

ECG PVC Classification Algorithm based on Fusion SVM and Wavelet Transform

Huang Dong¹, Liao Zhengquan², Li Changbin¹, Li Dan¹ and Huang Wendong¹

¹*College of Mathematics and Computers Science, QinZhou University, QinZhou, China, 530099*

²*Internal Medicine-Cardiovascular Department, First People's Hospital of Qinzhou, QinZhou, Guangxi, China, 535000*
qzxyhd@163.com

Abstract

In the process of ventricular premature beat (PVC) and normal sinus rhythm (NSR) identification base on electrocardiogram (ECG), there exists problems like negative effect from ECG rhythm and low recognition rate. This paper proposes the electrocardiogram PVC classification algorithm based on support vector machine (SVM) and wavelet algorithm. The algorithm uses the wavelet transform to analyze ECG beating model, which is not influenced by the change of ECG waveform. The two feature sets respectively compose of statistical parameters of the wavelet coefficients and the selected wavelet coefficients. PVC and NSR are analyzed by using SVM. The experimental results show that this method improves the recognition rate of ECG.

Keywords: *Wavelet transform; Eelectrocardiogram; Support vector machine; PVC classification*

1. Introduction

ECG(Electrocardiogram; ECG)is an important physiological signal used in the diagnosis of heart disease. Premature ventricular contractions (premature ventricular contraction; PVC) is the most common form of cardiac arrhythmia disease. The understanding of the disease is the basis for forecasting the evolution of ventricular arrhythmias. Ventricular arrhythmias may increase the risk of patients with ventricular contraction occurred before, which will pose the patient a mortal threat. Some researchers in the detection of PVC ectopic beats have found that an ectopic pacemaker located in the ventricle, the ventricular depolarization is detected regularity assessment of ventricular tachycardia and is particularly effective.

There are many methods have been mentioned in the feature extraction and classification of ECG, for example, wavelet transform, fourier transform, lyapunov transformation, AR parametric model analysis. How to extract features from multi-dimensional ECG data is the key to these theories. For example, Reference [1] proposes a PVC recognition method based on wavelet transform and BP neural network, recognizing for PVC by selection of the characteristic parameters of the ECG signal, extracting feature parameters using wavelet transform, structural feature vector, then use the BP neural network trained. Reference [2] proposes a discriminant variable width template algorithm based multi-template matching premature ventricular contractions. Depending on the template to determine the width of the RR interval, queue-based template matching and template library that matches the second match strategy, using the correlation coefficient and the RR interval than the combination of algorithms to detect chamber PVC. Reference [3] proposes a method for the identification of premature ventricular contraction(PVC)base on wavelet transform and BP Neural Network. To achieve normal rhythm of the ECG signal and premature ventricular contractions of automatic detection

and identification by selection of the ECG signal characteristic parameter, using wavelet transform to detect the characteristic points, then detect feature points using wavelet transform to construct a three-layer BP neural network structure of 7-10-1. Using BP neural network which is trained on the identification of premature ventricular contractions.

Currently, the ECG feature extraction of PVC has become one of the heated topic of biomedical engineering and signal processing, especially the extraction of PVC wavelet coefficient under normal or unnormal conditions. But the challenges faced in the detection process are PVC beats are changing. This paper presents One kind of model is not affected by the beat detection algorithm, Using wavelet transform to extract two feature sets from ECG, Then identify analyzed by support vector machine (SVM).

2. ECG research Methods

2.1. ECG Filter and Heart Beat Detection

The data of Table 1 is quoted from MIT-BIH database, which is employed to test the correctness of the proposed method. ECG signal sampling frequency is 360Hz. The band pass filter with the lower limit of frequency band of 1Hz and upper limit of frequency band of 35Hz are used to carry out the wave filtering of ECG Signal, so as to eliminate the bother of baseline drift and power.

Table 1. Assessment Data of MIT-BIH Arrhythmia Database

Record number	waves heart(total)	The number of normal heart rhythm	The number of premature ventricular contractions
106	2027	1507	520
210	2617	2423	194
221	2427	2031	396
233	3061	2230	831

The first step in ECG feature extraction and classification is the beating heart beat detection. Challenges faced in the detection process are the change of heart beat rhythm waveforms. Proposed algorithm is an effective solution to this problem in this section. The heart beat will not be affected by the waveform by the algorithm.

By studying the characteristics of wavelet coefficients, we found ECG wavelet transform, Wavelet model significantly when the heart beats. This can be detected by beating beat, Steps of the algorithm are as follows:

(1) First with a moving window equally ECG signal 10, in the ECG signal is filtered on the use of such averaging methods.

(2) After the signal filtering the signal travels an average continuous wavelet transform, wavelet transform models are continuous wavelet transform a 5 layer.

(3) This transformation signals through linear approximation of data transmission (LADT) after compression algorithm, this algorithm unit distance greater than or equal to 5 points discarded, leaving it to the point of unit distance in 5 or less.

(4) Determine the two thresholds. Two threshold values will be approximately linear in a data matrix C is used, two threshold values are set by a 3-second window function to perform calculations in the following step is obtained:

$$CP = C > 0$$

$$CN = -C$$

$$BP = Mean(CP) \quad (1)$$

$$BN = Mean(CN)$$

(5) Positive threshold and negative threshold values are given by:

$$THP = 1/3(BP + Max(CP)) \quad (2)$$

$$THN = 1/3(NP + Max(CN)) \quad (3)$$

(6) Calculating the difference between CP and CN of the sample.

(7) If the result of step 6, the amplitude is higher than at any point THP value or lower value than THN, then, before 50 points and 50 points after the detection point will be to find out if there is a point in the amplitude after more than the previous one and a point, the point will be recorded in a temporary beat detection matrix.

(8) Since the low frequency wave amplitude fast beats, heart requires at least 200 ms to beat again. So the point if any of the following points of the interval less than 200 milliseconds, the detected will be discarded.

2.2. ECG Feature Extraction

In the present study, the set of features extracted from the ECG two features, a set of features including variance and entropy of the wavelet coefficients, And a continuous ECG beat RR ratio[3], characterized by a two-stage selection set 2 wavelet coefficients is further selected in the first stage the second stage of the pre-selected on the basis of the wavelet coefficients.

Characterized 1 can extract by the following set of steps [4,5]:

(1) Selected ECG signal is averaged out through a bandpass filter and a zero-phase moving window.

(2) Algorithms to detect the ECG signal spend mentioned section.

(3) With a range of 200 milliseconds before the QRS detection points to 400 milliseconds after the QRS wave detection point Hamming window to split ECG.

(4) With a 10-order Daubechies wavelet six discrete wavelet transform, decompose each ECG signal band.

By analyzing the results and considering the advantages of a combination of features, the feature set 1 has been defined as:

(1) The variance of the wavelet coefficients.

(2) The entropy of wavelet coefficients.

(3) Continuous ECG beat R-R ratio.

Feature Set 2 can be extracted by the following methods, including through a pre-selection method to the features to be preselected.

The first stage is characterized by a pre-selected, extracted from the wavelet transform might be a useful feature 412 coefficients, the wavelet coefficients according to the magnitude of these sort, you can find a lot of factors are similar to a value of 0, which means that the ECG most of the information is contained in the relatively few coefficients. The second stage selection record the highest average ranking coefficient as candidate features, the process is defined as follows:

(1) The potential of the wavelet coefficients of each record, which are arranged in descending order according to the sequence, each wavelet coefficient is assigned a respective scale factor.

(2) According to the records of the coefficients can know the sum of the sequence.

(3) Select the lowest sequence sum of wavelet coefficients.

Feature selection: realized by a front end to select the feature selection process, may identify a subset of all of the features characteristic ratio obtained with the same or better performance, with an empty set of features, all possible feature set will produce feature. Set on the performance of these features were evaluated, and retain the best performance

among the set. By this additional feature set to expand to the optimal set of two feature sets, these features sets again the classification performance evaluation, and to retain one of the best set of performance characteristics, repeating this process until after the additional feature set classification performance is not until then significantly improved, then this feature set is the set of features we need 2.

In this study, classification has been carried out by using support vector machine (SVM) approach, and property is classed by using linear classifiers, which will be described in the following sections.

3. The Basic Theory of SVM

The traditional method of conducting statistical pattern recognition machine learning, emphasizing the empirical risk minimization. The simple empirical risk minimization will produce "over learning problems", its poor ability to promote. It is only in the sample tends to infinity, to ensure its performance only theory. However, statistical learning theory (STL) capable of machine learning in finite samples, SVM used in this paper is the theoretical basis of statistical learning theory.

Support vector machines proposed by Vapnik, its optimization problem when setting constraints for training error, optimization objective to minimize the confidence range. SVM is a learning method based on minimization criterion, compared with some traditional learning methods, which clearly has the ability to promote optimum. In the SVM solution process will eventually translate into solving quadratic programming problems, so the solution of SVM is globally optimal solution obtained.

3.1. Discriminant Function

Discriminant function refers to a function by a linear combination of the various components of the x made.

$$g(x) = w^T x + w_0 \quad (4)$$

For the decision rules for two types of problems:

If $g(x) > 0$, it is determined that x belongs to A

If $g(x) < 0$, it is determined that x belongs to B

If $g(x) = 0$, then x can be any

3.2. Fisher Linear Discriminant Method

The main conception of Fisher linear discriminant method is to sample the multidimensional space projected onto a one-dimensional space, that is the dimension of the data is reduced to one-dimensional. However, the samples will be multidimensional space to the projected one-dimensional, multi-dimensional samples may overlap better together and cannot separately. But in general, it can be observed in a certain direction, so that it can well are on a straight line in the direction of the obtained samples are projected. As shown in Figure 1.

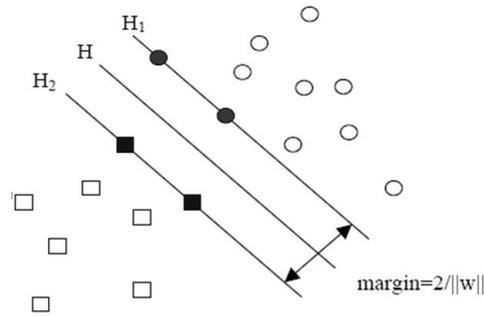


Figure 1. Fisher Linear Discrimination

3.3. Optimal Classification Plane

SVM is obtained from the evolution of optimal classification plane linearly separable scenarios, the basic idea of two-dimensional scene description available, as shown in Figure 2.

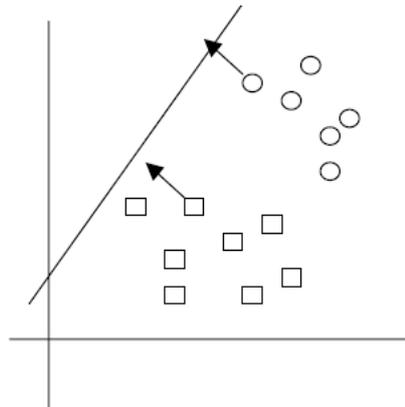


Figure 2. Linearly Separable Scenarios Optimal Classification Line

In Figure 2, the square dots and circles represent two different types of samples, H for sorting line, H1, H2, respectively, parallel to the line classification of the line H, These two types of straight lines pass through the sorting line H from the most recent sample, the distance between H1 and H2 is called a class interval. We will be able to correctly distinguish between two types of positive (training error rate = 0), and can maximize classification line spacing becomes optimal classification lines. If the evolution of the high-dimensional space, the optimal classification line will correspond to the optimal classification surface.

Problem solving optimal classification surface can be transformed into a constraint optimization problem, as follows:

$$\begin{aligned}
 & \text{minimize} \quad \frac{1}{2} \|w\|^2 \\
 & \text{subject to} \quad y_i [wx_i + b] - 1 \geq 0 \quad i = 1, 2, \dots, l
 \end{aligned} \tag{5}$$

Equation describes a convex quadratic programming problem, which due to the objective function and constraints are presented convex nature, According to optimization theory shows that there is only one global minimum of such planning issues. Using Lagrange multiplier method, and to satisfy the KKT (Karush-kuhn-Tucher) conditions:

$$a_i \{y_i [wx_i + b] - 1\} = 0 \quad (6)$$

So you can get an optimal classification function to solve the above problems:

$$f(x) = \text{sgn} \{w^* \cdot x + b^*\} = \text{sgn} \left\{ \sum_{i=1}^k a_i^* y_i (x_i \cdot x) + b^* \right\} \quad (7)$$

Where in a^*, b^* to determine optimal dividing hyperplane parameters, $(x_i \cdot x)$ is the dot product of two vectors.

3.4. Support Vector Machine

In the optimal classification function obtained in the previous section, only the classification of samples and training samples treated in support vector inner product operation. So, for a feature space, the feature space only needs to know the inner product operation, we can solve the optimal linear classification problems.

For nonlinear problems, it can be non-linear transformation, transformed into a high-dimensional space of linear problems. However, seeking to transform the high-dimensional space optimal classification was fairly complex, so under normal circumstances is not easy to adopt this approach.

According to the theory of functional analysis, as long as the kernel function $K(x_i, x_j)$ satisfies the Mercer condition, it will be transformed with a space corresponding to the product. Therefore, choosing the right inner product function $K(x_i, x_j)$ in optimal classification surface, you can classify linear space after the nonlinear transformation, Classification function corresponding evolution:

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^k a_i^* y_i K(x_i \cdot x) + b^* \right\} \quad (8)$$

The above description is the basic theory of SVM, Survey says, SVM is the first by the inner product function, nonlinear transformation of the input space into a high-dimensional space, and then seek the optimal classification surface of this high-dimensional space. SVM classification function and neural network classification similar, Its output is a linear combination of the intermediate point, wherein each intermediate point and the input sample and a support for a corresponding vector inner product, it can also become SVM support vector network.

3.5. Selection of Kernel Function

Use different kernel functions, you can enter the space to construct different types of non-linear decision-face learning machine, which can form different support vector algorithms. Under normal circumstances, the general is directly given kernel function type. Commonly used types of kernel functions are as follows:

- (1) Linear kernel: $K(x_i, x_j) = (x_i, x_j)$
- (2) Polynomial kernel function: $K(x_i, x_j) = (s(x_i, x_j) + c)^d$
- (3) RBF kernel function: $K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right)$
- (4) Sigmoid kernel function: $K(x_i, x_j) = \tanh(s(x_i \cdot x) + c)$

4. PVC Classification

SVM classifier is used herein to classify samples, Hyperplane way to support vector machine classifier described by the following sample into two categories [6, 7, 8]

$$w^T x + b = 0 \quad (9)$$

Where $x = [x_1, x_2, \dots, x_n]^T$ is an n-dimensional input vector, b is a scalar, $w = [w_1, w_2, \dots, w_n]^T$ is a weight vector, constrained optimization problem by the following decisions:

$$w = \arg \min_w \frac{1}{2} \|w\|^2 + c \sum_{i=1}^N \xi(i) \quad (10)$$

And subject to:

$$y(i)[w^T x(i) + b] \geq 1 - \xi(i), i = 1, 2, 3, \dots, N \quad (11)$$

N represents number of training samples, $x(i)$ represents another training samples of the first i vector, $y(i) \in \{+1, -1\}$ is the class label of $x(i)$, $C > 0$ is a scalar, $\xi(i) > 0$ is a slack variable, and its role is to relax the formula (10) a hard binding. Lagrangian method can be used to solve the problem (10) and (11) of constrained optimization problems, linear support vector machine discriminant function is:

$$f(x) = w^T x + b = \sum_{i=1}^L \alpha(i) y_s(i) S^T(i) x + b \quad (12)$$

$\alpha(i) (i = 1, 2, 3, \dots, L)$ is a positive Lagrange multiplier, L is $S(i)$ corresponding support vector machine total number of training samples $\alpha(i)$, $y_s(i)$ is the class standard, Its decision-making function is [9, 10]:

$$y = \text{sgn}[f(x)] \quad (13)$$

Tenfold cross validation was used to assess the performance of the classifier. Feature extraction data is divided into 10 equal-sized "pack" and 10 classifiers trained with different folding, while the other nine classifier training data used for comparison, the ultimate determination of classification performance is to take the average of 10 test results set.

5. Experimental Analysis

The ECG waveform processed is obtained from ECG data through the filter with a lower limit frequency of 1Hz the upper frequency band pass 35Hz, then the ECG are transformed by the five layers of wavelet continuously, and finally by using 10-point moveable window function signal averaging techniques. Its use LADT threshold algorithm after data compression method to detect heart beat rhythm [11]. Beating heart beat detection technology without affecting ECG waveform, ECG feature extraction from the ECG, get two feature sets used for classification using wavelet transform.

With the detection of ECG wavelet transform, wavelet coefficients of variance and entropy can be calculated, R-R ratio can also be deduced by continuous ECG signal band, which formed feature set 1.

ECG wavelet coefficients from the decay found the sharpest 31 attenuation coefficient is very fast[12]. This 31 coefficient has the lowest level and is pre-selected for the local features. Forward selection process applied to the 10 training set respectively, once a proper subset, it will conduct independent testing on the test set. Subset contains an average of nine elements, characterized in that the set 2 of 9 elements [13.14].

These features include two feature sets, as the SVM classifier input, using tenfold cross-classification algorithm, the overall classification accuracy rate of feature set 1 is 88.07%, as shown in Table 2 below:

Table 2. Classification of Feature Set 1(%)

Serial No.	PVC Detection Rate	Normal NSR Detection Rate	Average Detection Rate
106	87.73	89.37	88.55
210	83.97	94.85	89.41
221	89.81	87.63	88.72
233	84.91	86.27	85.59
Average		88.07%	

A similar set of one feature classification experiment, the average recognition accuracy of 77.13%, as shown in Table 3. The feature set 1 and set 2 characterized joint classification experiment, the average recognition accuracy of 93.17%, as shown in Table 4.

Table 3. Classification of Feature Set 2(%)

Serial No.	Detection Rate	Normal NSR Detection Rate	Average Detection Rate
106	76.86	76.17	76.50
210	80.29	75.15	77.72
221	77.69	79.94	78.82
233	74.67	76.25	75.46
Average		77.13%	

Table 4. Classification of Feature Set 1 and 2(%)

Serial No.	PVC Detection Rate	Normal NSR Detection Rate	Average Detection Rate
106	92.45	96.13	94.29
210	90.16	95.95	93.06
221	95.72	93.67	94.69
233	88.89	92.39	90.64
Average		93.17%	

6. Result Analysis

The purpose of this study is to extract the wavelet coefficients in the classification features, thereby improving the classification accuracy of PVC. ECG detection method is not affected by the proposed ECG waveform change. In fact, the key is how to detect the ECG rhythm ECG peaks detected, and therefore, the ECG detection method proposed in this paper can be used for ECG peak detection signal, the ECG signal into it has a significant effect on the ECG characteristics of a further extraction, further compression of ECG data.

The accuracy of the classification only depends on the characteristics of the ECG, but also depends on the classifier. In general, the signal characteristics compared to the high-dimensional, low-dimensional signal features for better classification results are possible and feasible. From the viewpoint of dimension, when compared to the original sequence of construction and wavelet coefficients represent ECG signal characteristics, but did not reduce the number of dimensions. If all the characteristics of wavelet coefficients as ECG, then they would be classified as fully contain the information, so it is difficult to make an ECG signal classification, and sometimes even get worse classification results.

Although in many classification modeling, for example, logic and neural network identification technology is more flexible [15, 16], but the linear discriminant method for ECG signal is fast and efficient in classification and discrimination.

In this study, the two feature sets are extracted. Feature Set 1 is composed of tolerance, entropy and continuous ECG beat R-R wavelet coefficients than the components. PVC-based classification feature set 1 identification number 106,210,221 and 233 data files ECG signal accuracy reached 88.55%, 89.41%, 88.72% and 85.59%, to achieve the overall average classification accuracy of 88.07%; feature set 2 is selected by a two-stage wavelet coefficients composition, PVC-based feature set two identification numbers 106,210,221 and 233 ECG signal data file classification accuracy of 76.50%, 77.72%, 78.82% and 75.46%, to achieve the overall average classification accuracy of 77.13%. Classification results show the results of a classification of the feature set is better than the feature set of classification 2, the data shown in Table 4 also shows that the use of a set of features in combination with the feature set can improve the classification result; the average accuracy of the classification reaches 93.17%.

Further experimental results show that, when dimension is reduced to 34, the classification results are very similar by using a combination of the wavelet coefficients and the feature set.

7. Conclusion

This thesis presents a fusion of SVM and wavelet algorithm PVC ECG classification algorithm. Using wavelet transform to extract two feature sets, analyzing PVC and NSR by SVM, the recognition result can reach 93.17%. This shows that it is possible that ECG signal features can be extracted by a low-dimensional wavelet coefficient. This method improves the accuracy of identification of ECG features.

Acknowledgements

Project Funding: Science and Technology Research in Universities of Guangxi, 2013(2013YB259).

References

- [1] Y. Wang, "Classification and treatment of premature ventricular contractions", *Clinical Focus*, vol. 3, 017, (1999).
- [2] Y. Bo and Z. Yue, "Multi template matching VPBdiscriminant algorithm", *On computer engineering*, vol. 16, no. 34, (2010), pp. 106-114.
- [3] X. Ping, X. Min and L. Yingjun, "Wavelet transform and BP neural network based on ventricular premature beat(PVC) China recognition", *Journal of medical physics*, vol. 27, no. 2, (2010), pp. 1762-1765.
- [4] H. Honghua and G. Zhiguo, "Application of wavelet entropy in ECG denoising processing and R wave detection", *Computer engineering and Application*, vol. 46, no. 9, (2010), pp. 116-119.
- [5] X. Hongyan and H. Minsong, "ECG feature extractionalgorithm", *Chinese Journal of scientific instrument*, vol. 29, no. 11, (2008), pp. 2362-2366.
- [6] Y. Cheng, S. Yujuan and L. Liuqi, "The lifting wavelet ECG denoising and QRS wave recognition based on fast algorithm", *Journal of Jilin University: Engineering science*, vol. 42, no. 4, (2012), pp. 1037-1043.

- [7] T. Liping, Z. Jian and L. Jian, "Support vector machine arrhythmia diagnosis research", based on Laser Journal, vol. 33, no. 4, (2012), pp. 80-82.
- [8] X. Liang and W. Xiaodong, "Improved fast independent component analysis and support vector machineclassification diagnosis ECG", based on Chinese Journal of scientific instrument, vol. 30, no. 12, (2009), pp. 2599-2604.
- [9] P. Xiujuan, Z. Xiaoping and H. Liang, "The fetal ECG extraction of least squares support vector machine", Journal of electronics and information, vol. 31, no. 12, (2009), pp. 2941-2947.
- [10] H. H. Chen, C. Y. Chiang and T. C. Chen, "Analysis and Design of On-sensor ECG Processors for Realtime Detection of Cardiac Anomalies Including VF, VT, and PVC", Journal of Signal Processing Systems, vol. 65, no. 2, (2011), pp. 275-285.
- [11] M. Bsoul, H. Minn and M. Nourani, "Real-time sleep quality assessment using single-lead ECG and multi-stage SVM classifier Engineering in Medicine and Biology Society (EMBC)", 2010 Annual International Conference of the IEEE, IEEE, (2010), pp. 1178-1181.
- [12] D. Shi-fei, X. Li and Z. Hong, "Research and progress of cluster algorithms based on granular computing", International Journal of Digital Content Technology and its Applications, (2010).
- [13] P. Samui, "Application of Least Square Support Vector Machine (LSSVM) for Determination of Evaporation Losses in Reservoirs", Engineering, (2011) April.
- [14] H.-Q. Wang, F.-C. Sun, Y.-N. Cai, L.-G. Ding and N. Chen, "An unbiased LSSVM model for classification and regression", Soft Computing, (2010) February.
- [15] L. Hong Lee, R. Rajkumar and D. Isa, "Automatic folder allocation system using Bayesian-support vector machines hybrid classification approach", Applied Intelligence, (2012) February.
- [16] Y. Xu-bing, C. Song-can and C. Bin, "Proximal supportvector machine using local information", Neurocomputing, (2009).

Author



MR. Huang Dong, (1972.12-), who is from Lingshan(Guangxi), is the member of CCF (E200027021M). He got his postgraduate in Guangxi Normal University, and the academic title of Associate Professor in 2013(Senior Engineer in 2011). His research areas include Intelligent Recognition, Data Mining, and Information System.