Robust multiple cardiac arrhythmia detection through bispectrum analysis

A. Lanatá *, G. Valenza, C. Mancuso, E.P. Scilingo

Department of Information Engineering and Interdepartmental Research Center "E. Piaggio", University of Pisa, Via Caruso 16, 56122 Pisa, Italy

ARTICLE INFO

ABSTRACT

This paper investigates the use of Higher Order Spectra parameters to identify the most common multiple cardiac arrhythmias. In detail, we calculated magnitude of bispectrum, three values of bispectrum entropy, mean and variance of the phase of bispectrum integrated over a particular region wherein no bispectrum aliasing is assumed. This set of features is used to distinguish normal QRS from five different classes of arrhythmia over a large amount of normal and pathologic ECG signals. An accurate parametric and non-parametric analysis of these feature distributions is also performed in order to identify the optimum classifier. Experimental tests were performed on signals gathered from the MIT-BIH Arrhythmias Database, and mean and standard deviation of all confusion matrixes obtained from 20 steps of cross validation are shown. Results showed that the bispectrum is high performance, reliable and robust method to identify arrhythmias.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The analysis of single waves of electrocardiogram (ECG) signal, such as QRS complex, P and T waves, is one of the essential tasks in the cardiovascular arrhythmia detection. Therefore, automatic arrhythmias recognition as medical decision support, is the most important research area in the ECG signal processing (de Chazal, Dwyer, Reilly, & Tabakov, 2004). Recently, a large amount of methods was developed to obtain efficient and robust features, extracted from QRS complexes, for automatic recognition by means of classifier algorithms. Generally, these algorithms can be based on three main analysis techniques: in the time, frequency and time–frequency domain. Regarding the time analysis, which is known as morphology based technique, the standard features are heartbeat interval, amplitude parameters (QRS, ST), duration parameters (QRS, QT, and PR), and combined parameters (Q/R ratio, S/R ratio). More recently, two QRS descriptors have been defined to recognize pathological events in real-time mode (Iliev & Krasteva, 2007), two indexes from QRS slope analysis have been introduced (Pueyo, Sörnmo, & Laguna, 2008) in order to quantifying ischemia-induced ECG changes in percutaneous transluminal coronary angioplasty (PTCA), whilst either advanced morphologic parameters from Karhunen–Loéve transform with an adaptive neurofuzzy logic classifier (Pang et al., 2005) or ECG parameters from the Dilated Discrete Hermite expansion (Gopalakrishnan, Acharya, & Mugle, 2004) were used in heart ischemia detection.

Regarding features extracted from QRS complex in the frequency domain, methods based on a Fourier transform have been developed in order to discriminate different kinds of rhythms by means of several classifiers where the most common are Artificial Neural Networks (ANNs) (Minami, Nakajima, & Toyoshima, 1999). Latest works used a Fourier Power Spectrum (FPS) of QRS complex, observed in the range of 0–20 Hz, to classify the cardiac arrhythmias with Grey Relational Analysis (GRA) (Lin, 2008).

Regarding the time–frequency techniques, Wavelet Transforms (WT) have been applied to extract features of cardiac arrhythmias (Addison, 1999). In particular, spatial transformation of QRS complex by using Morlet Wavelet has been proposed to obtain features for an ANN (Lin, Du, & Chen, 2008). Wigner–Ville distribution has been used to characterize atrial fibrillation in the ECG and to track the variations in fundamental frequency of the fibrillatory waves (Stridh, Sörnmo, Meurling, & Olsson, 2001).

From current literature, several features are obtained from Higher Order Spectra (HOS) of each QRS complex.

This choice is motivated by the following considerations (Mendel, 1991):

- Shape information of a signal resides primarily in the phase and not in the amplitude of its Fourier transform.
- HOS retain both amplitude and phase information from the Fourier transform of the signal.
- HOS are translation invariant. Functions can be defined from HOS, which satisfy desirable properties, such as scaling and amplification invariance. These functions can be then utilized as invariant features for pattern recognition.

* Corresponding author. Tel.: +39 050 2217462; fax: +39 050 2217051.
 E-mail addresses: antonio.lanata@ing.unipi.it, a.lanata@centropiaggio.unipi.it (A. Lanatá).

- HOS, of order greater than two, have zero expected value for Gaussian noise. Therefore, the obtained features have high immunity to additive Gaussian noise.
- Even if the signal is subjected to a time shift, the phase of bispectrum remains unaltered.

This work is focused on the following features: invariants of the bispectrum integrated in the triangular region showed in Fig. 3 (Chandran & Elgar, 1993), mean magnitude and phase of Shannon entropy (Chang & Sun, 2004; Grassberger, Schrieber, & Schraffrach, 1991) integrated in the same triangular region (Chua, Chandran, Acharya, & Lim, 2008).

Six classes of the most important arrhythmic QRS complexes are used as target of the classifiers: (i) Normal Beat (NB), (ii) Paced Beat (PB), (iii) Left Bundle Branch block Beat (LBB), (iv) Right Bundle Branch block Beat (RBB), (v) Premature Ventricular Contraction Beat (PVCB), (vi) Atrial Premature Contraction-1 Beat (APCB) (see Fig. 1).

Usually, the arrhythmias recognition tasks are performed by means of ANNs, it is due to their ability to separate feature spaces into non-linear regions (Kinnebrock, 1992). However, this paper aims at demonstrating a strong robustness of the extracted features from HOS, which do not require further space transformation. Accordingly, the optimum classifier is carried out from a study of the feature distributions, between non-parametric and parametric methods (Duda, Hart, & Stork, 2001; Friedman, Hastie, & Tibshirani, 2000; Heijden, Duin, Ridder, & Tax, 2004; Jain, Duin, & Mao, 2000; Schlesinger & Hlavac, 2002; Vapnik, 1998; Webb, 2002). A deep discussion on the distributions is also reported into Conclusion and Discussion section. Experimental tests were performed on signals gathered from the MIT-BIH Arrhythmias Database and in order to evaluate performance of each classifier, we calculated mean and standard deviation of all confusion matrixes obtained from 20 steps of cross validation (Kohavi & Provost, 1998).
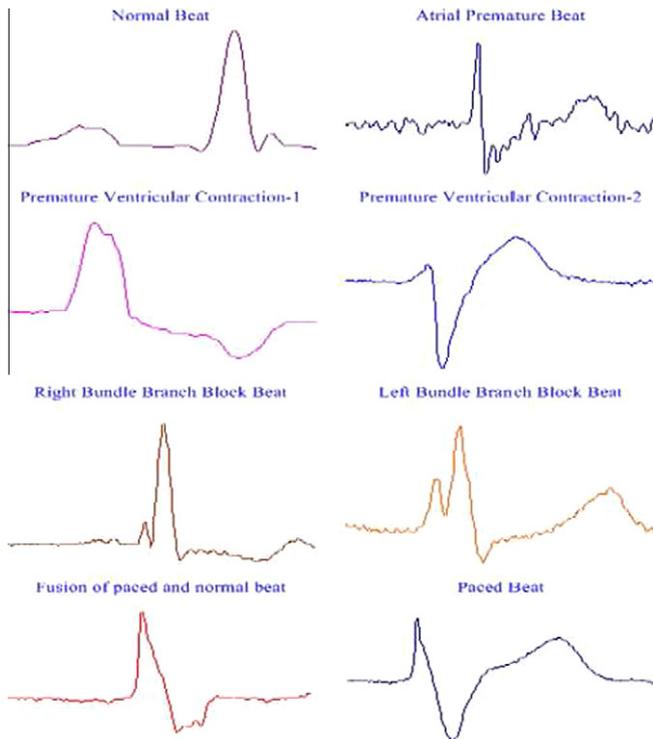
2. Materials and methods

'High Order Spectra' (HOS) or polyspectra of a stochastic process are defined from moments or cumulants of order greater than two. In particular, this work is focused on the two dimensional third order cumulant Fourier transform, called bispectrum (Mendel, 1991; Nikias, 1993):

$$B(f_1, f_2) = \int\int_{t_1, t_2 = -\infty}^{+\infty} c_3(t_1, t_2) exp^{-j(2\pi f_1 t_1 + 2\pi f_2 t_2)} df_1 \, df_2 \qquad (1)$$

with the condition:

$$|\omega_1|, |\omega_2| \leqslant \pi \quad \text{for } \omega = 2\pi f$$

The $c_3(t_1, t_2)$ variable represents the third order cumulant, which is defined as follows:

$$c_3(t_1, t_2) = E\{s(t_1)s(t_2)s(t_1 + t_2)\} \qquad (2)$$

where $s(t)$ is a square integrable stationary signal with zero mean. Thus, the bispectrum measures the correlation among three spectral peaks, ω_1, ω_2 and $(\omega_1 + \omega_2)$ and estimates the phase coupling.

Sometime bispectrum is unable to distinguish between pairs of frequency strongly coupled and pairs of frequency weakly coupled but at high frequency, because their bispectrum values are similar. In order to overcome this limitation it is possible to evaluate power and calculate the bicoherence function. The bicoherence function is the bispectrum normalized form with respect to its power spectrum:

$$B_{co}(f_1, f_2) = \frac{B(f_1, f_2)}{\sqrt{P(f_1)P(f_2)P(f_1 + f_2)}} \qquad (3)$$

where $P(f)$ is the estimated power spectrum of the $s(t)$ signal.

2.1. Bispectrum features

A previous study demonstrated that the bispectrum have the following symmetry properties (Nikias & Raghuveer, 1987):

$$B(f_1, f_2) = B(f_2, f_1) \qquad (4)$$
$$B(f_1, f_2) = B * (-f_2, -f_1) \qquad (5)$$
$$B(f_1, f_2) = B * (-f_1, -f_2) \qquad (6)$$
$$B(f_1, f_2) = B(-f_1 - f_2, f_2) \qquad (7)$$
$$B(f_1, f_2) = B(f_1, -f_1 - f_2) \qquad (8)$$
$$B(f_1, f_2) = B(-f_1 - f_2, f_1) \qquad (9)$$
$$B(f_1, f_2) = B(f_2, -f_1 - f_2) \qquad (10)$$

that divides the (f_1, f_2) plane in eight symmetric zones (see Fig. 2).

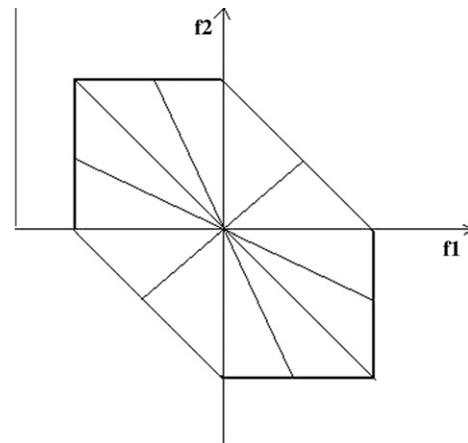**Fig. 1.** The most common cardiac arrhythmias from Lin et al. (2008).

Fig. 2. The eight symmetric zones of bispectrum.

Bispectrum of a real signal is uniquely defined by its values in the triangular region of computation, $0 \leqslant f_1 \leqslant f_2 \leqslant f_1 + f_2 \leqslant 1$, provided there is no bispectral aliasing (Brillinger, Rosenblatt, & Petropulu, 1967).

The bispectral feature set consists of: mean and variance of bispectral invariants (mean and variance of $P(a)$), mean magnitude (M_{mean}) of the bispectrum and the phase entropy (P_e), normalized bispectral entropy (P_1) and normalized bispectral squared entropy (P_2).

Specifically, let us introduce the bispectral parameter, $P(a)$, which is invariant to translation, dc-level, amplification, and scale. It is defined as follows:

$$P(a) = arctan\left(\frac{I_i(a)}{I_r(a)}\right) \tag{11}$$

where:

$$I(a) = I_r(a) + jI_i(a) = \int_{f_1=0^+}^{\frac{1}{1+a}} B(f_1, af_1) df_1 \tag{12}$$

for $0 < a \leqslant 1$ and $j = \sqrt{-1}$ where a is the slope of the straight line on which bispectrum is integrated.

In this work, only the mean and variance of this features are considered.

Mean magnitude and phase entropy (Chang & Sun, 2004) are calculated within the region defined in Fig. 3.

Mean magnitude is defined as:

$$M_{mean} = \frac{1}{L} \sum_{\Omega} |B(f_1, af_1)| \tag{13}$$

and phase entropy is:

$$P_{e=} \sum_n p(\Psi_n) log(p(\Psi_n)) \tag{14}$$

$$p(\Psi_n) = \frac{1}{L} \sum_{\Omega} 1(\Phi(B(f_1, af_1)) \epsilon \Psi_n) \tag{15}$$

$$\Psi_n = \{\Phi| - \pi + 2\pi n/N \leqslant \phi \leqslant -\pi + 2\pi(n+1)/N\} \tag{16}$$

with $n = 0, 1, \ldots, N - 1$.

L is the number of points within the region in Fig. 3, Φ refers to the phase angle of the bispectrum, Ω refers to the space of the defined region in Fig. 3, and $1(\cdot)$ is an indicator function which is equal to 1 when the phase angle Φ is within the range of bin Ψ_n in Eq. (16).

The mean magnitude of the bispectrum can be useful in discriminating between processes with similar power spectra but different third order statistics. However, it is sensitive to amplitude changes.

Normalized bispectral entropy (P_1) is equal to:

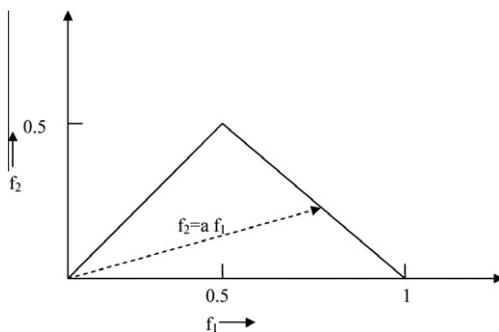$$P_1 = -\sum_n p_n log(p_n) \tag{17}$$

where:

$$p_n = \frac{|B(f_1, af_1)|}{\sum_{\Omega} |B(f_1, af_1)|} \tag{18}$$

and Ω is the region as in Fig. 3.

Normalized bispectral squared entropy (P_2) is calculated as:

$$P_2 = -\sum_n p_n log(p_n) \tag{19}$$

where:

$$p_n = \frac{|B(f_1, af_1)|^2}{\sum_{\Omega} |B(f_1, af_1)|^2} \tag{20}$$

and Ω is the region as in Fig. 3.

For each QRS extracted by the MIT-BIH arrhythmia database records, the bispectrum was computed using an indirect method as Fast Fourier Transform (FFT) of a third order cumulant estimation. All features are extracted as above described.

For each one of the six classes, the features distribution was studied by means of Royston test (Royston, 1983; Royston, 1992) which is an extension of the Shapiro–Wilk test. The Royston method combines each single univariate normality analysis into one omnibus statistic test for multivariate normality (see Table 6).

2.2. Classifier

In this work two classifiers are used. The first, which is named Mixture of Gaussian (MOG), represents a classifier based on parametric methods. The second, which is named is K-Nearest Neighbor (K-NN) classifier, is one of the most popular non-parametric classification tools. These classifiers are explained as follows:

2.2.1. MOG

Mixture of Gaussian algorithm uses a model-based approach to cluster the input data Fig. 4. Each class is represented by a known probabilistic distribution (Gaussian distribution) centered in the centroid of the cluster (see Fig. 4).

All input data are modeled by a combination of these distributions.

If mean and variance of all Gaussian distributions are known, MOG algorithm calculates a distance function as the probability that a new data refer to a specific class.

$$d(x_i, c_j) = P(x_i|G_j) = P(x_i|m_j, \sigma_j)$$

Otherwise, these parameters should be estimated by means of an iterative algorithm, which is commonly named expectation maximization (EM) (Xu & Jordan, 1992).
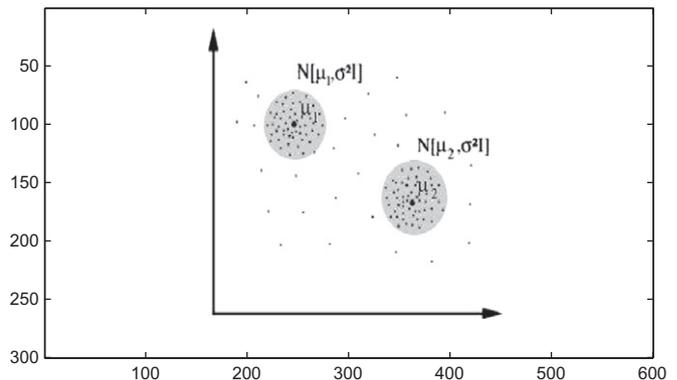

Fig. 3. Bispectrum invariants from Chua et al. (2008).

Fig. 4. Mixture of Gaussian.

2.2.2. K-NN–VDC

Non parametric *K*-Nearest Neighbor (*K*-NN) (Cover & Hart, 1967) algorithm uses the "proximity" concept between two examples belonging to the same class.

Our implementation was performed according to the following steps:

1. In the training phase, the *K*-NN just stores the training set data together with the labeling information.
2. In the test phase the *K*-NN calculates the Euclidean distance between the new test example and all the examples of the training set as follows:

$$D = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2}$$

$$= \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \qquad (21)$$

where $P = (p_1, p_2, \ldots, p_n)$ is a training set example, $Q = (q_1, q_2, \ldots, q_n)$ is the new test set example, n is the number of features.
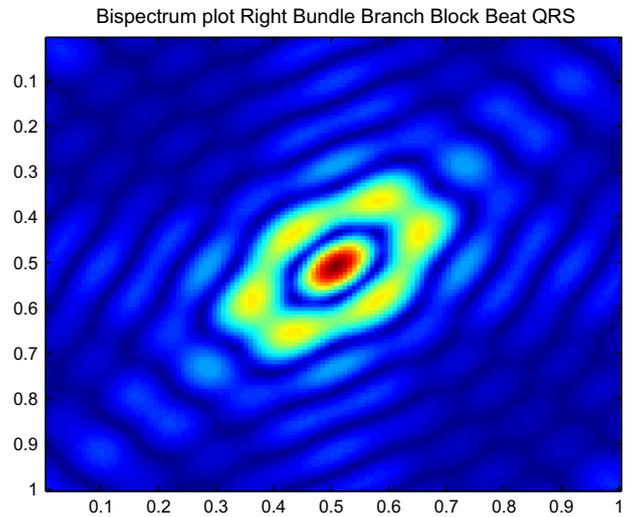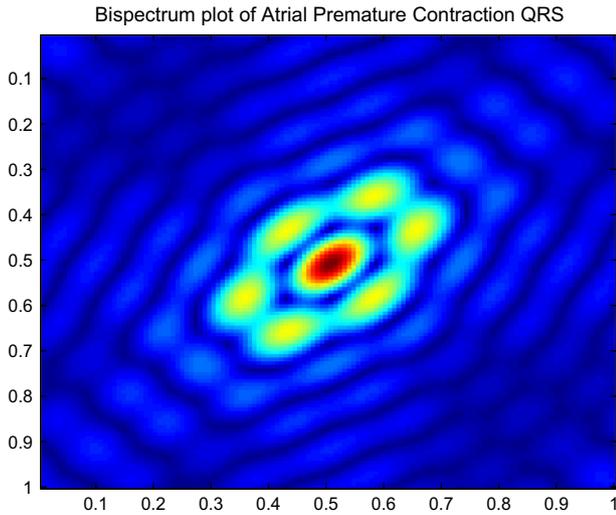
Bispectrum plot Right Bundle Branch Block Beat QRS

Fig. 7. Bispectrum class 3 Right Bundle Branch Block Beat.

Bispectrum plot of Atrial Premature Contraction QRS

Fig. 5. Bispectrum class 1 Atrium Premature Contraction.

Bispectrum plot of Left Bundle Branch Block Beat QRS
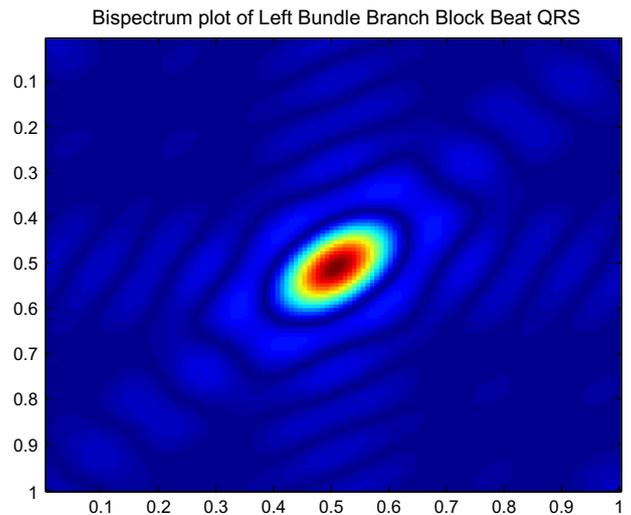
Fig. 8. Bispectrum class 4 Left Bundle Branch Block Beat.

Bispectrum plot of Premature Ventricular Contraction QRS
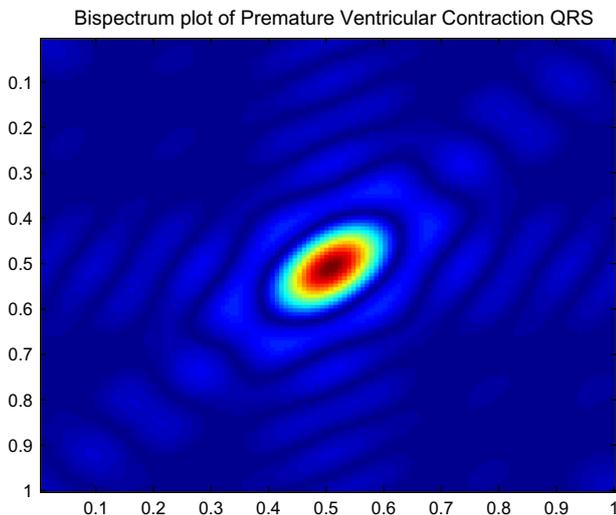
Fig. 6. Bispectrum class 2 Ventricular Premature Contraction.
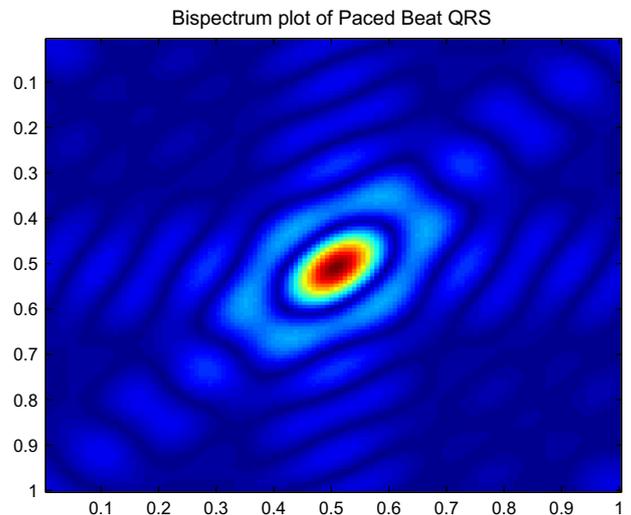
Bispectrum plot of Paced Beat QRS

Fig. 9. Bispectrum class 5 Paced Beat.

Afterwards, the K-NN, found the K examples that have the minimum distance, simply associates the more frequent class to the example test.

In order to verify the robustness of pattern recognition of the QRS complexes, the result of K-NN was compared with a simple Vector Distance classifier (VDC), which is a K-NN classifier with $K = 1$ (see Table VDC vs K-NN).

3. Experimental results

QRSs are extracted by the MIT-BIH arrhythmia database records, labelled as 107, 109, 111, 118, 119, 124, 200, 209, 212, 214, 217, 221, 231, 232 and 233. The feature set is showed in Table 5, where each row indicates the specific class and the columns reporting the kind of cardiac arrhythmia, the number of QRSs extracted, mean and standard deviation of all features for each six classes (from Fc1 to Fc6), respectively.

An estimation of each bispectrum per class is reported (see Figs. 5–10).
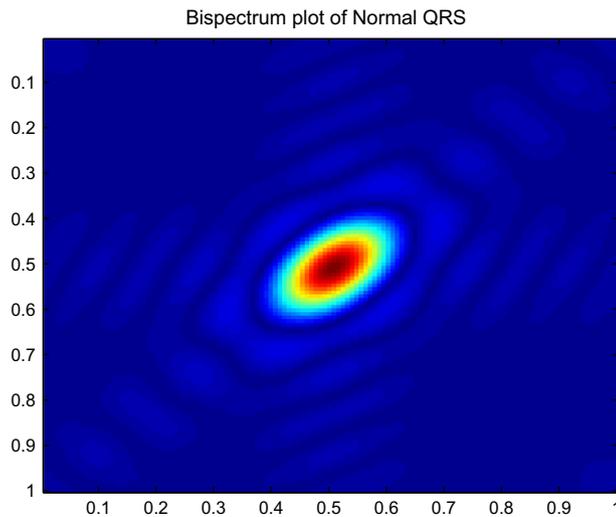

Fig. 10. Bispectrum class 6 Normal Beat.

In Tables 1–3 the confusion matrix of the used classifiers is showed; for each row and column the percentage of class recognition is reported in these tables the main diagonal represents the percentage of the recognition of that particular class (in bold), while the remaining data represent the errors of the recognition.

In Table 6, the study of the feature distributions is reported. In this table the rows refer to the classes, the columns refer to statistics Royston parameter, equivalent degrees of freedom, p-value, test-significance and normal distribution response.

4. Conclusion and discussion

In this paper, a method able to extract a robust set of features from the bispectrum was implemented, and the optimum classifier was determined and validated through six classes of normal and pathological QRSs. The bispectrum was calculated as Fourier transform of an estimation of the third order cumulant.

In order to choose the best classifier, a study of the feature distributions was carried out by means of Royston test. The results obtained from this test demonstrated that the feature distributions are Gaussian for all the considered classes (see Table 6). Accordingly, the MOG solution should theoretically be the most suitable classifier. Nevertheless, as shown in the same table, all p-values associated to each class, resulted under 0.6. It means that all Gaussian distributions have a high variance, therefore distribution tails are too overlapped (see confusion matrix Table 1) providing high numbers of positive and negative false, thus it did not result the most appropriate classifier. Instead, the use of non-parametric classifiers such as K-NN and VDC resulted the best choice, and in addition they did not required any statistical value from the distributions (see Tables 2 and 3). It means that a classifier which associates a new sample with a class, based on the occurrence of samples of that class into a region around the new sample, showed a better classification.

Regarding feature robustness, this paper showed very small differences between VDC (K-NN with $K = 1$) and classic K-NN (see Table 4) in terms of sensitivity. It means that the features were satisfactorily located that either the use the smallest K value ($K = 1$) or the optimum K value does not add significant differences. In addition VDC does not distort the feature space of inputs with additional transformations demonstrating the excellent feature placement. Worst outcome is from Atrium Premature

Table 1
Confusion matrix of MOG classifier.

Mean ± std of MOG					
69.8417 ± 6.9049	0.7480 ± 0.3745	7.5280 ± 2.2377	0.9704 ± 0.5285	0.4369 ± 0.2747	11.4181 ± 6.0215
1.2137 ± 0.6017	**65.8211 ± 17.3218**	0.7133 ± 0.3627	1.4632 ± 1.2770	5.9778 ± 2.1344	0.1293 ± 0.1419
6.1082 ± 1.2531	3.3496 ± 0.5837	**56.7483 ± 4.5267**	1.9522 ± 0.6563	3.4258 ± 0.8692	2.7651 ± 0.6002
2.1636 ± 0.7244	21.2846 ± 12.3493	20.6713 ± 1.3181	**77.8772 ± 3.8705**	22.0527 ± 2.7164	5.1207 ± 3.1554
0.7256 ± 0.4520	6.3252 ± 5.1367	0.7867 ± 0.2290	0.6141 ± 0.2602	**65.2705 ± 2.8818**	0.1616 ± 0.1161
19.9472 ± 6.5779	2.4715 ± 0.8900	13.5524 ± 5.8281	17.1228 ± 3.4720	2.8363 ± 2.0367	**80.4052 ± 7.3764**

Table 2
Confusion matrix of K-NN classifier.

Mean ± std of K-NN					
71.2137 ± 2.6797	0.7398 ± 0.3518	4.0594 ± 0.6186	0.1289 ± 0.0700	0.1318 ± 0.1310	0.8556 ± 0.1668
0.3958 ± 0.3370	**86.3496 ± 1.1333**	0.4650 ± 0.2132	1.5087 ± 0.3012	2.8433 ± 0.6029	0.1703 ± 0.0867
15.2902 ± 2.1116	3.6016 ± 0.4727	**87.5455 ± 0.8918**	6.9105 ± 0.7787	1.4910 ± 0.6612	2.0647 ± 0.2996
0.6069 ± 0.4618	4.0976 ± 0.5721	5.0140 ± 0.4813	**82.3768 ± 1.2211**	0.8460 ± 0.4097	2.5647 ± 0.3433
0.1187 ± 0.1596	2.4959 ± 0.7181	0.1364 ± 0.0698	0.4397 ± 0.1292	**94.4452 ± 1.3997**	0.0086 ± 0.0177
12.3747 ± 1.3369	2.7154 ± 0.4480	2.7797 ± 0.5042	8.6353 ± 0.8665	0.2427 ± 0.1736	**94.3362 ± 0.5241**

Table 3
Confusion matrix of VDC classifier.

Mean ± std of VDC					
66.2797 ± 2.1497	0.8699 ± 0.3089	4.5000 ± 0.4270	0.1668 ± 0.1002	0.1040 ± 0.0994	2.1099 ± 0.2626
0.7652 ± 0.3622	**86.4634 ± 1.3156**	0.8776 ± 0.2182	1.5049 ± 0.3303	1.9626 ± 0.4082	0.4267 ± 0.1135
19.5251 ± 1.4502	3.1545 ± 0.5955	**85.0420 ± 0.8167**	5.9666 ± 0.6966	0.5409 ± 0.2982	2.1078 ± 0.2548
0.4485 ± 0.2582	4.4472 ± 0.5881	6.1399 ± 0.6408	**84.0447 ± 0.8089**	0.2774 ± 0.1855	3.7392 ± 0.4141
0.0264 ± 0.0812	2.1301 ± 0.4014	0.1538 ± 0.0837	0.2957 ± 0.1177	**96.7753 ± 0.6688**	0.0129 ± 0.0246
12.9551 ± 1.6375	2.9350 ± 0.7562	3.2867 ± 0.4515	8.0212 ± 0.6497	0.3398 ± 0.2085	**91.6034 ± 0.5484**

Table 4
K-NN–VDC.

K-NN–VDC					
4.4987 ± 3.6595	0.1463 ± 0.4716	−0.8497 ± 0.8272	−0.1175 ± 0.1832	0.1526 ± 0.3015	−0.9784 ± 0.4203
−0.2902 ± 0.4440	−0.3821 ± 1.4989	−0.1888 ± 0.3201	−0.0152 ± 0.4142	0.2427 ± 0.7027	−0.1207 ± 0.1926
−5.7388 ± 2.4985	0.8537 ± 1.1090	2.4755 ± 1.0804	1.1259 ± 1.0175	0.3814 ± 0.3241	0.2845 ± 0.3685
0.2902 ± 0.2406	−0.2602 ± 1.0262	−0.8601 ± 1.1328	−1.8461 ± 1.1462	−0.0693 ± 0.3198	−0.6293 ± 0.5020
−0.0528 ± 0.2652	−0.1870 ± 0.5881	−0.0105 ± 0.0762	−0.0569 ± 0.1494	−0.9293 ± 0.8076	0.0129 ± 0.0578
1.2929 ± 2.7619	−0.1707 ± 0.7939	−0.5664 ± 0.5515	0.9098 ± 1.0136	0.2219 ± 0.2895	1.4310 ± 0.7337

Table 5
Features set.

Cardiac arrhythmia	QRSs	Fc1	Fc2	Fc3	Fc4	Fc5	Fc6
APCB	1895	−0.3132 ± 0.3415	0.6313 ± 0.3113	0.4903 ± 0.5975	−36,4895 ± 0.1729	5.5602 ± 0.5416	3.7145 ± 0.6230
PVCB	3071	−0.3530 ± 0.1726	0.7856 ± 0.1190	1.5760 ± 1.4491	−35,9997 ± 0.1626	5.7608 ± 0.1204	3.4871 ± 0.2303
RBB	7147	−0.2619 ± 0.2368	0.8336 ± 0.0979	1.0410 ± 1.1632	−36,5735 ± 0.3315	5.7957 ± 0.2693	3.7044 ± 0.4704
LBB	6593	−0.2959 ± 0.1725	0.8656 ± 0.1035	0.9543 ± 1.0815	−36,1659 ± 0.1269	5.6245 ± 0.1282	3.3827 ± 0.1119
PB	3604	0.0215 ± 0.1846	0.9076 ± 0.0778	0.4376 ± 0.3157	−35,9844 ± 0.1633	5.4687 ± 0.1485	3.3332 ± 0.0635
NB	11596	0.0887 ± 0.2138	0.7944 ± 0.1547	0.0808 ± 0.1123	−36,2950 ± 0.2234	5.5334 ± 0.2177	3.4793 ± 0.3528

Table 6
Royston's test.

Royston's test	Royston's statistic	Equivalent d.o.f.	p-Value	Significance	Normal distribution
Class 1	1.698433	1.053070	0.204650	0.050	Yes
2	0.804412	1.611792	0.567010	0.050	Yes
3	1.615098	1.280200	0.271557	0.050	Yes
4	0.743726	1.284151	0.487213	0.050	Yes
5	1.246641	1.462444	0.395291	0.050	Yes
6	1.497941	1.066260	0.237792	0.050	Yes

Contraction because there was not a sufficient training set of examples.

The outstanding result obtained opens a new scenario of low level hardware implementation (such as micro controllers) for portable electronics in smart home care system.

References

Addison, P. (1999). Wavelet transforms and the ecg: A review. *Physiological Measurement, 26.*

Brillinger, D. R., Rosenblatt, M., & Petropulu, P. (1967). *Computation and interpretation of kth order spectra. Spectral analysis of time series.* New York: Wiley. pp. 189–232.

Chandran, V., & Elgar, S. (1993). Pattern recognition using invariants defined from higher order spectra-one-dimensional inputs. *IEEE Transactions on Signal Processing, 41*(1).

Chang, T. N., & Sun, S. (2004). Blind detection of photomontage using higher order statistics. In *IEEE international symposium on circuits and systems (ISCAS), Vancouver, Canada.*

Chua, K. C., Chandran, V., Acharya, U. R., & Lim, C. M. (2008). Cardiac state diagnosis using higher order spectra of heart rate variability. *Journal of Medical Engineering and Technology, 32*(2), 145–155.

Cover, T., & hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory, 13*(1), 21–27.

de Chazal, P., Dwyer, M. O., Reilly, R. B., & Tabakov, S. (2004). Automatic classification of heartbeats using ecg morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering, 51*(7), 1196–1206.

Duda, R., Hart, P., & Stork, D. (2001). *Pattern classification* (2nd ed.). John Wiley and Sons.

Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression: A statistical view of boosting. *The Annals of Statistics, 38*(2), 337–374.

Gopalakrishnan, R., Acharya, S., & Mugle, D. (2004). Automated diagnosis of ischemic heart disease using dilated discrete hermite functions. In *Proceedings of the IEEE 30th annual northeast bioengineering conference.*

Grassberger, P., Schrieber, T., & Schraffrach, C. (1991). Nonlinear time sequence analysis. *International Journal of Bifurcation and Chaos, 1,* 512–547.

Heijden, F., Duin, R., Ridder, D., & Tax, D. (2004). *Classification, parameter estimation and state estimation.* John Wiley.

Iliev, I., & Krasteva, V. (2007). Real-time detection of pathological cardiac events the electrocardiogram. *Physiological Measurement.*

Jain, A., Duin, R., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(1), 4–37.

Kinnebrock, W. (1992). *Neural networks.* Munchen: Oldenburg Verlag.

Kohavi, R., & Provost, F. (1998). Glossary of terms. *Machine Learning, 30,* 271–274.

Lin, C. (2008). Frequency domain features for ecg beat discrimination using grey relational analysis based classifier. *Computers and Mathematics with Applications, 55.*

Lin, C., Du, Y., & Chen, T. (2008). Adaptive wavelet network for multiple cardiac arrhythmias recognition. *Expert Systems with Applications, 34.*

Mendel, J. (1991). Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications. *Proceedings of the IEEE, 79,* 278–305.

Minami, K., Nakajima, H., & Toyoshima, T. (1999). Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network. *IEEE Transactions on Biomedical Engineering, 46*(2).

Nikias, C. (1993). *Higher-order spectral analysis: A nonlinear signal processing framework.* USA: PTR Prentice-Hall, Inc..

Nikias, C., & Raghuveer, M. (1987). Bispectrum estimation: A digital signal processing framework. *Proceedings of the IEEE, 75*(7), 869–891.

Pang, L., Tchoudovski, I., Braecklein, M., Egorouchkina, K., Kellermann, W., & Bolz, A. (2005). Real time heart ischemia detection in the smart home care system. In *Proceedings of the 2005 IEEE engineering in medicine and biology 27th annual conference, Shanghai, China.*

Pueyo, E., Sörnmo, L., & Laguna, P. (2008). Qrs slopes for detection and characterization of myocardial ischemia. *IEEE Transactions on Biomedical Engineering, 55*(2).

Royston, J. P. (1983). Some techniques for assessing multivariate normality based on the Shapiro–Wilk w. *Applied Statistics, 32*(2).

Royston, J. (1992). Approximating the Shapiro–Wilk w-test for non-normality. *Statistics and Computing, 2*.

Schlesinger, M., & Hlavac, V. (2002). *Ten lectures on statistical and structural pattern recognition*. Kluwer Academic Publishers.

Stridh, M., Sörnmo, L., Meurling, C., & Olsson, S. (2001). Characterization of atrial fibrillation using the surface ecg: Time-dependent spectral properties. *IEEE Transactions on Biomedical Engineering, 48*(1).

Vapnik, V. (1998). *Statistical learning theory*. John Wiley.

Webb, A. (2002). *Statistical pattern recognition*. John Wiley and Sons.

Xu, L., & Jordan, M. (1992). On convergence properties of the em algorithm for Gaussian mixtures. *Neural Computation, 8*(1), 129–151.