Classification of ECG beats Using Cross Wavelet Transform and Support Vector Machines

Neenu Jacob
Dept. of Applied Electronics and Instrumentation Engineering
Rajagiri School of Engineering and Technology
Kochi, India

Liza Annie Joseph
Dept. of Applied Electronics and Instrumentation Engineering
Rajagiri School of Engineering and Technology
Kochi, India

***Abstract*—In this paper, heart beat classification is performed using cross wavelet transform (XWT), and support vector machines(SVM). XWT is used for the analysis and classification of electrocardiogram (ECG) signals. The cross-correlation between two time domain signals gives a measure of similarity between two waveforms. The proposed algorithm uses XWT to analyze ECG data and determine wavelet coherence(WCOH) and wavelet cross spectrum(WCS). WCOH and WCS obtained are used mathematically to determine the parameter(s) for the purpose of classification. SVM is used to classsify the beats based on the parameters calculated from WCOH and WCS. MIT-BIH arrhythmia database is used for evaluation of results. An overall accuracy of 94.8% for SVM based classification and 96.2% for two dimensional SVM based classification was obtained using the proposed method.**

Keywords*-Cross wavelet transform; Support Vector Machine; Heart beat classification; arrhythmia*

I. INTRODUCTION

Electrocardiography (ECG) is the recording of the electrical activity of the heart. This is an interpretation of the electrical activity of heart measured across the thorax or chest using electrodes attached to the skin for a period of time and can be recorded or displayed. ECG records electrical impulses generated by the polarization and depolarization of cardiac tissue and translates into a waveform. This waveform is used to detect presence of any damage to the heart or any abnormality in the electrical conduction system to the heart. The normal waveform of ECG is shown in Figure 1.

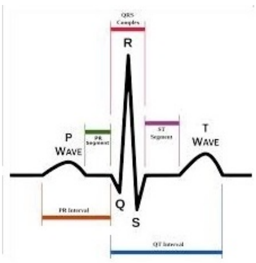

Figure 1. Normal ECG waveform

Cardiac arrhythmias are conditions which leads to abnormal activity or behavior in heart. Certain arrhythmia are life-threatening that can trigger cardiac arrest or even death, like ventricular fibrillation and tachycardia. Detection of arrhythmias, which are not life-threatening, but need medical attention and therapy to avoid deterioration is performed using the proposed method. The classification of heartbeats is an important step in detecting and classifying arrhythmias.

In this paper, we propose subjecting ECG signal to analysis using cross wavelet transform(XWT) and calculating parameter(s) using mathematical formula from wavelet coherence(WCOH) and wavelet cross spectrum(WCS). Then support vector machine(SVM) is used for classification of the normal and arrhythmia ECG signals. The motivation for this comes from a rule of thumb in that arrhythmia heartbeats can be differentiated from normal heartbeats in terms of both morphology and dynamics differences, as depicted in Figure 2. arrhythmia heartbeats are characterized by different abnormal patterns in the ECG waveform shape or missing important components in one heart cycle. Cardiac arrhythmias are associated with and identified by various irregularities in heart rhythm.

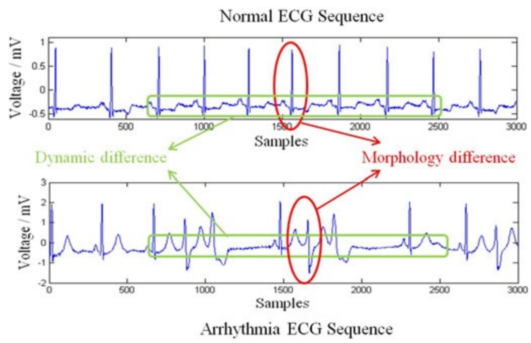

Figure 2. Differences between a normal beat and an arrhythmia beat.

II. OVERVIEW OF CROSS WAVELET TRANSFORM AND SUPPORT VECTOR MACHINES

A. Cross Wavelet Transform(XWT)

Wavelet transform is a linear transform that decomposes a signal into components that appear at different scales. XWT can be used to study the relation between pairs of time-domain signals. The XWT of two time domain siganls $x(n)$ and $y(n)$ is given by

$$W_{XY} = W_X W_Y^* \tag{1}$$

where * denotes complex conjugation. The cross wavelet power is defined as $|W_{XY}|$. The complex argument $arg(W_{XY})$ gives the local relative phase between $x(n)$ and $y(n)$ in time-frequency space. The WCOH of two time series is defined as

$$R_n^2(s) = \frac{\left|S(s^{-1}W_{XY_n}(s))\right|^2}{S(s^{-1}\left|W_{X_n}(s)\right|^2).S(s^{-1}\left|W_{Y_n}(s)\right|^2)} \quad (2)$$

where S is a smoothing operator and s is the scale. WCOH is the localized correlation coefficient in the time-frequency space.

B. Support Vector Machines(SVM)

SVM algorithms are used due to their good generalization capability derived from the structural risk minimization principle. Given a training dataset $\{(x_1, y_1), .., (x_N, y_N)\}$, where $x_i \in R^d$ and $y_i \in \{+1, -1\}$, SVM solves a quadratic optimization problem

$$min_{x,b,\xi_i} \frac{1}{2}\left\|w\right\|^2 + C\sum_{i=1}^{N}\xi_i \quad (3)$$

subject to $y_i(\langle\phi(x_i), w\rangle + b) - 1 + \xi_i \geq 0$

$$\xi_i \geq 0, i = 1, 2, \ldots, N \quad (4)$$

where $\phi(x_i)$ is a nonlinear transformation that maps training data to a higher dimensional space, ξ_i represent the losses and C is a regularization parameter.

By using Lagrange multipliers, the above equation can be written into its dual form, then the problem consists of solving

$$max_{\alpha_i} \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{N}\alpha_i y_i \alpha_j y_j K(x_i, x_j) \quad (5)$$

constrained to $0 \leq \alpha_i \leq C$ and $\sum_{i=1}^{N}\alpha_i y_i = 0$, where α_i are the Lagrange multipliers and $K(x_i, x_j) = \langle\phi(x_i), \phi(x_j)\rangle$ is the kernel function.

After obtaining the Lagrange multipliers, the SVM classification for a new sample x is simply given by

$$y = sgn\left(\sum_{i=1}^{N}\alpha_i y_i K(x_i, x) + b\right) \quad (6)$$

The free parameters of the SVM model γ and C have to be settled a priori.

III. Proposed Method

In the proposed method ECG signal is denoised and R peak is registered. This data is then segmented to individual beats for normalization and then subjected to XWT to get the WCS and WCOH values. Using mathematical formulae, parameter(s) are determined from WCS and WCOH, which are used to classify the beats. The flowchart for the proposed method is shown in Figure 3.

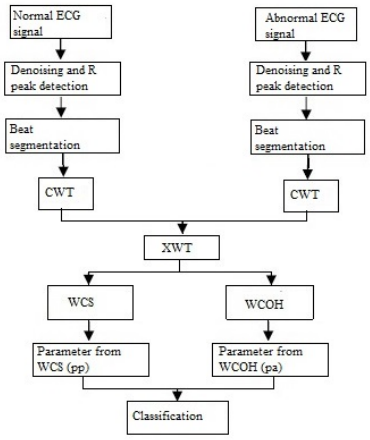

Figure 3. Flowchart of proposed method.

A. Dataset

MIT-BIH arrhythmia database developed as the standard test material for the evaluation of arrhythmia detectors is used. It is regarded as the benchmark database in arrhythmia detection and classification and has been extensively utilized for algorithm validation. The database contains two-lead ambulatory ECG signals from 47 subjects recorded over 48 half-hour period. The signals are digitized at 360 Hz. The header file associated with each record provides the reference annotations for each heartbeat.

B. Denoising Of ECG Data and R Peak Registration

Discrete Wavelet Transform(DWT) based decomposition and selective reconstructions of wavelet coefficients for denoising and QRS detection as developed in [2] is used to increase the efficiency of the algorithm. The ECG signal is decomposed to 10 levels using DWT. The noise is eliminated by identification of the noisy frequency band (D1 and D2) and and rejecting the corresponding coefficients. Similarly the baseline wandering can be eliminated by identification of corresponding frequency band (A10) and rejecting the coefficients. QRS frequency band is identified by selection of the detail coefficients (D4 and D5), because they contain most of the QRS information. This is used for R peak registration

C. Segmentation Of Beats

After identifyng the R peaks, the R-R interval is divided into 2:1 ratio as mentioned in [1]. Each segment consist of x points to the left and $2x$ to points to right of R peak. The number of samples(n) in one beat segment is given by

$$n = \frac{60.m}{H}$$

where m is the sampling rate and H is heart rate. The purpose of normalization is to get segments having equal number of samples for point-to-point correlation analysis. The length of each segment is fixed at 300 samples.

D. Application Of XWT On Beat Segments

Similarity between two waveforms can be measured using cross-correlation. XWT taken over the continuous wavelet transform(CWT) of two signals generates the WCS and WCOH and gives a relationship time-scale space. Due to morphological similarity with that of the QRS complex, Morlet wavelet is used as mother wavelet. 512 scales are considered in this analysis. Parameter extraction formulas are developed in [1] is used for classification of normal and abnormal classess after analysis of ECG beats.

E. WCS and WCOH Based Parameter Extraction

On application of XWT, matrices containing WCS and WCOH between two signals are generated. Equations are developed for feature extraction from WCS and WCOH. The scale range is denoted by $s1$ and $s2$ and over the segment $t1$ and $t2$. The following equations for parameter extraction is given by

$$pp = \sum_{s1}^{s2} \sum_{t1}^{t2} WCS(s,t)$$

$$pa = \sum_{s1}^{s2} \sum_{t1}^{t2} WCOH(s,t)$$

For Lead II the set of parameter (pa,pp) is extracted and this is used for classification described in the next section.

F. Classification

Data classification reduces to a two class problem creating a partition between normal and arrhythmia class. Threshold based classification mentioned in [1] was avoided due to poor accuracy. SVM based classification was used. SVM consists of building an optimal hyperplane that maximizes the separation margin between two different classes. Different pa and pp values for abnormal and normal classes were calculated. These values are used to train SVM. SVM was trained to generate 0 if pa test value belongs to normal class and 1 if pa test value belongs to abnormal class. Similarly SVM was trained to generate 0 if pp test value belongs to normal class and 1 if pp test value belongs to abnormal class.

After training the SVM, it was used to classify the pa and pp values for each beat under study. Label “Class1”is used for pa values and label “Class2”is used for pp values.If the sum of the “Class1”and “Class2”is 0 then the beat is classified as normal. Otherwise the beat is classified as abnormal since it belongs to arrhythmia class.

The generated SVM based classification rule is stated below.
If pa test value is
“normal”mark as “Class1=0”
else “abnormal”mark as “Class1=1”
end if
If pp test value is
“normal”mark as “Class2=0”
else “abnormal”mark as “Class2=1”
end if
If Class1+Class2=0 mark as “Normal”
else mark as “abnormal”
end if
This method gave an accuracy of 94.8%

Two dimensional SVM was also used as an alternative method for classification. In this method pa and pp values were given together to train the SVM. Once training is completed, pa test value and pp test value are given as a pair. If the test value pair is normal then a value 0 is assigned otherwise a value 1 is assigned.

The generated two dimensional SVM based classification rule is stated below.
If test pair value is
“normal”mark as “0”
else “abnormal”mark as “1”
end if
The two dimensional SVM gave an accuracy of 96.2%.

svmtrain.m and svmclassify.m were used for training and classifying of beats.

IV. Experimental Results

The input data for the proposed method was selected from MIT-BIH arrhythmia database. The results were tested on lead II and is described below. Three statistical indices are considered: Accuracy(Acc), Sensitivity(Se) and Specificity(Sp).

The accuracy of the classifier is given by

$$Acc = \frac{N_T - N_E}{N_T} \times 100$$

where the variables N_E and N_T represent the total number of classification errors and beats in the file, respectively.

Table I
Performance Evaluation Metric For Beat Classificaton

Method	No. of beats	TP	FP	TN	FN	Se%	Sp%
SVM	500	371	12	103	14	96.4	89.5
2D SVM	500	381	14	100	5	98.7	87.7

Sensitivity(Se) is the ratio of the number of correctly detected events, true positives (TP), to the total number of events, given by

$$Se = \frac{TP}{TP + FN} \times 100$$

where false negatives (FN) is the number of missed events.

The specificity (Sp) is the ratio of the number of correctly rejected nonevents, true negatives (TN), to the total number of non events, and is given by

$$Sp = \frac{TN}{TN + FP} \times 100$$

where false positives (FP) is the number of falsely detected events.

The performance evaluation metric is given in Table I. Performance evaluation is based on test performed on 500 beat segments extracted from ECG data collected from 43 patients. The sensitivity and specificity were found out to

Table II
CLASSIFICATON ACCURACY OF SVM CLASSIFIER

Method	No. of beats	Se%	Sp%	No. of Correct classifi-cations	No. of Misclas-sifica-tions	Acc%
SVM	500	96.4	89.5	474	26	94.8
2D SVM	500	98.7	87.7	481	19	96.2

be 96.4% and 89.5% for SVM based classification. Classification accuracy was found to be 94.8% for this method and is shown in Table II. For two dimensional SVM based classification a sensitivity of 98.7%, specificity of 87.7% and overall accuracy of 96.2% was obtained.

V. CONCLUSION

In the proposed method the ECG data is preprocessed for removing high frequency noise and baseline wandering as this can affect the performance of the classifier. DWT based method is used for this purpose. The denoised ECG signal is subjected to XWT to obtain WCS and WCOH values. From this parameter(s) are extracted using mathematical formulae and are used for classification using SVM. The proposed method gives an accuracy of 94.8% and 96.2% using SVM based classification and two dimensional SVM based classification respectively based on information measured across a single lead. The computation time required for classification reduces considerably.

The work reported in this paper can be further extended to identify each of the arrhythmias individually. Also additional features can be added during the training of SVM classifier to further improve the accuracy of the proposed method. The result obtained using the proposed method can be extended to 12 lead ECG based classification system and other abnormality issues.

REFERENCES

[1] S.Banerjee,M.Mitra “Application of Cross Wavelet Transform for ECG Pattern Analysis and Classification”, *IEEE transactions on Instrumentation and Measurement*, Vol.63 , No.2, pp.326-333,Feb 2014.

[2] S. Banerjee, R. Gupta, and M. Mitra,“Delineation of ECG characteristic features using multiresolution wavelet analysis method ”*Measurement*, vol. 45, no. 3, pp. 474487, Apr. 2012.

[3] Can Ye, B.V.K. Vijaya Kumar, and Miguel Tavares Coimbra,“Heartbeat Classification Using Morphological and Dynamic Features of ECG Signals ”*IEEE Transactions on Biomedical Engineering*, vol. 59, no. 10, pp. 29302941, Oct. 2012.

[4] S. Banerjee, M. Mitra,“ECG Feature Extraction and Classification of Anteroseptal Myocardial Infarction and Normal Subjects using Discrete Wavelet Transform ”*International Conference on Systems in Medicine and Biology*, 16-18 December 2010, IIT Kharagpur, pp. 55-60, Apr. 2010.

[5] Simon Haykin, *Neural Networks :A Comprehensive Foundation*, 2nd edition, Prentice Hall,2005.

[6] PTB Diagnostic ECG Database Directory, Physiobank Archive Index, PTB Diagnostic ECG Database [Online]. Available: ¡http://physionet.org/physiobank/database¿

[7] MIT-BIH Arrhythmias Database. [Online]. Available: http://www.physionet.org/physiobank/database/mitdb/.

[8] C.Lin, C.Yang,“Heartbeat Classification Using Normalized RR Intervals and Wavelet Features ”*International Symposium on Computer, Consumer and Control*, pp. 650-653, 2014.