An ECG Classification Model based on Multilead Wavelet Transform Features

M Llamedo Soria[1], JP Martínez[2]

[1]Universidad Tecnológica Nacional, Buenos Aires, Argentina
[2]Comm Tech Group, Aragon Inst of Eng Research / CIBER-BBN, Univ of Zaragoza, Aragon, Spain

Abstract

The objective of this work is to develop a model for ECG classification based on multilead features. The MIT-BIH Arrhythmia database was used following AAMI recommendations and class labeling. We used for classification classical features as well as features extracted from different scales of the wavelet decomposition of both leads integrated in an RMS manner. Step-wise and a randomized method were considered for feature subset selection, and linear discriminant analysis (LDA) was also used for additional dimensional reduction. Three classifiers: linear, quadratic and Mahalanobis distance were evaluated, using a k-fold like cross validation scheme. Results in the training set showed that the best performance was obtained with a 28-feature subset, using LDA and a Mahalanobis distance classifier. This model was evaluated in the test dataset with the following performance measurements global accuracy: 86%; for supraventricular beats, Sensitivity: 86%, Positive pred.: 20%; for ventricular beats Sensitivity: 71%, Positive pred.: 61%. This results show the feasibility of classification based on the multilead wavelet features, although further development is needed in subset selection and classification algorithms.

1. Introduction

The analysis of the electrocardiographic signal (ECG) provides a noninvasive technique to analyze the heart function for different cardiac conditions. Particularly, automatic classification algorithms focus on ECG rhythm and morphology analysis and their disturbances. Disturbances in the rate, regularity, site of origin or conduction of the electrical impulses are known as arrhythmias [1]. While some types of arrhythmias represent a life threat in the short term (e.g. ventricular fibrillation), there are other types that appear less frequently and represent a long-term threat without proper treatment. It is in those later cases, which require carefully inspection of long ECG recording, where the use of automatic algorithms represents a significant help for diagnostic.

Many algorithms for ECG classification were developed in the last decade [2, 3, 4, 5], achieving very good performance training a global classifier on the well-known MIT-BIH arrhythmia database [6]. Some of these algorithms improved their performance when information from a small excerpt of the recording under classification was used for local adaptation [2, 7, 5]. However, the features extracted from the ECG signal (the model) are quite similar in most of the proposed approaches found up to the moment. Some authors found improvements in the classification performance when features from other leads were accounted in the model [7], or including new features that integrates information present in both leads, like the vectocardiogram (VCG) maximum value (VCG_{max}) and VCG angle (VCG_{angle}) [5]. Our group is currently working in the multilead (ML) approach for ECG delineation based in the wavelet transform [8, 9], and the improvement respect to the single-lead (SL) approach was confirmed by other groups [10]. Since the methodology for feature extraction generally requires previous ECG delineation, we propose an ECG classifier that uses a set of robust features already used in the delineation stage. This is attractive for an efficient real-time implementation, like an implantable or portable device. The objective of this work is to develop and evaluate a model for ECG classification based on ML features, under the hypothesis that these features should provide a better classification performance.

2. Methods

2.1. ECG database

In this work we used the MIT-BIH Arrhythmia database [6] for training and evaluating the classifier. The database consists of 48 two-lead recordings of approximately 30 minutes and sampled at 360 Hz. The first 23 recordings were extracted from routine ambulatory recordings while the remaining 25 were selected because of the presence of less common complex ventricular, junctional and supraventricular arrhythmias. The two recorded leads are

not the same in all recordings, depending on the arrhythmia and physical limitation of the subject's body. The annotations provided with the database were used for training and testing purposes, following the recommendations and class-labeling of AAMI. We adopted the training (DS1) and test (DS2) division scheme used in [4] for comparative purposes. AAMI Class Q beats were discarded since they are poorly represented in the database. The division scheme is summarized in Table 1.

Table 1. Scheme of the division of the MIT-BIH database into training (DS1) and testing (DS2) sets. Recordings with paced beats were excluded.

	N	S	V	F	#Rec.
Full DB	89737	2757	6959	801	44
DS1	45675	926	3767	413	22
DS2	44062	1831	3192	388	22

2.2. Signal processing

Most of the considered features are based on the wavelet transform (WT) analysis of the ECG signal. The WT is defined for a signal $x(t)$ as:

$$W_a x(b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\psi\left(\frac{t-b}{a}\right) dt, a > 0 \quad (1)$$

The WT can be discretized using a dyadic scheme where the scale factor is $a = 2^k$ for $k \in Z^+$, with the same sampling rate at each scale (Algorithme à trous) and is easily implemented as a filter bank. We used a quadratic spline as the prototype wavelet $\psi(t)$, being the derivative of a smoothing function. As a result of this transformation, the original ECG signal can be analyzed as a smoothed derivative at different scales (frequency bands) and translations (times). This type of analysis has been successfully used for ECG delineation based in the analysis of the peaks and zero-crossings of the WT at the most appropriate scales (see [11] for a detailed description of the ECG delineator).

To integrate the information of the available channels, we use a similar concept to those proposed in [8, 10], creating a multilead signal for each scale $a = 2^k$ by computing the RMS of the wavelet transform in the $M = 2$ leads

$$W_{rms_a} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(W_a x_i[n])^2}. \quad (2)$$

This results in one positive signal for each scale. Figure 1 shows an example of a W_{rms} signal.
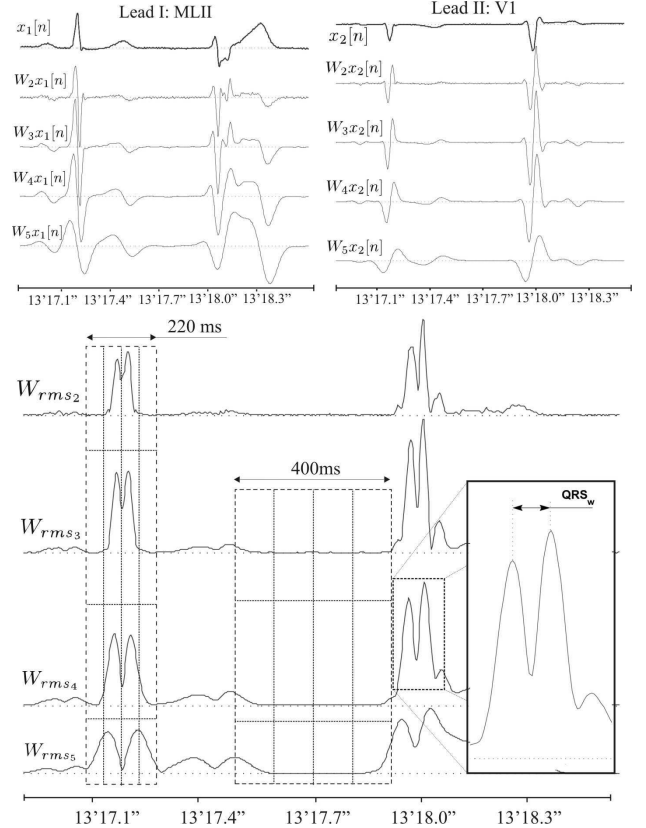

Figure 1. Excerpt of record 201 of MIT-BIH database. Normal (N) and ventricular (V) AAMI class heart beats. In the top figures both ECG leads are shown with their corresponding wavelet decomposition (scales 2-5). The lower panel depicts the RMS composition of both leads wavelet transform (W_{rms}). Some features measured in the W_{rms} signal are also shown.

2.3. Classification features

Following the conclusions of previous works [2, 4, 5], we included in our model both interval and morphological features. As interval features we used the typical features from the RR sequence $RR[i-1]$, $RR[i]$ and $RR[i+1]$, differences between adjacent RR intervals $dRR[i-1]$, $dRR[i]$ and $dRR[i+1]$, where $dRR[i] = RR[i] - RR[i-1]$, and the median of the last 10 RR intervals.

As morphological features we considered an estimate of the QRS width (QRS_W) measured in the $W_{rms_a}x[n]$ signal for scales $a = 2, 3, 4$ as the difference between the two maxima around the fiducial point (FP) (see Figure 1), the maximum modulus of the QRS loop (VCG_{max}) and the angle of the loop at this position (VCG_{angle}) as proposed in [5]. Analogously, we proposed the same concept evaluated in the QRS loop formed by $W_a x_i[n]$ in scales 2 to 5 ($WT_{max,a}$ and $WT_{angle,a}$). We also defined a window

of length 400 ms ending at the QRS onset (QRS_{on}). This window was divided in 4 sections of 100 ms, and the energy of the $W_{rms_a}x[n]$ signal was computed in each section in scales 3 to 5 ($preQRS_i$). This amounts 12 features per beat (4 sections $\times$ 3 scales). The objective of this group of features is to include robust information about the P wave region. The same concept was used to study the QRS complex morphology with a window starting at $FP - 100$ms, and ending at $FP + 120$ms for scales 2 to 5 (QRS_{Ei}), as can be seen in Figure 1 (16 features per beat). Finally the ratios between QRS_{Ei} at consecutive scales lead to 12 new features ($QRS_{Ri} = QRS_{E\,i+1}/QRS_{E\,i}$). All the aforementioned feature values (60 per beat) form together a feature vector (FV) for each heart beat in the database.

To deal with the wild variations that may happen between different recordings, we used the first 5 minutes of the recording to estimate a reference FV (FV_N). This FV_N is calculated as the median of all FV of the beats belonging to the normal class within the first 5 minutes. Then, a relative feature vector FV_{rel} is calculated as

$$FV_{rel} = \frac{FV - FV_N}{FV_N}. \qquad (3)$$

2.4. Model selection and dimensional reduction

Several models with lengths ranging from 5 to 28 features were selected using SPSS step-wise method and a randomized method implemented in Matlab. Those models derived from the SPSS were populated according to two parameters, one for the variables to enter (F_{enter}) the model and the other to leave it (F_{leave}). Other models were chosen after evaluating the performance of 2000 random models for a fixed length. Besides, the effect of dimensionality reduction by Fisher's linear discriminant analysis (LDA) was also studied for every tested model. Model performance analysis is described in the next sections.

2.5. ECG classification

We used both linear and quadratic discriminants for classification purposes. The general quadratic discriminant functions for i classes can be written as

$$g_i(\mathbf{x}) = -\frac{1}{2}\mathbf{x}^T \Sigma_i^{-1} \mathbf{x} + \mu_i^T \Sigma_i^{-1} \mathbf{x} \qquad (4)$$
$$-\frac{1}{2}\mu_i^T \Sigma_i^{-1} \mu_i - \frac{1}{2}\log(|\Sigma_i|) + \log(P(\omega_i))$$

where μ_i, Σ_i are the mean vector and covariance matrix of the feature vector for beats belonging to the i-th class, and $P(\omega_i)$ is the *a priori* probability of each of the classes. The values of μ_i and Σ_i were computed from DS1, and

Table 2. Performance for each model on DS1 separating all classes. The results are expressed in percentages.

		Model			Sensitivities (%)				Positive Pred. (%)			
		Size	FS	Cl	Acc	N	S	V	F	N	S	V	F
Random Subsets		**28**	**L**	**M**	**81**	**83**	**76**	**74**	**25**	**99**	**11**	**77**	**4**
		20	L	M	81	82	74	79	27	99	10	74	5
		28	G	L	75	77	55	63	5	98	8	76	0.4
Step Wise		28	L	L	95	99	22	61	59	97	25	90	45
		28	L	Q	86	88	50	73	64	99	12	65	13

FS: feature set. Local adapted (L) or global (G).
Cl: classifier. Linear (L), quadratic (Q) or mahalanobis (M).

equal *a priori* probabilities were considered. The classification rule assigns a FV $\mathbf{x}$ to the class i which results in the maximum $g_i(\mathbf{x})$.

A related classifier is that based on the minimum Mahalanobis distance, which assigns the class which minimizes $(\mathbf{x} - \mu_i)^T \Sigma_i^{-1} (\mathbf{x} - \mu_i)$.

When covariance matrices Σ_i are assumed equal for all classes, the discriminant functions become linear. In this case, a pooled covariance matrix Σ is calculated as in [4].

2.6. Performance evaluation

All models were evaluated in DS1 using a k-fold cross-validation scheme, where each fold is one of the 22 recordings present in DS1. Then performance was measured in terms of the global accuracy, specificity, and class sensitivity as defined in [4]. These performance measures resulted from three classifiers: linear, quadratic and minimum Mahalanobis distance; with the assumption of equal prior probability. Those models which maximized the criteria $j = S_{Se} + S_{+p} + V_{Se} + V_{+p}$ were considered for a final comparison, where X_{Se} and X_{+p} stand for sensitivity and positive predictivity of class $X \in \{S, V\}$ respectively, with S: supraventricular premature contractions and V: ventricular premature contractions.

The methodology used for the final test is the same proposed in [4] following AAMI recommendations. Thus DS1 was used to train the best performing model and DS2 exclusively for testing purposes.

3. Results

Results of the model selection procedure are shown in Table 2. This table shows the ability of best models separating all classes on DS1. The model with best overall figures was chosen (highlighted model in table 2). This model was used for the final performance measure on DS2 after training on DS1. Results and confusion matrix in the test set are shown on table 3 and 4 respectively.

Table 3. Final evaluation of the best performing model, on DS2 separating S and V from other classes. Results from previous works are shown for comparison. All results are expressed in percentages.

| | S | | | | V | | | |
	Acc	Se	+P	FPR	Acc	Se	+P	FPR
de Chazal	95	76	39	5	97	78	82	1
this work	86	86	20	13.6	95	71	61	3

Acc: Accuracy; Se: Sensibility; +P: Positive predictability
FPR: False positive rate

Table 4. Confusion matrix for the final evaluation of the best model on DS2.

| | | Algorithm | | | |
		N	S	V	F
Annotation	N	35309	6147	1340	1266
	S	133	1578	105	15
	V	5	317	2253	617
	F	52	15	4	317

4. Discussion and conclusions

In this work we proposed and studied a set of features derived from a multilead wavelet transform of the ECG. According to our results, the best performing model consisted in 28 features locally adapted (FV_{rel}), followed by LDA for dimensionality reduction and a minimum Mahalanobis distance classifier. For this model the first LDA dimension mostly separates classes N and V, and the features most correlated to this projection are those related to QRS morphology (QRS_W, QRS_{Ei} and QRS_{Ri}). The second and third LDA projections jointly separate classes N and S. The features more correlated with them are the $preQRS_i$. However, as it can be seen in Tables 3 and 4 the discrimination ability is still poor.

It was observed in the final test an unexpected improvement in S sensitivity and positive predictivity with respect to the training set performance, previously seen in [4]. This improvement could be caused by the different amount of class S examples present in DS1 and DS2. Most previous works present the classification performance as in table 2 or 3, following AAMI recommendations. However, in databases like MIT-BIH where one class is one or two order of magnitude greater than the others, some indicators like accuracy or false positive rate may become optimistically biased (see table VII in [4]), and sensitivity or positive predictivity should be used. Also the confusion matrix provides greater insight into the classification performance.

This preliminary work showed the feasibility of using the multilead wavelet analysis features for ECG classification, though more development is needed to achieve the performances obtained in previous works.

Acknowledgments

References

[1] Taylor GJ. 150 Practice ECGs: Interpretation and Review. Blackwell Science, 2002. ISBN 0-632-04623-6.

[2] Hu YH, Palreddy S, Tompkins W. A patient-adaptable ecg beat classifier using mixture of experts approach. IEEE Transactions on Biomedical Engineering 1997;44:891–899.

[3] Lagerholm M, Peterson C, Braccini G, Edenbrandt L, Sörnmo L. Clustering ecg complexes using hermite functions and self-organizing maps. IEEE Transactions on Biomedical Engineering 2000;47:838–848.

[4] de Chazal P, O´Dwyer M, Reilly RB. Automatic classification of heartbeats using ecg morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering 2004;51:1196–1206.

[5] Christov I, Jekova I, Bortolan G. Premature ventricular contraction classification by the kth nearest-neighbours rule. Physiological Measurement 2005;26:123–130.

[6] Mark R, Moody G. Mit-bih arrhythmia database 1997. http://ecg.mit.edu/dbinfo.html, 1997.

[7] de Chazal P, Reilly RB. A patient-adapting heartbeat classifier using ecg morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering 2006; 53:2535–2543.

[8] Llamedo Soria M, Martínez J, Laguna P. A multilead wavelet-based ecg delineator based on the rms signal. In Computers in Cardiology 2006. IEEE Computer Society Press, 2006; 153–156.

[9] Almeida R, Martínez JP, Rocha AP, Olmos S, Laguna P. Automatic multilead vcg based approach for qt interval measurement. In Computers in Cardiology 2006. IEEE Computer Society Press, 2006; 369–372.

[10] Lux RL, Gettes LS, Mason JW. Understanding proarrhythmic potential in therapeutic drug development: alternate strategies for measuring and tracking repolarization. Elsevier Journal of Electrocardiology 2006;39:S161–S164.

[11] Martínez J, Almeida R, Olmos S, Rocha A, Laguna P. A wavelet-based ecg delineator: Evaluation on standard databases. IEEE Transactions on Biomedical Engineering 2004;51:570–581.

Address for correspondence:

Mariano Llamedo Soria
Universidad Tecnológica Nacional – Dpto. de Electrónica
Medrano 951 - C1179AAQ – Capital Federal, Argentina
llamedom@electron.frba.utn.edu.ar