Proceedings of 2013 IEEE Conference on Information and Communication Technologies (ICT 2013)

ECG Beat Classification Using Wavelets and SVM

Shameer Faziludeen[1], Sabiq P.V[2]
[1,2]Department of Electronics and Communication Engineering, KMCT College Of Engineering, Calicut ,India.
[1]shameer503@gmail.com, [2]sabiq_pv@rediffmail.com

ABSTRACT

Electrocardiogram (ECG) is one of the most important noninvasive tools for the diagnosis of cardiac arrhythmia. Automatic beat classification in ECG is a topic of continuing research. In this paper, automatic classification of 3 beat types – normal sinus rhythm, premature ventricular contraction and left bundle branch block is implemented. QRS detection is done using the Pan Tompkins algorithm. Wavelet decomposition using daubechies 4 wavelet is done. 25 features are extracted for each beat from wavelet analysis, namely - mean, variance, standard deviation, minimum and maximum of detail coefficients and of approximation coefficients. 3 RR interval features are also extracted for each beat. Beat classification is implemented by using OAO (One Against One) SVM (Support Vector Machine). 3 SVM's are designed and final grouping is done by maximum voting. Novel method of feature selection is introduced. Feature selection for a particular SVM is done based on the beats to be classified by that SVM. ECG signals are obtained from the open source MIT-BIH cardiac arrhythmia database. 6355 beats (2036 LBB, 3865 N, 454 PVC) are used for testing the implementation. Accuracy of 98.46%, 98.47% and 99.92% are obtained for left bundle branch block, normal and premature ventricular contraction beats respectively.

Keywords: ECG beat classification, Support Vector Machine, Wavelets

1. INTRODUCTION

Cardio-vascular disease is one of leading causes of death. Any disturbance in normal rhythm of heart is called cardiac arrhythmia. Electrocardiogram (ECG) is one of the most important noninvasive tools for the diagnosis of cardiac arrhythmia. The electrocardiogram (ECG) signal represents the changes in electrical potential during the cardiac cycle recorded using surface electrodes on the body [1].

The basic ECG wave is shown in Fig.1. ECG wave consists of the P wave, QRS complex, ST and T wave.

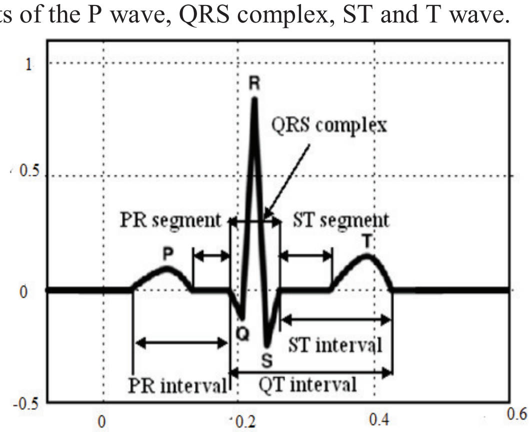

Fig. 1. Basic ECG wave

Arrythmia detection and classification can be done by analyzing ECG signals. Manual approach of ECG analysis is time consuming and requires expertise. Possibility of errors when manual analysis is done by general physician is high. ECG analysis is difficult in a small setup like clinics, health centres, e.t.c. Hence, method for automatic classification with high accuracy is needed. Also, automatic classification will be helpful to paramedical and emergency medical staff in cases where immediate action is to be taken.

In this paper, automatic classification of 3 different types of beats - normal sinus rhythm, premature ventricular contraction and left bundle branch block is implemented. QRS detection is done using the Pan Tompkins algorithm [11,12]. Wavelet decomposition of QRS complex of each beat using daubechies 4 wavelet is done [26]. 25 features are extracted for each beat from wavelet analysis, namely - mean, variance, standard deviation, minimum and maximum of detail coefficients (from level 1 to 4) and of approximation coefficients(level 4). 3 RR interval features are also extracted for each beat [24,25]. Beat classification is implemented by using OAO (One Against One) SVM (Support Vector Machine). 3 SVM's are designed and final grouping is done by maximum voting. Novel method of feature selection is introduced. Feature selection for a particular SVM is done based on the beats to be classified by that SVM. ECG signals are obtained from the MIT-BIH cardiac arrhythmia database [5].

2. QRS DETECTION

QRS detection is needed prior to feature extraction. QRS detection is done using the Pan Tompkins algorithm [11,12]. 3 RR interval features for each beat are obtained directly from QRS detection step as proposed by Tsipouras et al [24,25]. RR interval of a beat is the time interval between R peaks of that beat and the previous beat. The 3 RR interval features extracted for each beat are RR interval of that beat, RR interval of previous beat and RR interval of beat after.

3. WAVELETS

3.1: Introduction

The wavelet transform has emerged over recent years as a powerful time– frequency analysis and signal coding tool favoured for the interrogation of complex nonstationary signals. Both short duration, high frequency and longer duration, lower frequency information can be captured simultaneously using wavelets. Hence the method is particularly useful for the analysis of transients, aperiodicity and other non-stationary signal features where, through the interrogation of the transform, subtle changes in signal morphology may be highlighted over the scales of interest. Another key advantage of wavelet techniques is the variety of wavelet functions available, thus allowing the most appropriate to be chosen for the signal under investigation [23].

2 types of wavelet transform are Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT).

3.2: Discrete Wavelet Transform (DWT)

Accurate and efficient implementation of CWT is provided by DWT. FWT algorithm given by Mallat gives filterbank representation of DWT.

3.2.1: DWT algorithm

This algorithm provides a filterbank representation for computing the wavelet coefficients. Signal is split into approximation and detail coefficients as shown in Fig.2 [6].

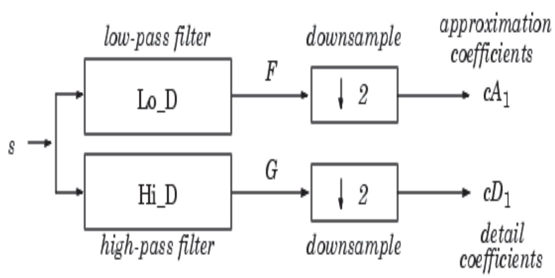

Fig. 2. Wavelet decomposition

The length of each filter is equal to 2n. If N= length (s), the signals F and G are of length N+ 2n - 1, and then the coefficients cA1 and cD1 are of length

$$\text{floor}\left(\frac{N-1}{2} + n\right) \tag{1}$$

This process can be continued for decomposition of signal into many levels [6].

3.2.2: IDWT algorithm

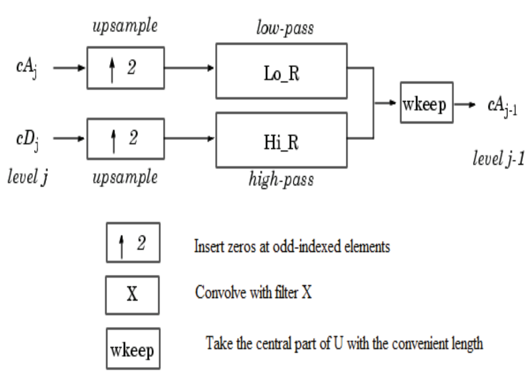

Fig. 3. Wavelet reconstruction

IDWT algorithm reconstructs the original signal from the decomposed signal, i.e., the approximation and detail coefficients. Reconstruction step is given in Fig.3 [6].

3.3: The Daubechies Wavelet Transform

Named after Ingrid Daubechies, the daubechies wavelets are a family of orthogonal wavelets defining a discrete wavelet transform and characterized by a maximal number of vanishing moments for some given support. With each wavelet type of this class, there is a scaling function (also called father wavelet) which generates an orthogonal multi resolution analysis[6].

3.4: Implementation

The daubechies wavelet of order 4 is used. Decomposition upto level 4 of QRS complex of each beat is done. 25 features are extracted for each beat from wavelet analysis, namely - mean, variance, standard deviation, minimum and maximum of detail coefficients (from level 1 to 4) and of approximation coefficients(level 4).

4. SUPPORT VECTOR MACHINES

4.1 Principles

Support Vector Machine (SVM) was introduced by Vapnik. Support Vector Machines (SVM's) are a relatively new learning method used for binary classification. The basic idea is to find a hyperplane which separates the d-dimensional data perfectly into its two classes [20]. SVM is a supervised classification method. Here, a set of known objects is called training set. Each object of the training set consists of a feature vector and a belonging class value. Based on the training data, the learning algorithm extracts a decision function to classify the unknown input data [18].

Let there are l training examples (x_i, y_i) i = 1,.... l , where each example has d inputs ($x_i \in R^d$), and a class label with one of two values ($y_i \in (-1,1)$) [20]. Now, all hyperplanes in R^d are parameterized by a vector (w), and a constant (b), expressed in the equation

$$w.x + b = 0 \tag{2}$$

w is the vector orthogonal to the hyperplane. Given such a hyperplane (w,b) that separates the data, this gives the function

$$f(x) = \text{sgn}(w.x + b) \tag{3}$$

which correctly classifies the training data. However, a given hyperplane represented by (w,b) is equally expressed by all pairs (λw, λb) for $\lambda \in R^+$. The canonical hyperplane is defined to be that which separates the data from the hyperplane by a distance of at least1 [20]. That is, we consider those that satisfy

$$y_i(x_i.w + b) \geq 1 \text{ for all } i \tag{4}$$

To obtain the geometric distance from the hyperplane to a data point, we must normalize by the magnitude of w [20]. This distance is given by

$$d((w,b),x_i) = \frac{yi(xi.w+b)}{\|w\|} \geq \frac{1}{\|w\|} \tag{5}$$

The hyperplane that maximizes the geometric distance to the closest data points is needed. This is accomplished by minimizing ||w|| (subject to the distance constraints) [20].

4.2 Multiclass SVM

Support vector machines are binary classifiers. However, in real cases, data is to be classified into more than 2 classes. This is done by using multiclass SVM. This can be done by using the One Against One(OAO) approach.

The "one against one" strategy, also known as "pairwise coupling", "all pairs" or "round robin", consists in constructing one SVM for each pair of classes. Thus, for a problem with c classes, c(c-1)/2 SVMs are trained to distinguish the samples of one class from the samples of another class. Classification of an unknown pattern is done according to the maximum voting, where each SVM votes for one class [17].

4.3 Implementation

Multi-class SVM is implemented using OAO strategy. Kernel used is linear kernel for all SVM's. Hence, 3 SVM's are to be constructed. Novel method of variable feature selection is used for training the SVM.

4.3.1: Training

SVM1: SVM 1 is to classify between LBB (Left Bundle Branch Block) and Normal(N) beats. For this SVM, the full feature set (28 features) extracted is used. For training the SVM, features of 19 normal beats and 14 LBB beats from MIT-BIH database are used.

SVM2: SVM 2 is to classify between Normal(N) and PVC (Premature Ventricular Contraction) beats. For this SVM, only the 3 RR interval features are used. 27 normal(N) and 7 PVC beats are used for training.

SVM3: SVM 3 is to classify between LBB and PVC (Premature Ventricular Contraction) beats. For this SVM, only the 3 RR interval features are used. 14 LBB and 7 PVC beats are used for training.

4.3.2: Testing

6355 beats(3865 N, 2036 LBB, 454 PVC) are used for testing the OAO implementation.

Classifier is evaluated using the parameters – sensitivity, specificity and accuracy.

$$\text{Sensitivity} = (TP/TP+FN)*100 \quad (6)$$

$$\text{Specificity} = (TN/TN+FP)*100 \quad (7)$$

$$\text{Accuracy}=(TP+TN)/(TP+TN+FP+FN)*100 \quad (8)$$

where TP stands for true positive, TN for true negative, FP for false positive and FN for false negative.

5. RESULTS

Table 1. Beat classification details

Samples under test	Record no:	No: of beats	No: of beats classified	No: of false classified beats
1-64400(LBB)	111	206	203	Nil
87300-195800(349 LBB, 1PVC)	111	350	347	1
196100-649300 (LBB)	111	1490	1487	3
1-649700(N)	115	1952	1946	Nil
1-108000(11PVC)	116	395	392	1
1-649700(1538 N, 442 PVC)	119	1983	1980	92

Table 2. Comparison parameter values for implemented classifier

	No: of beats	Sensitivity (in %)	Specificity (in %)	Accuracy (in %)
LBB	2036	99.85	97.8	98.46
Normal	3865	97.57	99.88	98.47
PVC	454	99.34	99.97	99.92

6. CONCLUSION

ECG beat classifier has been implemented to classify 3 beat types – normal sinus rhythm, left bundle branch block and premature ventricular contraction. Feature extraction is done by using daubechies wavelet and classification by using the OAO SVM strategy. 6355 beats (2036 LBB, 3865 N, 454 PVC) are used for testing the implementation. Accuracy of 98.46%, 98.47% and 99.92% are obtained for left bundle branch block, normal and premature ventricular contraction beats respectively.

REFERENCES

[1] Goldberger A.L., "Clinical Electrocardiography-A simplified approach", 2007, Elsevier.
[2] G.D.Clifford, F.Azuaje, P.E.Mcshary, "Advanced methods and tools for ECG data analysis", 2006, Artech House.
[3] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals", Circulation Vol. 101, Jun.2000, pp.215-220 http://circ.ahajournals.org/cgi/content/full/101/23/e215
[4] Moody GB, Mark RG, "The impact of the MIT-BIH Arrhythmia Database", IEEE Engineering in Medicine and Biology, Vol.20,May-June 2001, pp.45-50.
[5] Massachusetts Institute of Technology, MIT-BIH arrhythmia database, http://www.physionet.org/physiobank/ database/mitdb.
[6] MATLAB Help. http://www.mathworks.in/help/wavelet/
[7] Mallat S, "A Wavelet Tour of Signal Processing – The Sparse Way", 2009, Elsevier.
[8] Rangayyan R.M., "Biomedical Signal Analysis- A case-study approach", 2002, Wiley.
[9] "Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms", Association for the Advancement of Medical Instrumentation, Arlington, VA, 1999.
[10] A. Graps, "An Introduction to Wavelets", IEEE Computational Science and Engineering,Vol.2, 1995.
[11] J. Pan, W.J. Tompkins, "A Realtime QRS Detection Algorithm", IEEE Transactions on Biomedical Engineering", Vol.BME-32, No.3, Mar.1985.
[12] W.J. Tompkins, "Biomedical Digital Signal Processing – C-Language Examples and Laboratory Experiments for the IBM PC", New Delhi, Prentice Hall of India Pvt. Ltd, 2008.
[13] Bert-Uwe Köhler, C. Hennig, R.Orglmeister, "The Principles of Software QRS Detection", IEEE Engineering in Medicine And Biology, Jan/Feb.2002.
[14] Z. Liu, F. Shi, S. Zheng, X. Zhang, "Support Vector Machine-based ECG Beats Automatic Diagnosis System", Journal of Information & Computational Science, Vol.9, pp.3805-3812, 2012.
[15] M. Moavenian, H. Khorrami, "A qualitative comparison of Artificial Neural Networks and Support Vector Machines in ECG arrhythmias classification", Expert Systems with Applications, Vol.37, pp.3088-3093, 2010.
[16] M. Tavassoli, M.M. Ebadzadeh, H.Malek, "Classification of cardiac arrhythmia with respect to ECG and HRV signal by genetic programming", Canadian Journal on Artificial Intelligence, Machine Learning and Pattern Recognition, Vol. 3, No. 1, January 2012.
[17] J. Milgram, M. Cheriet, R.Sabourin, "One Against One or "One Against All": Which One is Better for Handwriting Recognition with SVMs?"
[18] G.G. Sanchez, "Examination of the applicability of Support Vector Machines in the context of ischaemia detection", Diploma Thesis, University of Dresden, Nov.2010.
[19] N. Kohli, N.K. Verma, "Arrhythmia classification using SVM with selected features", International Journal of Engineering, Science and Technology, Vol. 3, No. 8, pp. 122-131, 2011.
[20] "An Introduction to Support Vector Machines", Dustin Boswell, Aug.2002.
[21] S.R. Gunn, "Support Vector Machines for Classification and Regression", University of Southampton, May 1998.
[22] C.J.C. Burges., "A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery 2, pp.121-167, 1998.
[23] P.S. Addison, "Wavelet Transforms and the ECG: A Review", Physiological Measurement, Vol.26, pp.155-199, 2005.
[24] M.G. Tsipouras, D.I. Fotiadis, D.Sideris, "Arrythmia Classification using the RR – Interval Duration Signal", Computers in Cardiology, Vol.29, pp. 485-488, 2002.
[25] M.G. Tsipouras, D.I. Fotiadis, D.Sideris, "An Arrythmia Classification System based on the RR – Interval Signal", Artficial Intelligence in Medicine, Vol.33, pp. 237-250, 2005.
[26] P.D. Chazal, B.G.Celler, R.B.Reilly, "Using Wavelet Coefficients for the Classification of the Electrocardiogram", Proceedings of World Congress on Medical Physics and Biomedical Engineering, Chicago, July 2000.