35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Detection of Myocardial Ischemia Episode Using Morphological Features*

Cheng-Hsiang Fan, Yu Hsu, Sung-Nien Yu, *Member, IEEE*, and Jou-Wei Lin

***Abstract*—In this study, we propose to use morphological features that are easy to identify to differentiate myocardial ischemic beats from normal beats. In general, myocardial ischemia causes alterations in electrocardiographic (ECG) signal such as deviation in the ST segment. When the ST segment level deviates from a certain voltage, the beat would be diagnosing as myocardial ischemia. To emphasize on ST variations, the QRS complex of the ECG signal was first subtracted and replaced with a straight line. Five–level discrete wavelet transform (DWT) followed to decompose the waveform into subband components and the A5 subband, which is most sensitive to the changes in the ST segment, was reconstructed for the calculation of 12 morphological features. The support vector machine (SVM) and the 10-fold cross-validation method were employed to evaluate the performance of the method. The results show high values of 95.20%, 93.29%, and, 93.63% in sensitivity, specificity, and accuracy, respectively, that were demonstrated to outperform the other methods in the literature.**

I. Introduction

"Myocardial ischemia is the pathological state underlying ischaemic heart disease. It can lead to myocardial infarction (commonly known as heart attack) which in its acute form can lead to the death of the affected person." [1]. The most important cause of myocardial ischemia is coronary artery stenosis or obstruction. Myocardial ischemia causes alterations in electrocardiographic (ECG) signal such as deviation in the ST segment [2]. When the ST segment deviates more than a certain level, the beat would be diagnosing as myocardial ischemia.

Recently, several studies have been conducted to developing computer-aided diagnosis algorithms for the diagnosis of myocardial ischemia. Exarchos and coworkers proposed to use rule-based mining technology to identify myocardial ischemia heartbeat from normal heartbeat in 2006 [3]. Khoshnoud and coworkers used subband ECG signal decomposition with multi-level wavelet analysis and claimed that their method provided an easier way to locate the important points of the waveform for myocardial ischemia diagnosis with probabilistic neural network (PNN) [4].

*Research supported by the National Science Council and the Ministry of Education, Taiwan, Republic of China.

Cheng-Hsiang Fan, Yu Hsu, and Sung-Nien Yu are with the Department of Electrical Engineering and the Advanced Institute of Manufacturing with High-tech Innovations, National Chung Cheng University, Chiayi County, Taiwan (phone: +886-5-2720411 ext 33205; fax: +886-5-2720862; e-mail: fanchenshan@gmail.com; anderson23i@hotmail.com; ieesny@ccu.edu.tw (corresponding author)).

Jou-Wei Lin is with the Cardiovascular Center, National Taiwan University Hospital Yun-Lin Branch and the College of Medicine, National Taiwan University, Dou-Liou City, Yun-Lin County, Taiwan. (e-mail: jouweilin@yahoo.com).

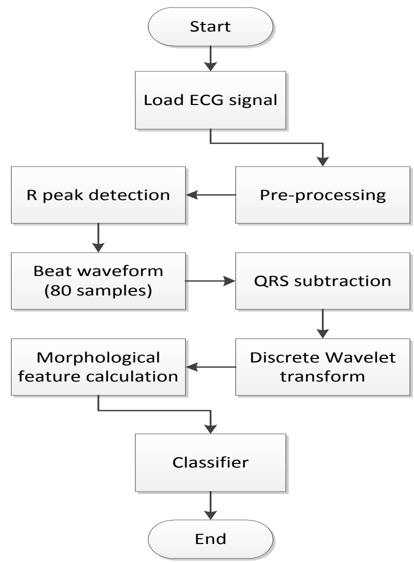

Figure 1. Experimental procedure.

However, conventional methods usually required to find the fiducial points, e.g. T wave, R peak, ISO point, J point, as features for myocardial ischemia detection [5], but the fiducial points may not be easy to locate when the ECG signal is noisy. Therefore, in this study, we proposed to use morphological features that were calculated from the entire heartbeat waveform. With this method, only the R point of the heartbeat, which is the easiest to locate, is to be located and the requisite of extremely clean signal for accurately locating several key points is loosened. The performance of the method was validated using support vector machine classifier and 10-fold cross-validation method.

II. Methods

A. Database:

The data used in the experiments were obtained from the "European Society of Cardiology (ESC) ST-T database" [6]. This database includes 78 data files recorded from myocardial ischemia patients. Each file contains two-lead, two-hours ECG signals sampled at 250 Hz. The start and end times of the ST-segment changes (myocardial ischemia episode; MI episode) were clearly annotated in the files.

B. Discrete Wavelet Transform

Discrete wavelet transform (DWT) was employed to decompose the ECG signals into subband components. The DWT provides a good time-frequency representation of a signal by using variable sized windows. Long time windows are used to get a finer low frequency resolution. Short time windows are used to get high frequency information. WT is suitable for the analysis of non-stationary signals such as ECG.

The ECG signal can be decomposed into finer details by multi-level discrete wavelet transform (DWT) using high-pass (*g[n]*) filter, low-pass (*h[n]*) filter, and downsampling (↓2). [7]. After the first level decomposition, two signals representing the detail (high-frequency) and the approximate (low-frequency) are obtained. The approximate signals are further decomposed into the detail and the approximate after the second level decomposition, et al, as depicted in Fig. 2 (a). Subband components can be reconstructed back to the length of the original signal *x[n]* by inverse DWT (IDWT), as depicted in Fig. 2 (b). This process can also be used to eliminate noises by setting components in certain subbands to be zero and perform the IDWT.

C. Preprocessing

The aim of the preprocessor was to remove the baseline wander and noise artifacts frequently observed in ECG signals. ECG baseline wander usually caused by breathing or unexpected movement of experimental settings, which usually cover the frequency range below 1Hz [8]. In order to eliminate baseline wander, we used DWT to decompose signal to the seventh level and then set the approximate coefficient A7 to zero and perform the IDWT. In this manner, subband components below 0.97 Hz were removed from the signal.

The second part of the preprocessor was to remove noise artifacts.The soft-thresholding method proposed by Donoho [9] was adopted for this purpose. Seven levels of DWT were applied to the signal first. The subband coefficients with minor values were considered noise and were set to zero before the application of IDWT to eliminate the noise.

D. R Peak Detection

In order to accurately locate the R peaks in ECG, we focused only on the D_3 and D_4 subband components that show the most significant features of the QRS complex [10]. First of all, all the other subband components, except that of D_3 and D_4, were set to zero. Moreover, to highlight the location of the QRS complex, the coefficient values in both D_3 and D_4 were squared and the smaller values (threhold= standard deviation of the reconstruction signal) eliminated before performing IDWT. The location of the peaks in the reconstructed signal were the tentative positions of the R peaks, as depicted in Fig. 3 (a). However, since DWT sometimes causes minor shift of the waveform, the positions of the real R peaks (Fig. 3 (b)) were determined by searching the highest peaks in the vicinity (±10 samples) of the tentative R peaks (Fig. 3 (a)) in the preprocessed ECG signal.

E. Calculation of Morphological Features

After the R peak has been located, a 80-point waveform, with 35 samples before and 44 samples after the R peak (Fig. 4), was segmented as the representative waveform of a

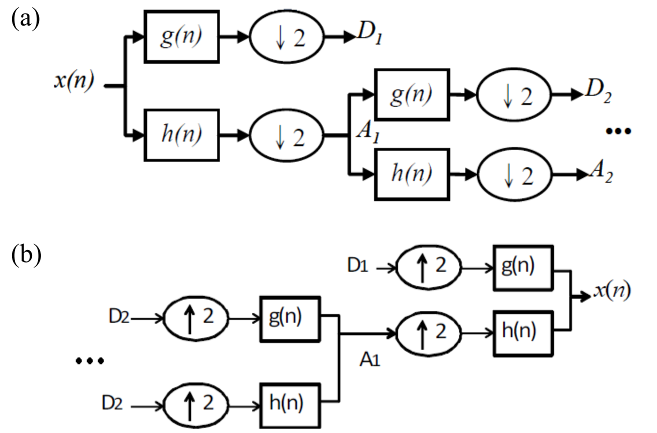

Figure 2. Discrete wavelet transform (DWT). (a) forward DWT (decomposition); (b) inverse DWT (reconstruction).

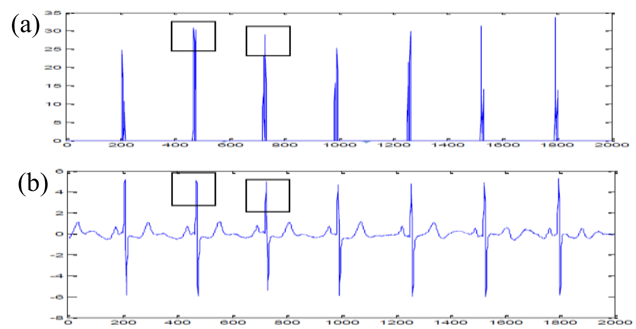

Figure 3. R peak detection. (a) tentative R peaks; (b) real R peaks after vicinity searching.

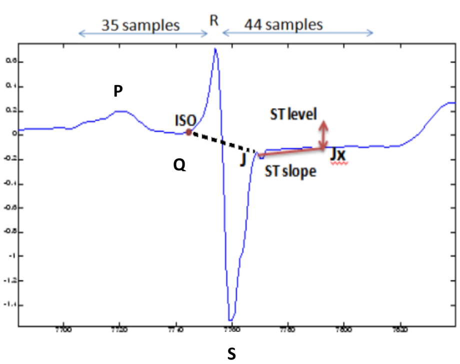

Figure 4. 80-sample ECG waveform of a heartbeat.

heartbeat. The three points *J*, *JX*, and *ISO* closely related to the R peak were first located on the waveform according to the heart rate, as shown in Table I and Fig. 4. The QRS complex was defined as the part of waveform between the *ISO* and *J* points. In order to concentrate only on the features associated with ST segment, the QRS complex was first removed from the original waveform and replaced with a straight line (dash line in Fig. 4).A five-level DWT followed to decompose the QRS subtracted waveform into different subbands. The low-frequency part (A5) of the 5^{th}-level DWT was reconstructed using IDWT. Six features were exploited to characterize the reconstructed A5 component, namely (1) the power the A5 component (Power), (2) the power ratio of the A5 to the original signal (Power ratio), (3) JX potential, (4) ST level, (5) ST deviation, and (6) ST slope, as summarized in Table II.

In order to characterize the variation of the waveform from a typically "normal" one, a reference waveform was generated by calculating the average waveform of the 80-point beat waveforms in the first 30 sec record, as adopted by the European Society of Cardiology to calculate the "normal" waveform for the database [6]. Six features associated with the relationship between the reconstructed A5 components from the test and the reference waveforms were calculated, including (1) the correlation coefficient (CC), (2) the mean (Mean_CF) and (3) the standard deviation (SD_CF) of the cross-correlation function, and (4) the mean (Mean_D), (5) the standard deviation (SD_D), and (6) the power (Power_D) of the difference waveform between the test and reference waveforms.

Each feature was normalized by subtracting the mean value from the feature and dividing by the feature's standard deviation. This process intended to normalize all the features to the same level.

F. Support Vector Machine Classifier

Support vector machine (SVM) maps the training samples from the input space into a higher-dimensional feature space via a mapping (kernel) function [11]. Any product between vectors in the optimization process can be implicitly computed to generate a hyperplan to categorize the samples into two classes.

For a training set of instance-label pairs (x_i, y_i), i=0,…l, where $x_i \in R$ and y_i=[-1,1], and a non-linear operator mapping with kernel function φ, the optimization problem becomes

$$\min_{w,b,\xi} \frac{1}{2} w^T w + S \sum_{i=1}^{l} \xi_i \tag{1}$$
$$\text{subject to} \quad y_i(w^T \varphi(x_i) + b) + \xi_i - 1 \geq 0, \xi_i \geq 0$$

where $S>0$ is the penalty parameter for the error term and ξ_i is the set of slack variables that is introduced when the training data is not completely separated by a hyperplane. To solve this problem, Vapink [11] has shown that the solution can be found by minimizing both the errors on the training set (empirical risk) and the complexity of the hypothesis space. Consequently, the decision found by SVM is a tradeoff between error and model complexity. Numerous studies have demonstrated the superiority of using SVM classifier over other classifiers in pattern classification tasks. Consequently, we employ the SVM classifier in the study. The radial basis function (RBF) was empirically selected as the kernel function of the SVM classifier.

III. Results And Discussions

Fourteen data files were selected from the database for experiments. Based on the information about myocardial ischemia episode provided by the database, 3970 ischemic and 28890 normal heartbeat waveforms were segmented from the data files for analysis.

TABLE I. Locations of the Key Points

Heart rate (HR)	***J***	***JX***	***ISO***
HR > 120 bpm	R+40ms	R+60ms	R-40ms
120 bpm > HR > 100 bpm	R+40ms	R+60ms	R-40ms
HR < 120 bpm	R+60ms	R+80ms	R-40ms

TABLE II. Morphological Features Calculated from the QRS Removed waveform

Feature	***Feature Description***	
Power	Power of the A5 subband signal	Feature 1
Power ratio	Power ratio of the A5 subband to the original signal	Feature 2
JX potential	Potential of the *JX* point	Feature 3
ST level	Potential difference between *J* and *ISO*	Feature 4
ST deviation	ST level change from the normal	Feature 5
ST slope	Slope of the segment between *J* and *JX*	Feature 6
CC	Correlation coefficient of the test and reference waveforms	Feature 7
Mean_CF	mean of the cross-correlation function	Feature 8
SD_CF	Standard deviation of the cross-correlation function	Feature 9
Mean_D	Mean of the difference waveform of the test and reference waveforms.	Feature 10
SD_D	Standard deviation of the difference waveform	Feature 11
Power_D	Power of the difference waveform	Feature 12

The performance of the classifier was measured by three statistics indices, namely (1) specificity: the percentage of correctly classified normal beats among the total normal beats; (2) sensitivity: the percentage of correctly classified myocardial ischemia beats among the total ischemic beats; (3) accuracy: the percentage of correctly classified beats among all the beats.

The ten-fold cross-validation method [12] was employed to evaluate the performance of a classifier. The test sample beats were firstly divided into ten test sample groups with the same distribution of attribute. Each sample group was alternatively reserved as the test group. The other nine groups were used to train the classifier and the performance of the classifier was measured by using the reserved group as test samples. This procedure repeats until all the sample groups had been reserved once as test samples. The performance of the classifier was evaluated by the average values of the three indices in the ten trials.

The results were summarized in Table III. The proposed morphological features and SVM classifier achieved a sensitivity of 84.69% and a specificity of 97.25%, resulting in an accuracy of 95.22%. The results were impressive, especially with the high specificity and accuracy. However, we have noticed that the sensitivity was much lower than the specificity. This phenomenon was caused by the imbalanced data sets, which would favor the major (normal) class and ignore the minor (ischemic) class.

Therefore, we sought to resolve this problem with over-sampling [13], which increases the number of samples in the minor class with data interpolation (or over-sampling) based on the real samples to the same level of the major class. The performance of the classifier using over-sampling in the training phase is demonstrated in Table IV. Comparing the performance in Table III and Table IV, a dramatic increase in the sensitivity was observed with the over-sampling technique. Only a minor decrease in specificity was observed, which was believed to be the compensation caused by oversampling in the minor (ischemic) class. The two effects resulted in a classifier equally effective in recognizing normal and ischemic ECG beats.

The performance of the proposed system was compared to that of two representative methods published in the literature, although the databases were not exactly the same. One is the rule-based mining method proposed by Exarchos and coworkers [3], which achieved 87% in sensitivity and 93% in specificity. The other is the method proposed by Khoshnoud and coworkers, who used subband ECG signals for locating important myocardial ischemic points for classification with probabilistic neural network (PNN) [4]. A sensitivity of 96.67% and a specificity of 89.18% were reported. The comparative results were summarized in Table V. It is impressive that the proposed method with over-sampling achieved sensitivity and high specificity, which is superior to the other methods that only show large value in either sensitivity or specificity. This property of the method is believed to be favorable for a computer-aided myocardial ischemia diagnosis system.

IV. Conclusion

We proposed a method for the detection of ischemic heartbeats based on morphological features. The objective of the study was to use only the R point which is easily identifiable and bypass the need to identify the key points that are apt to be buried in noise and might be difficult to be correctly located, such as the S and T points. Easily identifiable key points only depending on the location of the R point and the heart rate were used instead (Table I). Morphological features were calculated from the A5 subband components of the QRS complex subtracted waveform. This approach minimized the interference of the QRS complex in the calculation of ST-related features and only focused on the variation of the ST segment in characterizing ischemic waveform.

Impressive performance was observed with the morphological features. The application of oversampling technique to balance the samples in the two data sets further improved the performance of the classifier. The results demonstrate the effectiveness of the proposed method in accurately detecting ischemic beats using morphological features that are easy to calculate.

TABLE III Experimental results

Beat Type	Number of Samples (Train+Test)	Sensitivity (%)	Specificity (%)	Accuracy (%)
Ischemic	6489+722	84.69	97.25	95.22
Normal	33724+3748			

TABLE IV The effect of over-sampling in Ischemic class

Beat Type	Number of Samples (Train+Test)	Sensitivity (%)	Specificity (%)	Accuracy (%)
Ischemic	33724+722	95.40	93.29	93.63
Normal	33724+3748			

TABLE V Comparison with other studies

Method	Sensitivity	Specificity	Accuracy
Rule-based [3]	87%	93%	90%
PNN classifier [4]	96.67%	89.19%	90.75%
Proposed method	95.40	93.29	93.63

Acknowledgment

This study was supported in part by the grants NSC 100-2221-E-194-063 and NSC101-2221-E-194-019 from the National Science Council and a grant from the Ministry of Education, Taiwan, Republic of China.

References

[1] Wikipedia: Myocardial Ischemia http://en.wikipedia.org/wiki/Myocardial_ischemia

[2] M. S. Thaler, *The Only EKG Book You'll Ever Need*, 6th ed. New York: Lippincott Williams and Wilkins, 2009.

[3] T. P. Exarchos, C. Papaloukas, D. I. Fotiadis, and L. K. Michalis, "An association rule mining-based methodology for automated detection of ischemic ECG beats," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 8, Aug. 2006.

[4] S. Khoshnoud, T. Mohammad, and M. A. Shoorehdeli, "Probabilistic neural network oriented classification methodology for ischemic beat detection using multi resolution wavelet analysis," *Proc. of the 17th Iranian Conference of Biomedical Engineering (ICBME2010)*, pp.1-4, 2010.

[5] K. Daskalov, I. A. Dotsinsky, and I. I. Christov, "Developments in ECG acquisition, preprocessing, parameter measurement, and recording," *IEEE Eng. Med. Biol.*, vol. 17, no. 2, pp. 50–58, Mar./Apr. 1998.

[6] European Society of Cardiology (ESC) ST-T Database: http://physionet.org/physiobank/database/edb/

[7] M. Weeks, *Digital Signal Processing Using MATLAB and Wavelets, 2nd ed.* Sudbury, MA: Jones and Bartlett Publishers, 2011.

[8] Z. Donghui, "Wavelet approach for ECG baseline wander correction and noise reduction," *27th Annual International Conference of the IEEE Engineering in Medicine and Biology Soiciety*, Shanghai, China, pp. 1212 - 1215, 2005.

[9] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Information Theory*, vol. 41, no. 3, May, 1995.

[10] Z. Dabanloo, "Detection of QRS complexes in the ECG signal using multiresolution wavelet and thresholding method," *Computing in Cardiology*, pp. 805 - 808, 2011.

[11] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995.

[12] R. M. Rangayyan, *Biomedical Signal Analysis: A Case-Study Approach*. New Jersey: John Wiley & Sons, 2002.

[13] H. Haibo and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowledge and Data Eng.*, vol. 21, pp. 1263-1284, 2009.