technique holds out the observation to be classified, deriving the classification function from the remaining observations (see [5] for more details). The procedure is repeated of each member of the sample and an overall error rate is determined.

When dealing with univariate time series, one can use features of the time series such as autocorrelations, periodogram coefficients, wavelet features etc., in a standard discriminant analysis. Several authors have proposed discriminant analysis of univariate time series. See for example, [6] and [7]. With multivariate time series, the task is more complex because as well as taking into account the multiple series associated with each object, one has to also take into account the relationships between the components of each multivariate time series. Kakizawa et al. [8] developed approaches for classifying stationary multivariate time series using spectral matrices with Kullback–Leibler (K–L) and Chernoff discrepancy measures. Shumway [9] extended this with the Kullback–Leibler (K–L) discrepancy to locally stationary time series.

2.2. Wavelet Analysis and Wavelet Features

In what follows, using the notation of [10], we give a brief description of wavelet analysis and the associated features of wavelet variances and wavelet correlations.

The Discrete Wavelet Transform (DWT), which is an orthonormal transform, re-expresses a time series of length T in terms of coefficients that are associated with a particular time and with a particular dyadic scale as well as one or more scaling coefficients. The j-th dyadic scale is of the form 2^{j-1} where $j = 1, 2, \ldots, J$, and J is the maximum allowable number of scales.

The number of coefficients at the j-th scale is $T/2^j$, provided $T = 2^J$. In general the wavelet coefficients at scale 2^{j-1} are associated with frequencies in the interval $[1/2^{j+1}\ ,1/2^j]$. Large time scales give more low frequency information, while small time scales give more high frequency information about the time series. It is possible to recover the time series $X_t, t = 1, 2, \ldots, T$ from its DWT by synthesis. That is, the multi-resolution analysis (MRA) of a time series is expressed as

$$X_t = \sum_{j=1}^{J} d_j + s_J, \tag{2.1}$$

where d_j is the wavelet detail (series of inverse wavelet coefficients at scale j) and s_J is the smooth series which is the inverse of the series of scaling coefficients. Hence a time series and its DWT are actually two representations of the same mathematical entity.

The maximum overlap discrete wavelet transform (MODWT) is a modification of the DWT. Under the MODWT, the number of wavelet coefficients created will be the same as the number of observations in the original time series. Because the MODWT decomposition retains all possible times at each time scale, the MODWT has the advantage of retaining the time invariant property of the original time series. The MODWT can be used in a similar manner to the DWT in defining a multi-resolution analysis of a given time series.

In contrast to the DWT, the MODWT details and smooths are associated with zero phase filters making it easy to line up features in a MRA with the original time series more meaningfully.

If $\{\tilde{h}_{j,l}, l = 0, 1, \ldots, L_j\}$ is the j-level MODWT wavelet filter of length L_j, associated with scale $\tau_j \equiv 2^{j-1}$ then if $\{X_t\}$ is a discrete parameter stochastic process and

$$W_{j,t} \equiv \sum_{l=0}^{L_j} \tilde{h}_{j,l} X_{t-l} \tag{2.2}$$

represents the stochastic process by filtering $\{X_t\}$ with the MODWT filter $\{\tilde{h}_{j,l}\}$, and if it exists and is finite, the time independent MODWT wavelet variance at the j-th dyadic scale $\tau_j \equiv 2^{j-1}$ is defined as

$$\nu_X^2(\tau_j) \equiv \mathrm{var}\{W_{X,j,t}\}. \tag{2.3}$$

It can be shown that

$$\sum_{j=1}^{\infty} \nu_X^2(\tau_j) = \mathrm{var}\{X_t\}, \tag{2.4}$$

i.e., the wavelet variance decomposes the variance of the stochastic process across scales (see [10], p296-298 for more details).

Given a time series $x_t, t = 1, 2 \ldots, T$, which is a realization of the stochastic process X_t, an unbiased estimator of $\nu_X^2(\tau_j)$ is

$$\widehat{\nu}_X^2(\tau_j) \equiv \frac{1}{M_j} \sum_{t=L_j-1}^{T-1} \widehat{W}_{X,j,t}^2, \tag{2.5}$$

where $\widehat{W}_{j,t}^2$ are the MODWT coefficients associated with the time series x_t and $M_j = N - L_j + 1$ is the number of wavelet coefficients excluding the boundary coefficients that are affected by the circular assumption of the wavelet filter.

Given two appropriate stochastic processes $\{X_t\}$ and $\{Y_t\}$ with MODWT coefficients $W_{Xj,t}$ and $W_{Yj,t}$, respectively, the wavelet covariance is defined as

$$\nu_{XY}(\tau_j) \equiv \mathrm{cov}\{W_{Xj,t}, W_{Yj,t}\}, \tag{2.6}$$

and it gives the scale-based decomposition of the covariance between $\{X_t\}$ and $\{Y_t\}$. The wavelet covariance can be standardized to yield the wavelet correlation

$$\rho_{XY}(\tau_j) \equiv \frac{\nu_{XY}(\tau_j)}{\nu_X(\tau_j)\nu_Y(\tau_j)}. \tag{2.7}$$

For time series x_t and y_t which are realizations of $\{X_t\}$ and $\{Y_t\}$ respectively, replacing the wavelet variances and covariance by their unbiased estimators, we get the estimated wavelet correlation

$$\widehat{\rho}_{XY}(\tau_j) \equiv \frac{\widehat{\nu}_{XY}(\tau_j)}{\widehat{\nu}_X(\tau_j)\widehat{\nu}_Y(\tau_j)}. \tag{2.8}$$

In what follow, we will apply discriminant analysis to sets of multivariate time series via their wavelet variances and wavelet correlations.

2.3. Remark

Assumptions of multivariate normality are made about the probability distribution of the group feature variables in linear and quadratic discriminant analysis. It can be shown that under the assumption of multivariate normality, the sample linear and quadratic discriminant functions are asymptotically optimal in the presence of homoscedasticity and heteroscedasticity, respectively (see [11]) .

Table 1
Maximum allowable number of scales for $T = 2^J$.

Wavelet filter	DB2	DB4	DB6	DB8	SYM8	CF6
Number of Scales	*J*	*J-1*	*J-2*	*J-2*	*J-2*	*J-2*

In our application and simulation studies, the group feature variables of the time series, that are used are MODWT wavelet variances and wavelet correlations. Serroukh et al. [12] have shown that MODWT wavelet variance estimators are asymptotically normal for linear processes, while Serroukh and Walden [13] have shown that for bivariate linear processes, the MODWT wavelet covariance estimators are asymptotically normal. It follows that the MODWT wavelet correlation estimators are also asymptotically normal. While a situation is which all variables under consideration are shown to exhibit univariate normality may help achieve multivariate normality, it will not guarantee it. Thus, the group feature variables of the combined MODWT wavelet variances and wavelet correlations may not necessarily be asymptotically multivariate normal.

In most real applications of discrimination analysis, the assumption of multivariate normality might not be strictly met. Many authors have conducted studies on the robustness of the discriminant functions and have found that some of them are fairly robust to departures from assumed models with little or no modification (see e.g., [14], [15]. Furthermore since asymptotic normality of the MODWT wavelet variance and of the MODWT wavelet covariance are based on the assumption that the underlying univariate and bivariate processes are linear (see [12] and [13], this assumption of linearity will not necessarily be met for the ECG data under consideration in Sections 3, 4 and 5. However, we will proceed with using the wavelet feature variables in linear and quadratic discriminant analysis even though the multivariate normality assumption may not be strictly satisfied by these feature variables.

3. Discriminant Analysis of ECG data

The data analyzed corresponds to the PTB Diagnostic ECG Database set available at the Physionet website http://www.physionet.org/physiobank/database/ptbdb. The freely available data is a small subset of the database used by *cardioPATTERN - Telemedical ECG-Evaluation and Follow up* that have a patented procedure for ECG classification (see details in http://radib.dyndns.org).

In this application, we are interested in distinguishing between ECG signals of individuals with myocardial infarction and those of healthy controls. The available dataset consists of 200 records of the conventional 12 leads (i, ii, iii, avr, avl, avf, v1, v2, v3, v4, v5, v6) for 148 patients with myocardial infarction and 52 healthy controls.

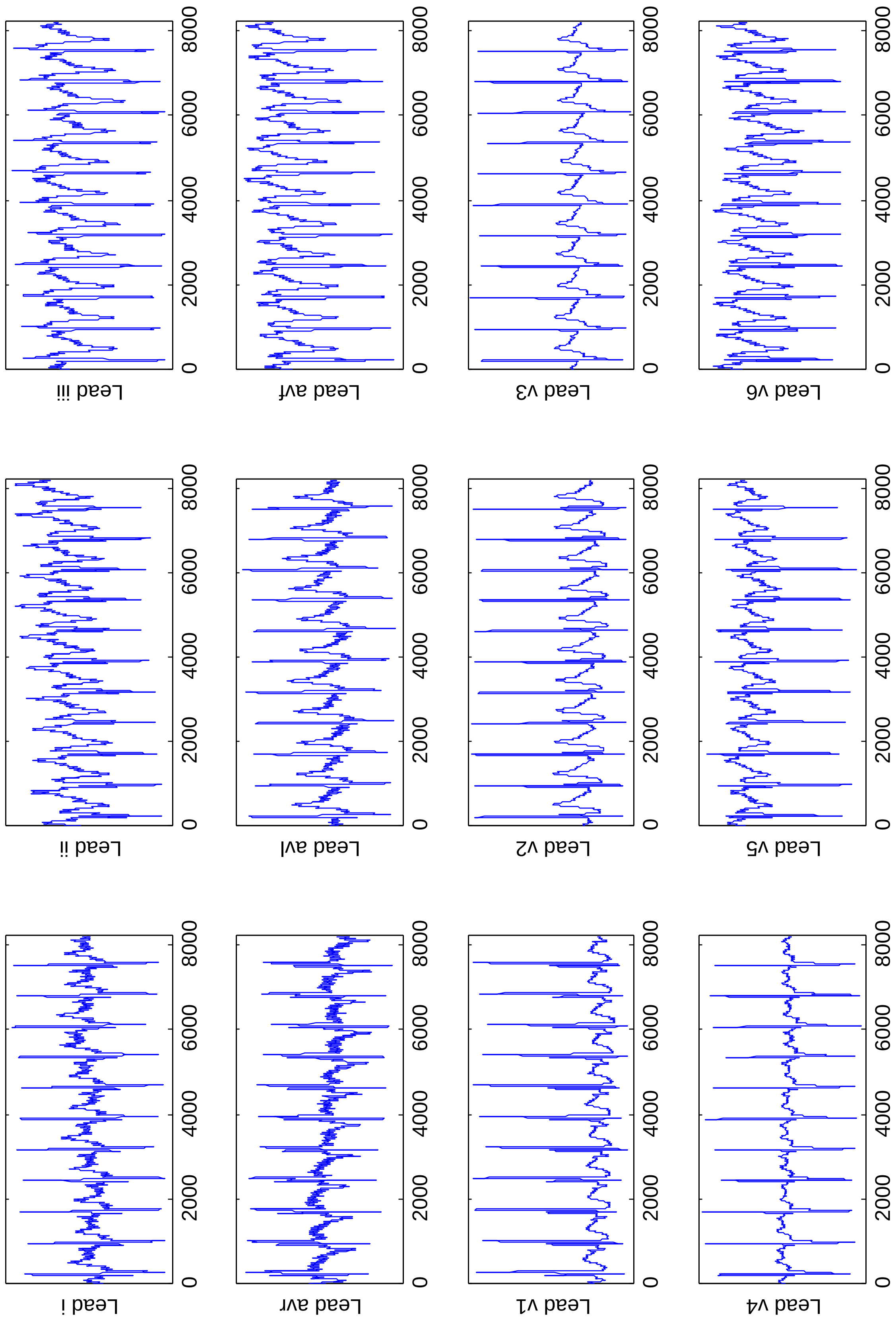
For each record, we read the first $2^{12} + 2^{13}$ observations. We discard the first 2^{12} observations since they could have some exogenous anomalies. The remaining $2^{13} = 8192$ observations correspond to around eleven heart beats. Figure 1 on page 6 and Figure 2 on page 7 present the 12-lead ECG signals for a patient with myocardial infarction and for a healthy control, respectively. It is clear there are some difference between the patterns of two type of signals.

Wavelet filters of lengths 2, 4, 6 and 8 of the Daubechies family (DB2, DB4, DB6, DB8), of length 8 from the Symmletts family (SYM8, and of length 6 from the Coiflets family (CF6) were used to generate the MODWT coefficients, and hence the MODWT variances and correlations of the signal. For more details on the wavelet filters refer to Chapter 4 of [10]. We used the stepwise implementation of [16] that selects the relevant variables (in this case the variables being the wavelet variances and correlations from the various scales) in order to minimize the misclassification error. We used the hold-out-one cross-validation technique to evaluate the performance of our method.

First, we apply the proposed procedure to the 12-dimensional signal. Table 1 on page 5 shows the maximum allowable number of scales for each of the filters for series length 2^J (see [10] p. 136 for more details). This is to ensure that the boundary coefficients which have an effect on the estimated scale by scale wavelet variance and correlation coefficients are excluded. In this case $J = 13$.

Table 2 on page 8 shows the misclassification rates for patients with myocardial infarction, for healthy controls and the overall misclassification rates using linear and quadratic discriminant analyses, with the wavelet variances (var), wavelet variances and wavelet correlations (var-corr) and wavelet correlations (corr). Table 2, also shows the results using the Kullback-Leibler discrimination information and the Chernoff

Fig 1. *ECG of a patient with myocardial infarction.*



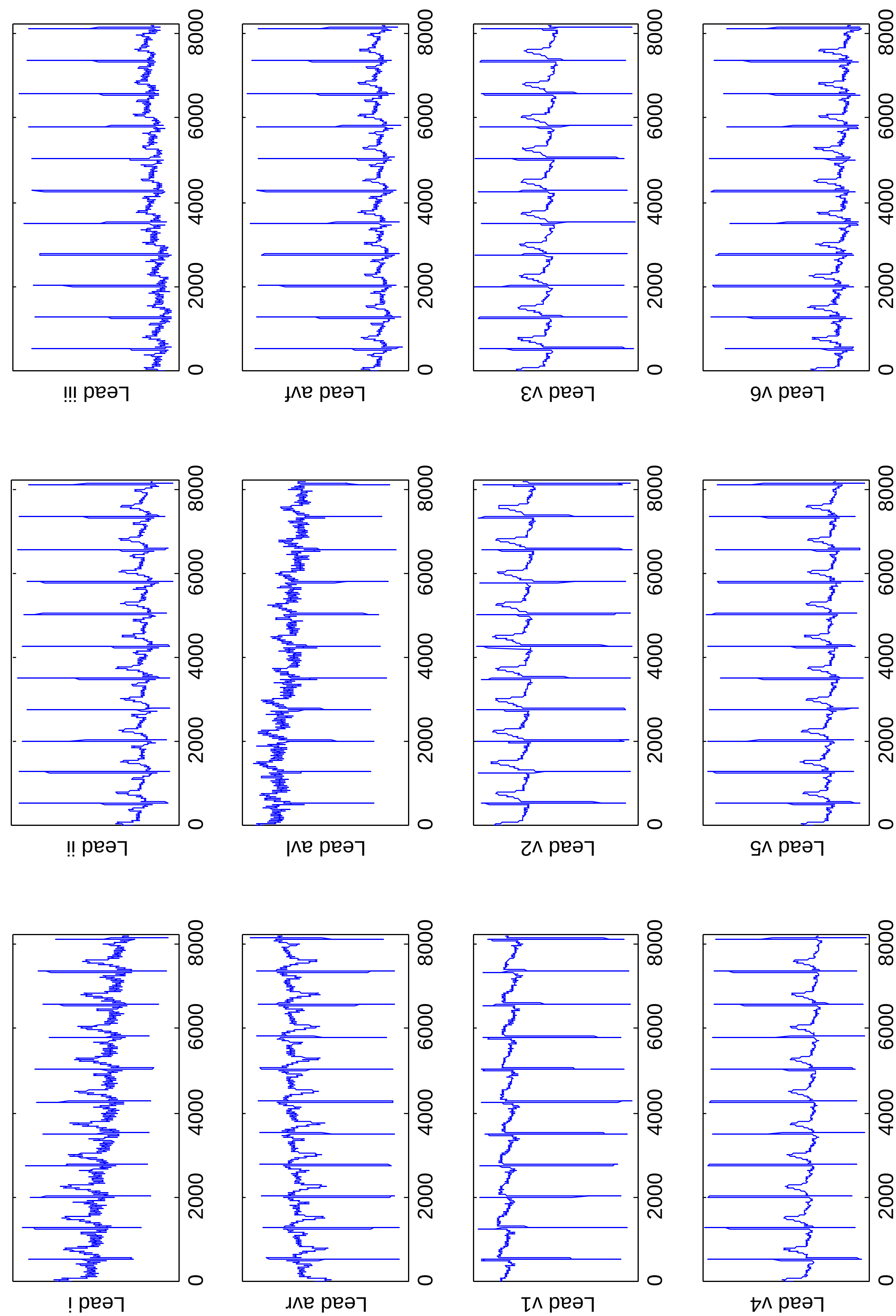


FIG 2. *ECG of a healthy control.*

Table 2
Misclassification rates for stepwise discriminant analysis and for the Kullback-Leibler and the Chernoff information measures applied to 12-lead ECG signals.

			Myocardial infarction	Healthy controls	Overall
Linear	DB2	var	12.80	13.50	13.00
		varcor	**4.70**	**5.80**	**5.00**
		cor	4.70	5.80	5.00
	DB4	var	12.20	15.40	13.00
		varcor	**4.10**	**3.80**	**4.00**
		cor	3.40	7.70	4.50
	DB6	var	13.50	11.50	13.00
		varcor	**5.40**	**7.70**	**6.00**
		cor	6.10	7.10	6.50
	DB8	var	14.90	14.90	13.50
		varcor	**5.40**	**5.40**	**5.80**
		cor	6.10	6.10	5.80
	SYM8	var	14.90	14.90	13.50
		varcor	**5.40**	**5.40**	**5.80**
		cor	6.10	6.10	5.80
	CF6	var	13.50	13.50	11.50
		varcor	**4.10**	**4.10**	**3.80**
		cor	5.40	5.40	9.60
Quadratic	DB2	var	14.20	26.90	17.50
		varcor	**2.00**	**23.10**	7.50
		cor	6.10	11.50	7.50
	DB4	var	10.80	26.90	15.00
		varcor	**3.40**	**28.80**	**10.00**
		cor	4.10	26.90	10.00
	DB6	var	13.50	23.10	16.00
		varcor	**4.10**	**23.10**	**9.00**
		cor	4.10	30.80	11.00
	DB8	var	12.20	25.00	15.50
		varcor	**4.10**	**19.20**	**8.00**
		cor	1.40	32.70	11.50
	SYM8	var	12.50	25.00	15.50
		varcor	**4.10**	**19.20**	**8.00**
		cor	4.10	32.70	11.50
	CF6	var	13.50	23.10	16.00
		varcor	**4.70**	**26.90**	**10.50**
		cor	3.40	34.60	11.50
Kullback–Leibler			9.60	43.90	35.00
Chernoff			1.90	93.90	70.00

information measures. These measures have been described in detail by Kakizawa et al. [8] and by Shumway and Stoffer [17].

The best results were obtained by our approach when the wavelet variances and correlations were both the input variables. When only wavelet variances were the input variables, the misclassification rates were much higher. When only wavelet correlations were the input variables, in a few cases the misclassification rates were similar to, or sometimes smaller than when both wavelet variance and correlations were input together. It is clear that the wavelet correlations provide useful information about the relationships between the leads of each ECG signal and hence make an important contribution to distinguishing between the patterns of ECG signals of individuals with myocardial infarction, and those of healthy controls.

Regarding the overall classification rates, the linear procedure generally outperformed the quadratic procedure. Only in few cases does the quadratic procedure outperform the linear procedure when classifying myocardial infarction ECGs. In general, the error rates were fairly similar for the different wavelet filters. All wavelet-based discriminant procedures outperformed the Kullback-Leibler information and Chernoff information procedures. For Kullback-Leibler information and Chernoff information procedures, we consider different values for the bandwidth used to estimate the spectral densities. The considered bandwidths were in the range [0.001, 0.01] that corresponds to from 9 to 81 contiguous fundamental frequencies that are close to the frequency of interest (see [17] p. 197 for more details).

These classification results imply sensitivities of 94.6-95.9% (95.3-98.0%) and specificities of 92.3-96.2%

TABLE 3
Misclassification rates for stepwise discriminant analysis and for the Kullback-Leibler and the Chernoff information measures applied to 3-factor ECG signals.

			Myocardial infarction	Healthy controls	Overall
Linear	DB2	var	18.20	23.10	19.50
		varcor	**18.20**	**21.20**	**19.00**
		cor	28.40	28.80	28.50
	DB4	var	19.60	19.20	19.50
		varcor	**14.90**	**19.20**	**16.00**
		cor	24.30	28.80	25.50
	DB6	var	23.00	21.20	22.50
		varcor	**16.20**	**21.20**	**17.50**
		cor	25.70	25.00	25.00
	DB8	var	22.30	22.30	21.20
		varcor	**16.20**	**16.20**	**17.30**
		cor	27.00	27.00	30.80
	SYM8	var	22.30	22.30	21.20
		varcor	**16.20**	**16.20**	**17.30**
		cor	27.00	27.00	30.80
	CF6	var	24.30	24.30	21.20
		varcor	**14.20**	**14.20**	**17.30**
		cor	24.30	24.30	26.90
Quadratic	DB2	var	9.50	50.00	20.00
		varcor	**22.30**	**25.00**	**23.00**
		cor	13.50	59.60	25.50
	DB4	var	18.20	30.80	21.50
		varcor	**7.40**	**42.30**	**16.50**
		cor	9.50	55.80	21.50
	DB6	var	14.20	42.30	21.50
		varcor	**6.10**	**51.90**	**18.00**
		cor	10.10	50.00	20.50
	DB8	var	16.20	40.40	22.50
		varcor	**7.40**	**59.60**	**21.00**
		cor	9.50	57.70	22.00
	SYM8	var	16.20	40.40	22.50
		varcor	**7.40**	**59.60**	**21.00**
		cor	9.50	57.70	22.00
	CF6	var	14.90	46.20	23.00
		varcor	**8.80**	**46.20**	**18.50**
		cor	12.20	53.80	23.00
Kullback–Leibler			17.30	62.80	48.00
Chernoff			5.70	87.70	63.50

(73.1-80.8%) when using the linear (quadratic) procedure with wavelets variances and correlations together. The procedures based on Kullback-Leibler information and Chernoff information have good sensitivities (90.4 and 98.1, respectively) but very poor specificities (56.1 and 6.1, respectively). The classification of a new signal took, using a wavelet based procedure, around five seconds.

Second, it order to consider a low dimensional signal and to reduce the computational time, we performed a dynamic factor analysis and we extracted the first three common factors for each ECG. Then, the proposed procedure was applied to the 3-dimensional signals formed by the three extracted common factors.

Table 3 on page 9 shows the misclassification rates using linear and quadratic discriminant analyses, respectively, with the wavelet variances, wavelet variances and correlation and wavelet correlations. This table also shows the results using the Kullback-Leibler discrimination information and the Chernoff information measure. In general, the results were better when we combined wavelets variances and wavelets correlations. In this case, the linear procedure is clearly superior to the quadratic procedure. In fact, in some cases, the quadratic procedure produces poor results.

While the classification of a new signal takes less than one second, i.e., the computational time was considerably reduced, it is clear that the overall misclassification rates for the 3-factor ECG signals were not as good when using the 12-dimensional signal. Reducing the 12-lead signals to 3-factor signals appears to result in the loss of useful information. The classification results for the 3-dimensional signals imply sensitivities of 81.8-85.8% and specificities of 78.8-82.7% when using linear procedure with wavelet variances and correlation

TABLE 4
Parameters of the synthetic ECG model.

Index(i)	1	2	3	4	5	6	7	8	9	10	11
α_i^x (mV)	0.03	0.08	-0.13	0.85	1.11	0.75	0.06	0.10	0.17	**0.39**	0.03
b_i^x (rad)	0.09	0.11	0.05	0.04	0.03	0.03	0.04	0.60	0.30	0.18	0.50
θ_i^x (rad)	-1.09	-0.83	-0.19	-0.07	0.00	0.06	0.22	1.20	1.42	1.68	2.90
α_i^y (mV)	0.04	0.02	-0.02	0.32	0.51	-0.32	0.04	**0.08**	0.01		
b_i^y (rad)	0.07	0.07	0.04	0.06	0.04	0.06	0.45	0.30	0.50		
θ_i^y (rad)	-1.10	-0.90	-0.76	-0.11	-0.01	0.07	0.80	1.58	2.90		
α_i^z (mV)	-0.03	-0.14	-0.04	0.05	-0.40	0.46	-0.12	-0.20	**-0.35**	-0.04	
b_i^z (rad)	0.03	0.12	0.04	0.40	0.05	0.05	0.80	0.40	0.20	0.40	
θ_i^z (rad)	-1.10	-0.93	-0.70	-0.40	-0.15	0.10	1.05	1.25	1.55	2.80	

together. Again, the procedures based on Kullback-Leibler information and Chernoff information have good sensitivities (82.7% and 94.3%, respectively) but very poor specificities (37.2% and 12.3%, respectively).

The cardioPATTERN procedure, using a database of 8500 ECGs from 3781 patients with different cardiac diseases, and 4719 normal persons, obtained a sensitivity of 72% and a specificity of 80% when classifying myocardial infarction (http://radib.dyndns.org/ekg/fverfahren_9_1e.html). Although this database as well as those used by [4], [3], [2] and [1] are not comparable with the PTB database, we can use their sensitivity and specificity results as reference values. Clearly, the results obtained with our procedure with the 12-lead ECG signals are quite favourable. Furthermore, given that for most of the time we achieved our best results when both wavelet variances and correlations were input together in the discriminant procedures, it is clearly an indication that the relationships between the components of the multi-lead ECG signals provide useful information for classification.

4. Discriminant Analysis of Synthetic ECG data

In this section, we evaluate the performance of the discriminant analysis with the wavelet features using synthetically generated ECG data. We use the generator available at the Open-Source Electrophysiological Toolbox (http://www.oset.ir/) that implements the dynamic model developed by McSharry et al. [18] and extended to multichannel ECG by Sameni et al. [19] and Clifford et al. [20]. The model is a three-dimensional formulation of single dipole of the heart:

$$\begin{aligned}
\dot{\theta} &= \omega \\
\dot{x} &= -\sum_i \frac{\alpha_i^x \omega}{(b_i^x)^2} \Delta\theta_i^x \exp\left[\frac{-(\Delta\theta_i^x)^2}{(b_i^x)^2}\right] \\
\dot{y} &= -\sum_i \frac{\alpha_i^y \omega}{(b_i^y)^2} \Delta\theta_i^y \exp\left[\frac{-(\Delta\theta_i^y)^2}{(b_i^y)^2}\right] \\
\dot{z} &= -\sum_i \frac{\alpha_i^z \omega}{(b_i^z)^2} \Delta\theta_i^z \exp\left[\frac{-(\Delta\theta_i^z)^2}{(b_i^z)^2}\right],
\end{aligned}$$

Figure 3 on page 11 shows a three lead ECG generated by this procedure.

In order to generate two different population using this procedure, we generate a sample of random ECG signals with the parameters in Table 4 and another sample using the same parameters with the exceptions of $a_{10}^x = 0.39 * \lambda$, $a_8^y = 0.08 * \lambda$ and $a_9^z * \lambda$ with $\lambda > 1$.

Figure 4 on page 12 illustrates the differences in the two generated signals using $\lambda = 1$ and $\lambda \in \{1.25, 1.5, 1.75, 2\}$. The signals generated with $\lambda > 1$ show a hyperacute T wave which is the first manifestation of acute myocardial infarction. For each population, we generate 100 ECG of length $T = 2^{12} = 4096$ and we apply the different discriminant procedures with the wavelet features and the Kullback-Leibler discrimination information and the Chernoff information procedures. As in the previous section, for the wavelet-based procedures we used wavelet filters from the Daubechies family (DB2, DB4, DB6, DB8), from the Symmletts family (SYM8) and from the Coiflets family (CF6) to generate the MODWT coefficients, and hence the MODWT variances and correlations.

Refer to Table 1 on page 5 for the maximum allowable number of scales for each of the filters for series of lengths $T = 2^{12}$. As before we used the stepwise implementation of [16]. One hundred simulations were carried out each time and an average overall misclassification rate was determined.

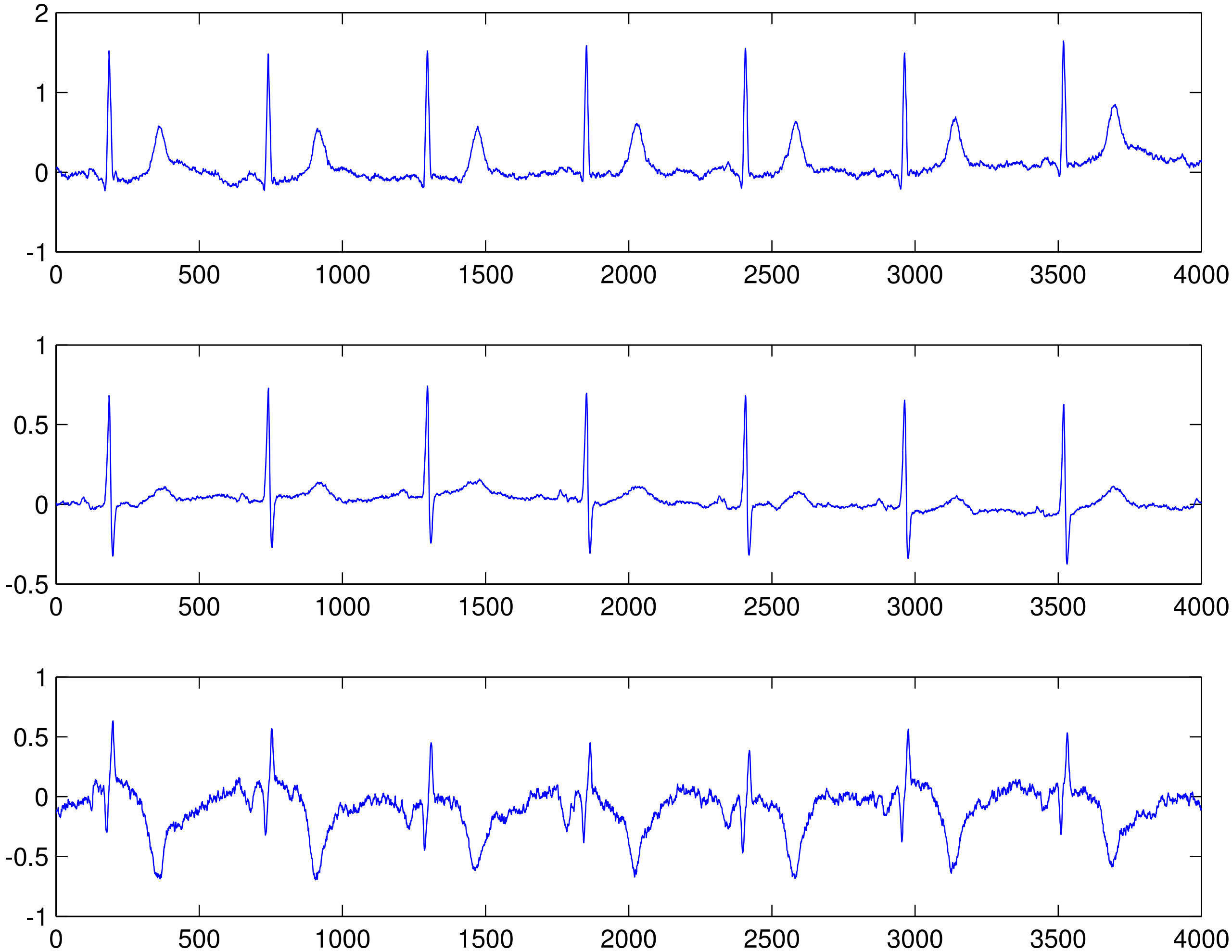


Fig 3. *Synthetic ECG signals.*

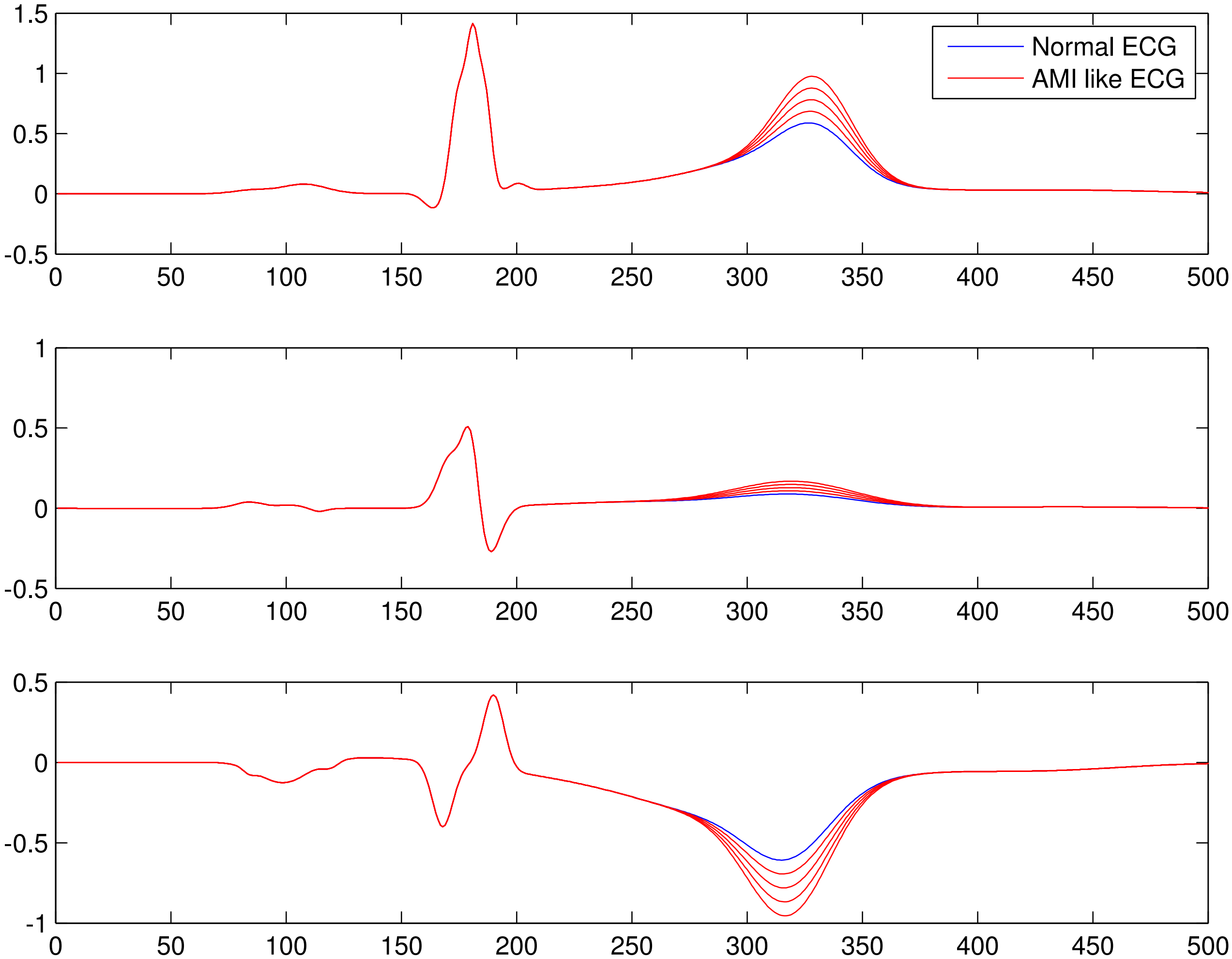


Fig 4. *Single beat of a synthetic ECG signals: Normal and Acute Myocardial Infarction.*

Table 5
Overall average misclassification rates for stepwise discriminant analysis and for the Kullback-Leibler and the Chernoff information measures applied to synthetic ECG signals.

			λ=1.25	λ=1.50	λ=1.75	λ=2.00
Linear	DB2	var	25.66	8.42	2.61	0.66
		varcor	**23.97**	**8.62**	**2.98**	**0.70**
		cor	38.12	27.72	20.85	15.46
	DB4	var	23.76	7.05	1.86	0.37
		varcor	**22.15**	**6.65**	**1.78**	**0.39**
		cor	38.80	28.47	22.15	16.82
	DB6	var	23.71	7.00	1.66	0.39
		varcor	**21.20**	**6.02**	**1.36**	**0.35**
		cor	39.91	29.20	22.92	17.45
	DB8	var	23.79	6.95	1.67	0.39
		varcor	**20.78**	**5.81**	**1.44**	**0.33**
		cor	39.93	29.53	22.92	17.72
	SYM8	var	23.79	6.95	1.67	0.39
		varcor	**20.78**	**5.81**	**1.44**	**0.33**
		cor	39.93	29.53	22.92	17.72
	CF6	var	23.94	7.00	1.77	0.43
		varcor	**22.16**	**6.52**	**1.76**	**0.40**
		cor	39.92	28.79	22.51	17.09
Quadratic	DB2	var	35.23	13.41	4.64	1.13
		varcor	**31.03**	**12.00**	**3.91**	**1.06**
		cor	43.68	31.60	23.19	15.67
	DB4	var	33.32	12.03	3.51	0.74
		varcor	**29.05**	**9.63**	**2.52**	**0.54**
		cor	44.05	32.90	24.95	16.95
	DB6	var	32.96	11.86	3.23	0.73
		varcor	**28.39**	**8.62**	**2.13**	**0.47**
		cor	44.64	33.61	25.49	18.49
	DB8	var	32.97	11.66	3.13	0.65
		varcor	**27.73**	**8.39**	**2.23**	**0.50**
		cor	44.50	33.75	26.09	19.05
	SYM8	var	32.97	11.66	3.13	0.65
		varcor	**27.73**	**8.39**	**2.23**	**0.50**
		cor	44.50	33.75	26.09	19.05
	CF6	var	33.67	12.16	3.31	0.75
		varcor	**29.48**	**9.62**	**2.57**	**0.53**
		cor	44.69	33.54	25.26	17.08
Kullback–Leibler			46.47	39.12	33.88	27.94
Chernoff			47.41	41.09	36.35	30.98

In Table 5 on page 13 we show the average misclassification rates for the synthetic ECG signals using linear and quadratic discriminant analyses, respectively, with the wavelet variances (var), wavelet variances and wavelet correlations (var-corr) and wavelet correlations (corr) and also using the Kullback-Leibler discrimination information and the Chernoff information measures. From this table, we observe that the best results were obtained by using the wavelet variances and correlations together. In this case, when only wavelet correlations were the input variables, the misclassification rates were much higher. However, wavelet correlations provide useful information to discriminate since the combined results are better than results using only wavelet variances. All wavelet based discriminant procedures outperformed the Kullback-Leibler information and Chernoff information procedures. For Kullback-Leibler information and Chernoff information procedures, we consider different values for the bandwidth used to estimate the spectral densities. The considered bandwidths were in the range [0.001, 0.01] that corresponds to from 5 to 41 contiguous fundamental frequencies that are close to the frequency of interest (see [17] p. 197 for more details).

In Figure 5 on page 14, we present the boxplots of the misclassification rate estimates for $\lambda = 1.25$ and 1.5 which are the most challenging scenarios. We present the results for linear and quadratic methods using the wavelet variances and correlations together. The figure illustrates the stable behavior of the proposed methods across the different wavelet filters. Also, linear procedures seems preferable to quadratic procedures. Kullback-Leibler information and Chernoff information procedures are outperformed by wavelet based procedures. As expected, Table 5 on page 13 and Figure5 on page 14 show that all methods improve when λ increases.

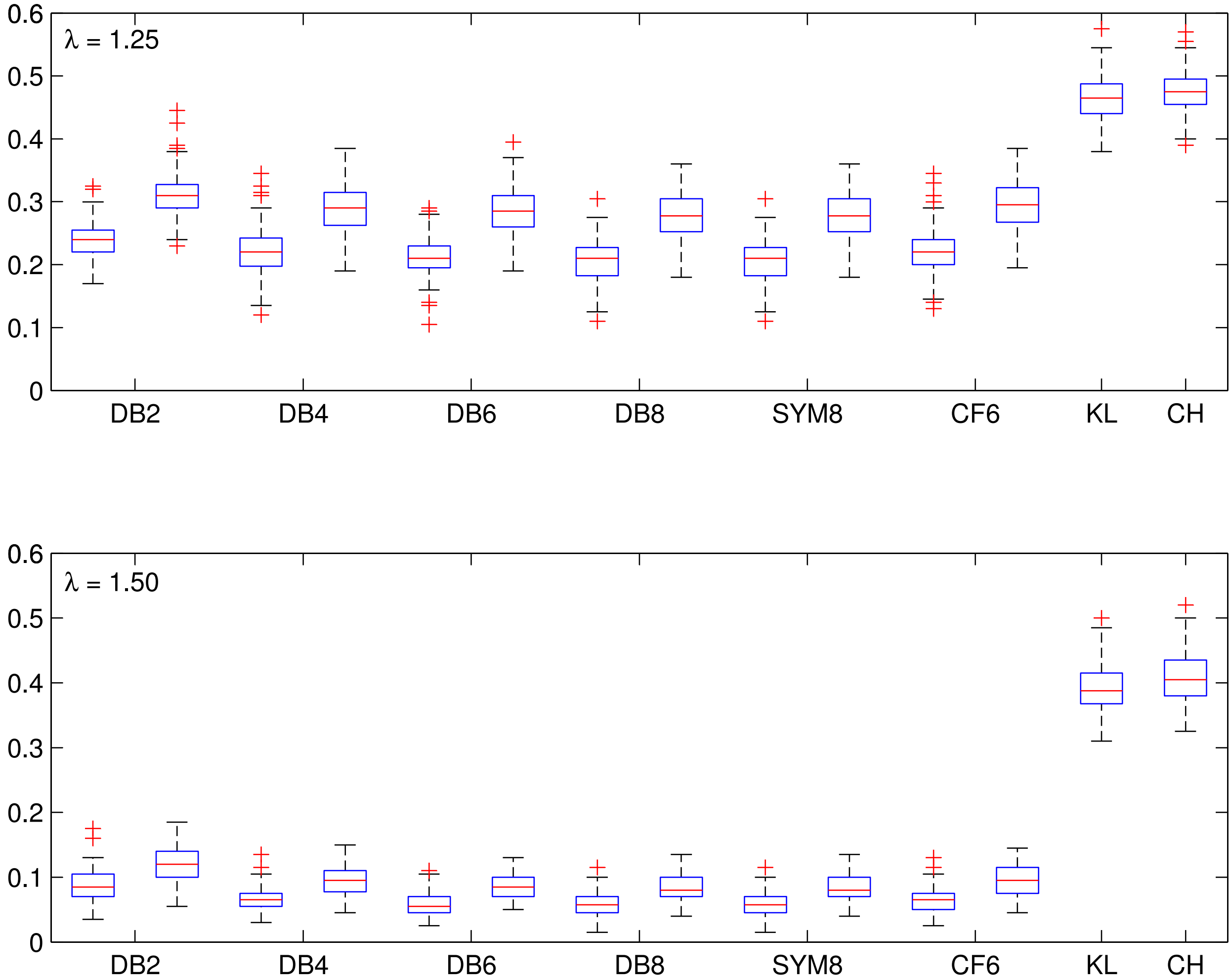


Fig 5. *Boxplots with the misclassification rates of the simulation with synthetic ECG signals.*

5. Other Simulation Studies

In order to evaluate the performance of the use of discriminant analysis with the wavelet features when applied to multivariate time series with patterns different from that of ECG signals, we conducted two sets of simulation studies in which we used the stepwise implementation of Strauss [16] in linear and quadratic discriminant procedures. For the first study, 50 pairs of bivariate time series of lengths $T = 2^9 = 512$ and $T = 2^{11} = 2048$ were generated from each of a vector autoregressive model of order one , VAR(1),

$$\mathbf{X}_t = \Phi \mathbf{X}_{t-1} + \epsilon_{t,1}, \tag{5.1}$$

where $\Phi = \begin{bmatrix} 0.5 & 0.1 \\ 0.7 & 0.5 \end{bmatrix}$ and from a vector autoregressive moving average model, VARMA(1,1),

$$\mathbf{X}_t = \Phi \mathbf{X}_{t-1} + \Theta \epsilon_{t-1} + \epsilon_{t,2}, \tag{5.2}$$

where $\Phi = \begin{bmatrix} 0.5 & 0.1 \\ 0.7 & 0.5 \end{bmatrix}$ and $\Theta = \begin{bmatrix} -0.2 & 0.1 \\ 0.3 & 0.1 \end{bmatrix}$.

The noises $\epsilon_{t,j}$ with $j = 1$ and 2 are bivariate $N(\mathbf{0}, \Sigma_j)$. We considered three noise structures in generating the bivariate time series, viz.,

1: $\Sigma_1 = \Sigma_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

2: $\Sigma_1 = \Sigma_2 \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$.

3: $\Sigma_1 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} 1 & 0.75 \\ 0.75 & 1 \end{bmatrix}$.

As in the previous two sections, wavelet filters from the Daubechies family (DB2, DB4, DB6, DB8), from the Symmletts family (SYM8) and from the Coiflets family (CF6) were used to generate the MODWT coefficients, and hence the MODWT variances and correlations. Refer to Table 1 for the maximum allowable number of scales for each of the filters for series of lengths $T = 2^9 = 512$ and $T = 2^{11} = 2048$.

We considered the three scenarios for input variables for each time series into the stepwise discriminant analysis procedure in order to select the optimal number of variables, which would produce the smallest overall error rate in the hold-out-one cross-validation procedure for discriminating between the two sets of bivariate time series. The total number of input variables for the first scenario were wavelet variances only; for the second, wavelet variances and scale by scale correlations; for the third scenario only the wavelet correlations. One-hundred simulations were carried out each time and an average overall misclassification rate was determined.

For the second simulation study, the same procedures as for the first simulation study were carried out, except that we introduced non-stationarity in the variance in the VAR and VARMA processes from which the time series were generated. This was done by multiplying each component of each bivariate time series by an exponential function, viz.,

$$\mathbf{Y}_t = D_t \times \mathbf{X}_t, \tag{5.3}$$

where

$$D_t = \exp \frac{-(t - 500^2)}{2 \times 200^2}. \tag{5.4}$$

Figure 6 on page 16 and Figure 7 on page 16 show single realizations of the bivariate time series generated with the third noise structure from each study. While, at a glance, there does not appear to be much difference between the overall patterns of the time series generated from the VAR(1) and VARMA(1,1) processes, differences over specific time period can be observed.

Table 6 on page 17 and Table 7 on page 18 show the average misclassification rates using linear and quadratic discriminant analyses for $T = 512$ and $T = 2048$ respectively, with the wavelet variances (var), wavelet variances and wavelet correlations (var-corr) and wavelet correlations (corr) and also using the Kullback-Leibler discrimination information and the Chernoff information measures.

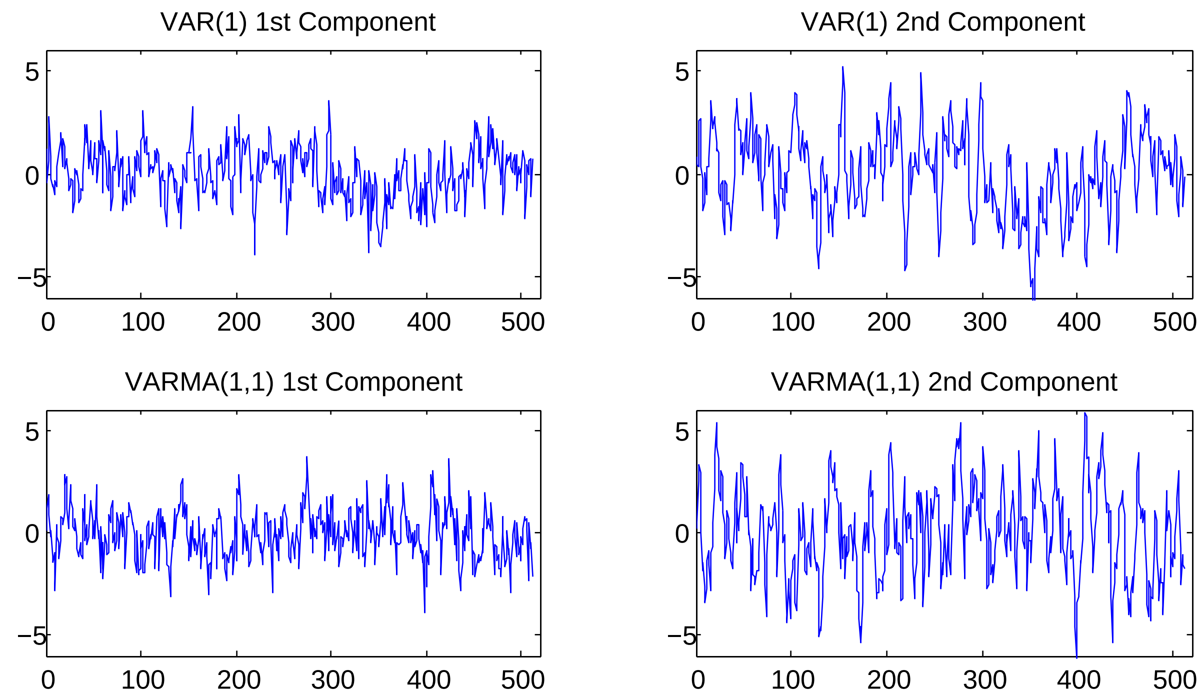


FIG 6. *Components of a mean and variance stationary bivariate series.*

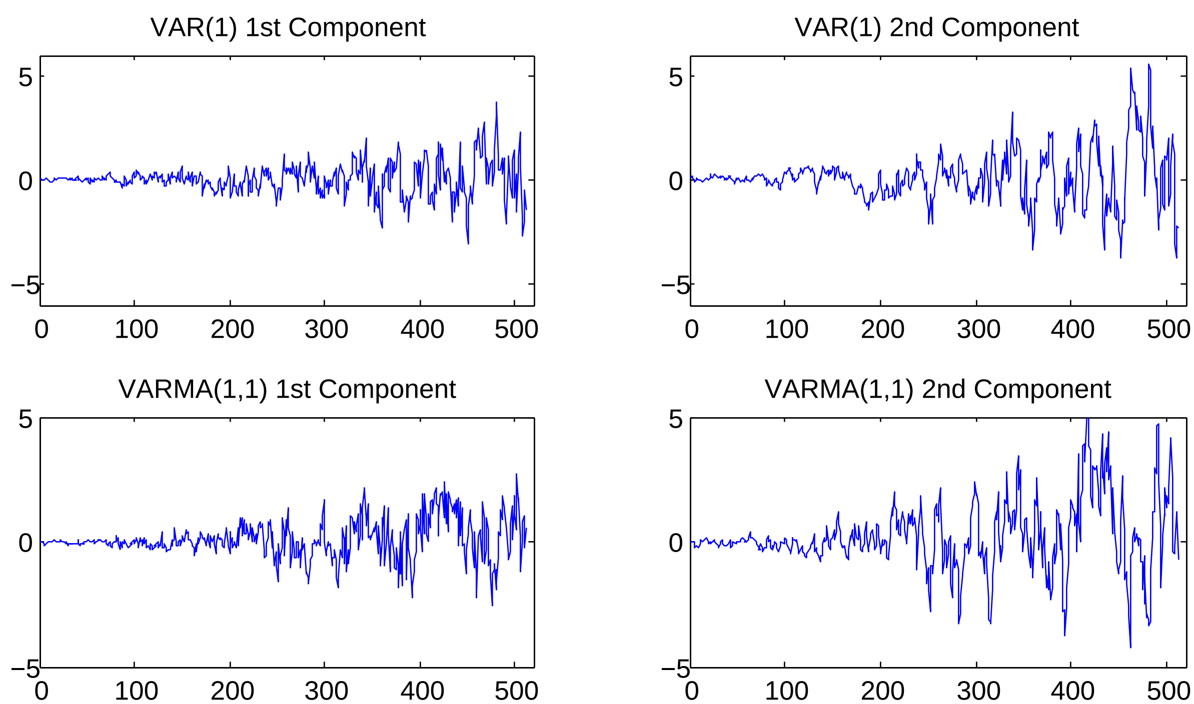


FIG 7. *Components of a variance non-stationary bivariate series..*

Table 6
Overall average misclassification rates for stepwise discriminant analysis and for the Kullback-Leibler and the Chernoff information measures applied to bivariate stationary time series: T=512 .

	Wavelet Filter	Feature	Stationary			Non-stationary		
			Noise 1	Noise 2	Noise 3	Noise 1	Noise 2	Noise 3
Linear	DB2	var	2.08	3.69	0.43	8.34	9.73	2.71
		varcor	**0.08**	**0.02**	**0.00**	**1.24**	**0.54**	**0.24**
		cor	0.58	0.10	0.15	3.81	1.35	1.49
	DB4	var	4.13	6.56	0.80	13.28	14.89	4.09
		varcor	**0.85**	**0.23**	**0.17**	**4.64**	**2.50**	**1.31**
		cor	0.96	0.34	0.58	5.81	2.14	2.16
	DB6	var	2.03	3.64	0.53	8.14	10.00	3.23
		varcor	**0.09**	**0.02**	**0.00**	**1.41**	**0.60**	**0.25**
		cor	0.58	0.07	0.09	3.70	1.34	1.03
	DB8	var	3.34	5.60	0.59	11.57	13.79	4.27
		varcor	**0.58**	**0.17**	**0.10**	**4.20**	**2.24**	**1.01**
		cor	0.93	0.23	0.27	5.07	1.93	1.38
	SYM8	var	1.94	3.67	0.52	7.70	10.03	3.24
		varcor	**0.10**	**0.03**	**0.00**	**1.25**	**0.56**	**0.23**
		cor	0.57	0.08	0.10	3.70	1.41	0.98
	CF6	var	2.60	4.93	0.47	10.32	12.95	3.70
		varcor	**0.27**	**0.10**	**0.02**	**3.04**	**1.64**	**0.51**
		cor	0.74	0.18	0.14	4.49	1.99	0.85
Quadratic	DB2	var	2.02	3.82	0.59	7.90	10.31	3.44
		varcor	**0.10**	**0.03**	**0.00**	**1.25**	**0.62**	**0.23**
		cor	0.59	0.09	0.08	3.71	1.44	0.91
	DB4	var	2.74	5.12	0.52	10.34	13.24	3.82
		varcor	**0.31**	**0.11**	**0.02**	**3.14**	**1.72**	**0.53**
		cor	0.75	0.20	0.14	4.43	2.03	0.85
	DB6	var	2.02	3.82	0.59	7.90	10.31	3.44
		varcor	**0.10**	**0.03**	**0.00**	**1.25**	**0.62**	**0.23**
		cor	0.59	0.09	0.08	3.71	1.44	0.91
	DB8	var	2.74	5.12	0.52	10.34	13.24	3.82
		varcor	**0.31**	**0.11**	**0.02**	**3.14**	**1.72**	**0.53**
		cor	0.75	0.20	0.14	4.43	2.03	0.85
	SYM8	var	1.87	3.42	0.45	7.52	9.84	2.94
		varcor	**0.09**	**0.03**	**0.00**	**1.23**	**0.52**	**0.22**
		cor	0.53	0.07	0.11	3.62	1.37	1.07
	CF6	var	2.52	4.66	0.44	10.06	12.36	3.43
		varcor	**0.24**	**0.08**	**0.03**	**2.91**	**1.53**	**0.52**
		cor	0.70	0.14	0.19	4.43	1.80	0.88
Kullback–Leibler			7.53	7.83	7.17	8.53	7.80	8.07
Chernoff			8.33	8.30	7.79	12.12	10.87	12.18

TABLE 7
Overall average misclassification rates for stepwise discriminant analysis and for the Kullback-Leibler and the Chernoff information measures applied to bivariate stationary time series: T=2048.

	Wavelet Filter	Feature	Stationary			Non-stationary		
			Noise 1	Noise 2	Noise 3	Noise 1	Noise 2	Noise 3
Linear	DB2	var	0.01	0.01	0.00	2.60	3.79	0.54
		varcor	0.00	0.00	0.00	**0.06**	**0.01**	**0.01**
		cor	0.00	0.00	0.00	0.51	0.09	0.17
	DB4	var	0.09	0.12	0.01	5.91	9.38	1.68
		varcor	0.01	0.00	0.00	**0.67**	**0.33**	**0.26**
		cor	0.00	0.01	0.02	1.76	0.80	1.19
	DB6	var	0.02	0.01	0.00	2.27	3.78	0.56
		varcor	0.00	0.00	0.00	**0.09**	**0.00**	**0.01**
		cor	0.00	0.00	0.00	0.59	0.09	0.14
	DB8	var	0.03	0.07	0.00	5.11	8.34	1.41
		varcor	0.00	0.00	0.00	**0.80**	**0.44**	**0.30**
		cor	0.00	0.00	0.01	1.53	0.59	0.78
	SYM8	var	0.02	0.02	0.00	2.02	4.11	0.55
		varcor	0.00	0.00	0.00	**0.12**	**0.02**	**0.03**
		cor	0.00	0.00	0.00	0.63	0.10	0.12
	CF6	var	0.04	0.06	0.00	4.05	7.07	0.97
		varcor	0.00	0.00	0.00	**0.97**	**0.36**	**0.23**
		cor	0.00	0.00	0.01	1.30	0.46	0.48
Quadratic	DB2	var	0.02	0.03	0.00	2.05	4.23	0.57
		varcor	0.00	0.00	0.00	**0.12**	**0.03**	**0.02**
		cor	0.00	0.00	0.00	0.63	0.11	0.13
	DB4	var	0.04	0.06	0.00	4.10	7.27	1.08
		varcor	0.00	0.00	0.00	**1.01**	**0.36**	**0.19**
		cor	0.00	0.00	0.01	1.35	0.46	0.51
	DB6	var	0.02	0.03	0.00	2.05	4.23	0.57
		varcor	0.00	0.00	0.00	**0.12**	**0.03**	**0.02**
		cor	0.00	0.00	0.00	0.63	0.11	0.13
	DB8	var	0.04	0.06	0.00	4.10	7.27	1.08
		varcor	0.00	0.00	0.00	**1.01**	**0.36**	**0.19**
		cor	0.00	0.00	0.01	1.35	0.46	0.51
	SYM8	var	0.02	0.01	0.00	2.03	3.86	0.52
		varcor	0.00	0.00	0.00	**0.12**	**0.01**	**0.02**
		cor	0.00	0.00	0.00	0.62	0.07	0.13
	CF6	var	0.03	0.05	0.00	4.04	6.74	0.88
		varcor	0.00	0.00	0.00	**0.83**	**0.30**	**0.19**
		cor	0.00	0.00	0.01	1.20	0.41	0.45
Kullback–Leibler			0.01	0.00	0.00	**0.40**	**0.14**	**0.15**
Chernoff			0.00	0.00	0.00	1.90	0.96	1.01

For $T = 512$ for both linear and quadratic methods, for both stationary and variance non-stationary series, and for all noise structures, across the different wavelet filters, the average misclassification rates were always smallest when both the wavelet variances and correlations were input together as the discrimination variables than when only wavelet variances or only wavelet correlations were the input variables. In particular, for this scenario, the misclassification rates ranged between 0% and 0.85% with the linear discriminant for stationary series and between 0.24% and 4.64% for variance non-stationary time series. With the quadratic discriminant the misclassification rates ranged between 0% and 0.31% for stationary series and between 0.22% and 3.14% for variance non-stationary series. The misclassification rates were generally lower for the third noise structure compared to the other noise structures for all combinations of linear and quadratic discriminators and stationary and variance non-stationary series. All wavelet based discriminant procedures clearly outperformed the Kullback-Leibler information and Chernoff information procedures.

For $T = 2048$ for stationary series with both linear and quadratic discriminators, across all noise structures, all wavelet-based methods as well as the Kullback-Leibler information and Chernoff information procedures. achieved 0% or close to 0% misclassification rates. On the other hand for non-stationary series amongst the wavelet-based methods, the misclassification rates were always smallest when both the wavelet variances and correlations were input together as the discrimination variables than when only wavelet variances or only wavelet correlations were the input variables. In particular, across the three noise structure and different wavelet filters, the misclassification rates ranged between 0.1% and 0.97% with the linear discriminator and between 0.02% and 0.83% with the quadratic discriminator. The Kullback-Leibler information procedure was quite competitive achieving misclassification rates of between 0.15% and 0.40% whereas the Chernoff information procedure achieved higher misclassification rates of between 0.96% and 1.90%

In summary for both studies, the best results were generally achieved when both wavelet variances and wavelet correlations were input together as the discriminating variables. As was also the case in the studies in Sections 3 4, when only wavelet variances were the input variables, the overall misclassification rates were generally higher. It is therefore clear that wavelet correlations provide useful information about the interrelationship between the components of each multivariate time series for the purposes of classification when combined with the wavelet correlations.

6. Concluding Remarks

We proposed a method of classifying multivariate ECG signals based on conventional discriminant analysis with wavelet features, namely, wavelet variances and wavelet correlations as the discriminating variables. We demonstrated in the application to ECG data, the study with synthetic ECG data, and in the other simulation studies, that when both wavelet variances and correlations were used as the discriminating variables, this method produces much better results that when only wavelet variances were used or in most cases when only wavelet correlations were used. We also demonstrated that overall, our procedure outperforms the other multivariate discriminant procedures proposed by Kakizawa et al. [8]. Using the results of other ECG data classifying methods as reference values, it is clear our method displays very favourable performance. Hence, the main contributions of this paper are i) the inclusion of the interrelationship features between components of the ECG signals in the form of wavelet correlations provides useful information for the purposes of classifying myocardial infarction; ii) the generally favourable performance of our approach when compared to the other well-known methods for classifying multivariate time series.

Acknowledgements

The first author began working on this paper while she was on sabbatical leave at the Universidad Carlos III de Madrid. She acknowledges the support provided by Monash University and the Universidad Carlos III de Madrid during this period.

We thank Professor David Stoffer for his guidance on the implementation of the routines for Kullback-Leibler and Chernoff-based procedures and Professors Gari Clifford and Reza Sameni for their guidance on the use of ECG synthetic generator.

All results for the application and simulations were generated using MATLAB software. To obtain the wavelet variances and correlations, we used the WMTSA Wavelet Toolkit developed by Cornish [22]. To

obtain ECG synthetic signals, we used the Open-Source Electrophysiological Toolbox developed by Sameni and co-authors [19], [21] and [20].

References

[1] Al-Naima, F. and Al-Timemy, A. (2009), Neural network based classification of myocardial infraction: A comparative study of wavelet and Fourier transforms in Pattern Recognition, Y. Peng-Yeng (Ed.). *InTech.pp.338-352.*

[2] De Chazel, P. and Reilly, R.B. (2000), A comparison of the ECG classification performance of different feature sets. *Computers in Cardiology* **27**, 327–330.

[3] Heden, B., Ohlin, H., Rittner, R. and Edenbrandt, L. (1997), Acute myocardial infarction detected in 12-lead ECG by artificial neural networks. *Circulation: American Heart Association Inc.* **96**, 1798–1802.

[4] Bozzola, G., Bortolan, G., Combi, C., Pinciroli, F. and Brohet, C. (1996), A hybrid neuro-fuzzy system for ECG classification of myocardial infraction. *Computers in Cardiology* , 241–244.

[5] Lachenbruch, P.A. and Mickey, M.R. (1968), Estimation of error rates in discriminant analysis.. *Technometrics* **10**, 1–10.

[6] Huang H., Hernando O., Stoffer D.S. 2004. Discrimination and Classification of Nonstationary Time Series using the SLEX Model . *J.Amer. Statist. Assoc.* **99 (467)**, 763–774.

[7] Maharaj, E.A., Alonso, A.M. (2007), Discrimination of locally stationary time series using wavelets. *Comput. Statist.Data Anal.* **52** , 879–889.

[8] Kakizawa Y., Shumway R.H. and Taniguchi M. (1998), Discrimination and clustering for multivariate time series . *J. Amer. Statist. Assoc* **93 (441)**, 324–340.

[9] Shumway, R.H. (2003), Time-frequency clustering and discriminant analysis . *Statist. Probab. Let.* **63 (3)**, 307–314.

[10] Percival, D.B. and Walden, A.T. (2000), Wavelet Methods for Time Series Analysis. *Cambridge University Press, Cambridge.*

[11] McLachlan, G.J. (2004), Discriminant Analysis and Statistical Pattern Recognition. *John Wiley and Sons, Inc. Hoboken, New Jersey.*

[12] Serroukh A., Walden A.T., Percival D.B. (2000), Statistical properties and uses of the wavelet variance estimator for the scale analysis of time series.*J. Amer. Statist. Assoc.* ***95 450*** *, 184–196.*

[13] Serroukh, A. and Walden, A.T. (2000), Wavelet scale analysis of bivariate time series II: Statistical properties for linear processes . *Nonparametric Statistics* ***13*** *, 37–56.*

[14] Chinganda, E.F. and Subrahaniam, K. (1979), Robustness of the linear discriminant function to nonnormality: Johnson's system. *J. Statist. Planning Inference* **3**, 69–77.

[15] Fatti, L.P., Hawkins, D.M., Raath, E.L. (1982), Discriminant Analysis. In Topics in Applied Multivariate Analysis, D.M. Hawkins (Ed.). *Cambridge University Press, Cambridge.pp.1-71.*

[16] Strauss, R.E. (2004), Statistical Matlab Library. *Available from: http://www.faculty.biol.ttu.edu/Strauss/Matlab/matlab.htm.*

[17] Shumway R.H. and Stoffer D.S. (2006), Time Series Analysis and Its Applications With R Examples. *Springer, New York.*

[18] McSharry P. E., Clifford G. D., Tarassenko L. and Smith L. A. (2003), A Dynamic Model for Generating Synthetic Electrocardiogram Signals. *IEEE Trans. Biomed. Eng* ***50****, 289–294.*

[19] Sameni R., Clifford G. D., Jutten C., Shamsollahi M. B. (2007), Multi-Channel ECG and Noise Modeling: Application to Maternal and Fetal ECG Signals. *IEURASIP Journal on Advances in Signal Processing. Article ID 43407, 14 pages, 2007. doi:10.1155/2007/43407.*

[20] Clifford G., Nemati, S. and Sameni, R. (2010), An Artificial Vector Model for Generating Abnormal Electrocardiographic Rhythms. *JPhysiological Measurements.* ***31 5****, 595–609.*

[21] Sameni R. (2008), Extraction of Fetal Cardiac Signals from an Array of Maternal Abdominal Recordings," Ph.D. dissertation, Sharif University of Technology - Institut National Polytechnique de Grenoble, July 2008. *[Online]. Available: http://www.sameni.info/Publications/Thesis/PhDThesis.pdf.*

[22] Cornish C. (2004), WMTSA wavelet toolkit for MATLAB. *available from: http://staff.washington.edu/dbp/wmtsa.html*