CLASSIFICATION OF CARDIAC VASCULAR DISEASE FROM ECG SIGNALS FOR ENHANCING MODERN HEALTH CARE SCENARIO

K.Vimala and Dr.V.Kalaivani

Department of Computer Science and Engineering(PG), National Engineering College, India

ABSTRACT

"Why to be in frustration we will do new creation for salvation". Based on these words we grapes your attention towards saving a life of a heart patient with the use of ECG in Public Health Care Center by transmitting ECG signals to nearby hospital server. In this paper we analyze the abnormalities found in the ECG signals by identifying the Normal, Bradycardia Arrhythmia, Tachycardia Arrhythmia and Ischemia signal using the method of Neuro Fuzzy Classifier. Daubechies Wavelet Transforms is used for feature extraction and Adaptive Neuro Fuzzy Inference System (ANFIS) is used for classification. The compression algorithm is performed by using Huffman coding.

KEY WORDS

ECG, DWT, ANFIS, Huffman coding

1. INTRODUCTION

Electrocardiogram (ECG) is used for diagnosing and monitoring the patient body. This analysis system can bring the possibility to record the heart condition at early stage, which the problem is being hard interpretation for non-trained people. Therefore the importance in developing the system that make this interpretation easier for non-trained people and the system could detect the disease with high levels of accuracy because many people who died cause of heart disease showed no outward symptoms. This research is used by the other health care professionals including physicians, nurses, therapists and technicians to bring together knowledge from many technical sources to develop new procedures, or to solve clinical problems. The compression algorithm is used to transmit the ECG signal to hospital server efficiently. The compression minimizes the data amount during the compression. The Cardiac abnormalities can be identified fast and efficiently using ECG signals. In this paper, it is proposed to perform real-time classification of Cardiac Vascular Disease (CVD).Here the features selection for the ECG is done using feature extraction technique. The methods presented here are divided into three pieces of work. Firstly, procedures to identify and annotate of ECG signal for Normal, Bradycardia Arrhythmia, Tachycardia Arrhythmia and Ischemia characteristic. Secondly, a strategy is presented for extracting the features vector for each sample of selected heart disease

using an algorithm that exploits the coefficient derived from Discrete Wavelet Transform. Lastly, this part presented the procedures of classification process using Adaptive Neural Fuzzy Inference System modeling.

2. RELATED WORKS

There are many techniques introduced by the researches to overcome the heart problems in day-today-life. The holter monitor system which records ECG signal continuously in ambulatory condition for a sizable time and system transmit the recorded data to hospital when the recorded period is completed. The feature extraction is done for classification of the heart disease like Normal, Bradycardia Arrhythmia, Tachycardia Arrhythmia and Ischemia characteristics. The compression technique is done for transmitting the signal from the health care services to hospital server.

The ECG analysis technique required the feature extraction and classifier stage. Before feature extraction the noise is removed from the ECG signal. For noise reduction many reduction techniques are proposed by different researchers. An approach to Cardiac Arrhythmia Analysis using Hidden Markov Models[18]. This technique classified by detecting and analyzing QRS complex and determining the R-R intervals to determine the ventricular arrhythmias. The Hidden Markov modeling addresses the problem of detecting low amplitude P waves in typical ambulatory ECG recordings.

 DWT-based feature extraction technique yields superior performance. ECG analysis using wavelet transform method[21] can distinguish the between the QRS wave and P, T wave. This technique also can distinguish noise, baseline drift and artifacts. So it can characterize the signal information analysis very well and suitable to process time-varying biomedical signals.

The two morphological feature extraction methods[28] which are higher-order statistics and hermite basis   functions. They used the support vector machines to compare the feature extraction methods and classification method to evaluate the generalization performance. But the use of higher order models need more computation cost and caused over fitting problem in generalization performance. In term of accuracy, they found that their hierarchical classification method showed better classification performance than the conventional multiclass classification method with despite the loss in accuracy and sensitivities certain classes. The hierarchical classification improved the mean values of sensitivity mean. It agreed that their classification method can distinguish the multiclass heartbeats with the unbalanced data distribution.

A new approach to feature extraction which is Karhunen Lo'eve Transform (KLT) which is an attractive and powerful approach [27] to the feature-extraction and shape representation process. It has the solution if the probability densities of population of pattern vectors of a problem domain are unknown. The problem about this method is it is too sensitive to noisy pattern of ECG signal.

 In ECG training and classification analysis stages, some researchers have tried to maximize the detection level of accuracy in many different ways such as digital signal analysis, Fuzzy Logic

methods, Artificial Neural Network, Genetic Algorithm, Support Vector Machine, Self-Organizing Map, Bayesian and other method with each approach exhibiting its own advantage and disadvantages.

The Neuro-Fuzzy approach[24] to the recognition and classification of the heart rhythms on the basis of ECG waveforms. It uses the new approach of heart beat recognition. The resolution for the problem of less sensitivity to the morphological variation of the ECG is brought here. It combining two techniques which are characterization of the QRS complex of ECG by Hermite polynomials and using the coefficients of Hermite kernel expansion as the features of the process and the application of the modified neuro-fuzzy TSK network for ECG pattern recognition and classification.

3. METHODOLOGY

This system provides useful technique for fast identification and classification of the heart disease.
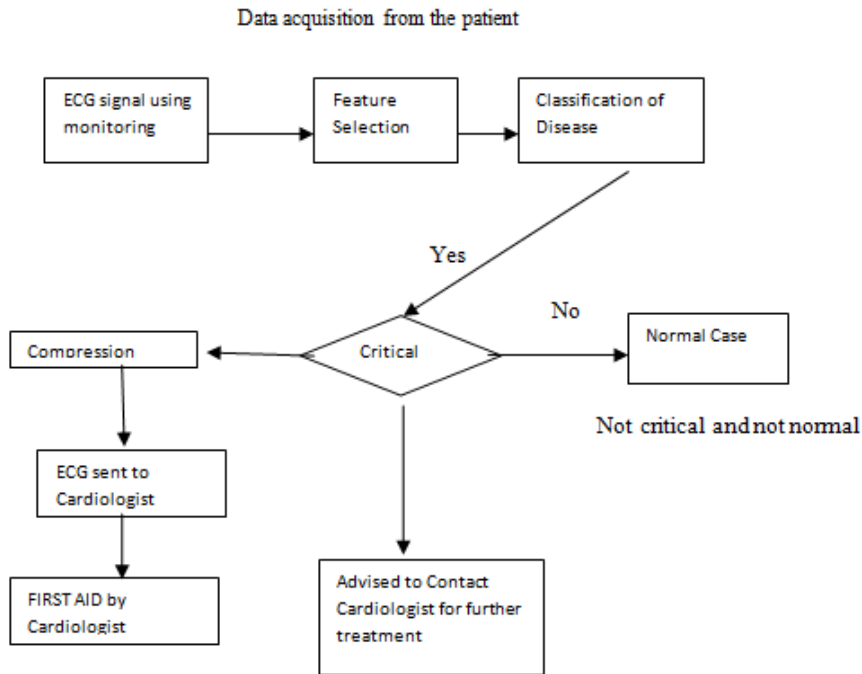

Figure 1: Cardiovascular Abnormality Diagnosis Process

This paper provides the detection of cardiovascular disease in ECG signal with the help of health care centers in rural areas. In figure 1 the signal is recorded from the patient's body and it is monitored by the physicians. The signal is then classified by the physicians by applying Neuro-Fuzzy classifier technique. This shows the result as whether the patient's is in the normal, mild or severe condition. If the patient's condition is critical, then the signal is transmitted to the hospital server by compressing the ECG signals using Huffman coding. For the severe cases physician will provide a tablet that could maintain the heart beat as normal for 6 hours which would help him to reach the Cardiac hospital. Then the patient can undergo for medical treatment by the cardiologist. If the patient's condition is mild then they are advised to contact the cardiologist.

3.1. DWT TECHNIQUE FOR FEATURE EXTRACTION

Feature extraction was conducted by applying wavelet techniques to patient data, thus providing ECG characteristic point detection capabilities.

The input selection of feature extraction methods applied in this thesis has to select well to make sure which components of a input gives the best representation of the given pattern of ECG signals. The detail wavelet coefficients contains some amount of significant informations of the signal, the detail wavelet coefficients of ECG signal of each subject were computed. The procedures of DWT implementation is describe as follow in Figure 2:
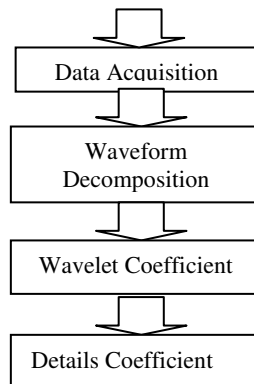

figure 2: feature extraction techniques

2.2. FEATURES EXTRACTION PROCEDURES

Selection of appropriate wavelet and the number of decomposition level is very important in DWT. The levels are chosen based on the parts of the signal that correlated with the frequencies required for classification . The general wavelet decomposition of DWT procedure involves three steps. The result of decomposed signal will shows the important details and approximation coefficients, which represents the original signal. The basic procedure follows the steps described below

- The wavelet type is chosen.
- The wavelet name is chosen
- Number of levels required for decomposition are chosen.

The DWT wavelet types have been chosen in the features extraction and the ECG signals were decomposed into time-frequency representations using single-level one-dimensional wavelet decomposition. The wavelet names of Daubechies wavelet filters db4 have been choosing and the number of decomposition levels was chosen to be 5. Result of decomposition, the ECG signals were decomposed into the details coefficients D1-D5 and one final approximation coefficient, A5.

The results of applying the Daubechies wavelet of order 4(db4) which is more suitable to detect changes of ECG signal is evaluated.

2.3. COEFFICIENTS EXTRACTION

The computed details and approximation wavelet coefficients of the ECG signal were used as the features vector representing the signals.

In this study, from the original intervals of ECG signal, five standard measures parameters are used. A signal of 75 discrete data was selected as considered ECG signals data. For each ECG signals, the detail wavelet coefficients of fourth level (75 coefficients) were computed. To reduce the dimensionality of feature vectors, the set of the wavelet coefficients were used. The following features are used to represent the time and frequency distribution of the ECG signals:

1. Energy of the wavelet of each ECG signals sample.
2. Maximum of the wavelet coefficients of each ECG signals sample.
3. Minimum of the wavelet coefficients of each ECG signals sample.
4. Mean of the wavelet coefficients of each ECG signals sample.
5. Standard deviation of the wavelet coefficients of the each ECG signal sample.

2.4. CLASSIFICATION USING NEURO FUZZY

Decision making of classification was performed in two stages: selection of coefficients computing by DWT and the ANFIS classifiers. Four types of ECG beats (Normal, Tachycardia Arrhythmia, Bradycardia Arrhythmia, and Ischemia) obtained from the PhysioBank databases will be classified by ANFIS classifiers.

2.5. ANFIS IMPLEMENTATION IN CLASSIFYING HEART DISEASE

The classification was performed using the ANFIS in Fuzzy Logic Toolbox. ANFIS were trained with the back propagation gradient decent method in combination with the least squares method. The block of featured processed in ANFIS were shown in Figure 3.
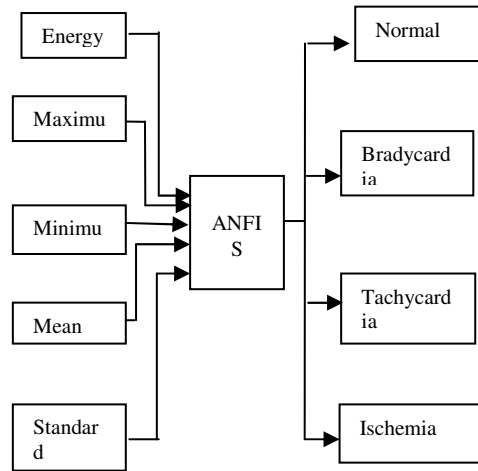
Figure 3: Block Diagram of Heart Disease Classification through ANFIS

The featured vector that are being computing from the DWT coefficient which are energy, Maximum, Minimum, Mean and Standard Deviation were defined as extracted features for ANFIS inputs and Normal, Bradycardia, Tachycardia Arrhythmia and Ischemia are defined as ANFIS outputs.

2.6. SKELETON OF ECG DATA

The ECG signals are divided into two parts
- crucial part
- plain part

The QRS complex waves and P and T waves are the important part of the cardiology system to identify the arrhythmia disease. The ECG signals are divided into smaller blocks and each block is encoded as an independent entity. The block length is decided according to the sampling rate as (1) shows

$$MB = t_{truncn}(f_{unit}S, [\log_2(\frac{S}{c})]) \qquad (1)$$

$$\text{Where} \quad t_{trunc}(x,n) = \frac{[x2^n + 0.5]}{2^n}$$

where *MB* represents the block length, *S* is the ECG sampling rate, *c* is the programmable constant between 101 and 200, f_{unit} is unit block size, and t_{trunc} is the rounding-truncation function with $2n$ In this work, the unit block size is selected in 0.05 s, which is the half duration of the QRS complex duration. It is a suitable period to detect the change the ECG signal precisely. After block division, the QLV of the block is calculated. For normal ECG signals, the QRS complex part can be regarded as a typical representative signal with high standard deviation $(Sta = \sqrt{\sum_{I=0}^{N_B-1}(xi - \bar{x})^2 / N_B})$ in comparison with the plain part [11].

The complex block with high Sta has more crucial information than the plain block with low Sta.

68

However, the Sta requires the complex calculations such as square root ($\sqrt{x}$) and squaring ($x2$). Therefore, the mean deviation (Mean) value is proposed to determine the QLV instead of the sta. The mean is defined as follows:

$$Mean \quad = \quad \frac{\sum\limits_{i=0}^{N_B - 1} |\, x_i \, - \, \bar{x} \,|}{N_B}$$
(2)

2.7. LOSSLESS COMPRESSION: HUFFMAN CODING

The compression was performed when the patient was in the severe state and the patient needs some medical treatment. In that case the patient ECG data is transferred to the doctor on the other end by compressing the signal.

Algorithm for Huffman coding

1. Compute the probability of each data.
2. Sort the set of data in ASCENDING order.
3. Create a new node where the left child is the lowest in the sorted list and the right is the second lowest in the sorted list.
4. Chop-off those two elements in the sorted list as they are now part of one node and add the probabilities. The result is the probability for the new node.
5. Perform insertion sort on the list with the new node.
6. REPREAT STEPS 3,4,5 UNTIL you only have 1 node left.

4. ECG SIGNALS DATASET

The number of data samples that represent each feature for Class 1,Class 2, Class 3 and Class 4 are 20 subjects, 15 subjects, 20 subjects and 15 subjects respectively. The rule based was created according the   expert knowledge using MATLAB rule base editor. Based on the three membership function (small, medium, large) that being used in this project, the number of rule based created by equation:

$a\,^\wedge b$ =c where; a is linguistic variables, b is number of input nodes, c is number of rules output.

It is used to generate the rules. There are 5 input nodes for ANFIS structure with 3 inputs of membership functions that processed by 243 rules to identify the desired output of heart disease.

5. RESULT AND PERFORMANCE ANALYSIS

The result and performance of feature extraction and classification was trained with the back propagation gradient descent method in combination with the least squares method has been made. The algorithm is composed of a two passes, such that forward pass and a backward pass.

- The least squares method (forward pass) - to optimize the consequent parameters.
- The gradient descent method (backward pass) - to adjust optimally the premise parameters.

The paper demonstrates the wavelet coefficients for the features which represents the ECG signals and the ANFIS trained on the input patterns to one of four classes.

In this study, training and test sets were formed by 171 data train the samples. The 100 data samples were used for training and 71 data samples were for testing. The training dataset was used to train the ANFIS, to verify the accuracy and the effectiveness of the trained ANFIS model for the detection of heart disease patients.

Table 1: Statistic of correct and incorrect Heart Disease Classification

Heart disease type	Class	Correct classified	Misclassified
Normal	1	43	1
Bradycardia Arrhythmia	2	40	2
Tachycadia Arrhythmia	3	44	0
Ischemia	4	41	1
Total		168	4

Table 1 and Figure4 show the correct classified and misclassified data samples of heart disease for each class. 43 samples from Class 1 were classified correctly and 1 data sample is incorrect classified. There are 40 samples out of 42 data samples of Class 2 are classified correctly and 41 samples from 42 samples from 42 data samples from class 4 were correctly classified. For Class 3, all of their 44 samples were classified correctly. The ANFIS misclassified 4 samples out of 172 data samples.
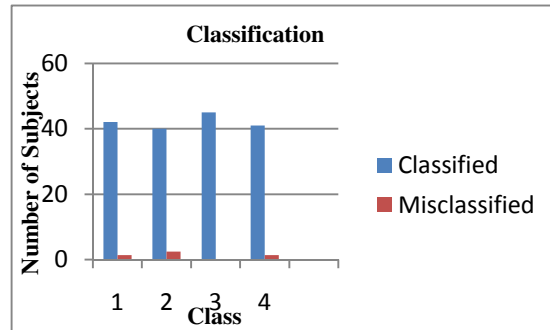

Figure 4 : Statistic of Heart Disease Classification

The MIT-BIH Arrhythmia Database is used to evaluate the performance [13]. The sampling rate and the resolution are 360 samples/s and 12 bits, respectively. In addition, the Gaussian white noise source is injected from -10 to 20 dB SNR for the noise stress test. The test vector is produced by injection of the noise source [14] with MIT-BIH record 100 [13]. The performances of compression coding are evaluated by the CR and the compression error rate of the percentage root mean square difference (PRD). And the R-peak detection performance is evaluated by the sensitivity (Se) and the positive productivity *(+P)*.

6. CONCLUSION

This paper indicate that by using DWT and ANFIS, the classification of Normal, Bardycardia Arrhythmia, Tachycardia Arrhythmia and Ischemia signals can be classified; therefore the primary objective of this study is achieved. The simulation results show the class of heart disease is well predicted using DWT and ANFIS system and the system working well since it achieve the 97.68% of classification accuracy rate. This result indicates that it has some potential and had been found to be successful in heart disease detection. The proposed QLV delineate the ECG signal, and it supports both the flows to achieve better performance with low computation complexity. By applying the QLV, the overall CR improves while maintaining the small compression error rate, and high-accuracy performance can achieve at the segmentation and the R-peak detection stage. The performance is evaluated by using the MIT/BIH Arrhythmia Database, and the noise robust test is also performed for the reliability of the algorithm.

7. REFERENCES

[1]   D. Jabaudon , J. Sztajzel, K. Sievert, T. Landis, and   R.   Sztajzel,   "Usefulness of ambulatory 7-day ECG monitoring for   the detection of atrial f ibrillation   and   flutter after acute stroke and transient ischemic attack,"   Stroke,   J.Amer. Heart Assoc., vol.35,pp.1647-1651, May 2004.

[2]   H.Kim, S. Choi, and H-j.Yoo, "A low power 16-bit RISC with lossless compression accelerator for body   sensor   network system," in Proc. IEEE ASSCC, Nov.2006, pp. 207-210.

[3]   S.-J. Song, N. Cho, S.Kim, J. Yoo, and H.-J.Yoo ,"A 2 Mb/s wideband pulse transceiver with direct-coupled interface for Human body communications," in Proc. IEEE ISSCC, Feb.2006,pp. 2278-2287.

[4]   .Zigel, A.Cohen, and A. Katz, "The weighted diagnostic distortion (WDD) measure for ECG signal compression," IEEE Trans. Biomed.Eng., vol.47.no.11,pp. 1422-1430, Nov. 2000.

[5]   T.A. Welch, "A technique for high-performance data compression," Computer, vol. 17, no.6, pp. 8-19, Jun. 1984.

[6]   Health   Informatics.   Standard   Communication   protocol.   Computer   assited Electrocardiography, British-Adopted European Standard BS EN1064,2005.

[7]   D.A. Dipersio and R.C. Barr, "Evaluation of the fan method of adaptive sampling on human electrocardiograms," Med. Biol. Eng. Comput., vol.23, pp. 401-410, Sep. 1985.

[8]   J.P. Abenstein and W.J. Tompkins, "A new data reduction algorithm for real time ECG analysis," IEEE Trans. Biomed.Eng.,vol.BME-29,no.1, pp. 43-48,Apr.1982.

[9]   W.C. Muller, "Arrhythmia detection program for an ambulatory ECG monitor, " Biomed. Sci. Instrum., vol.14, pp. 81-85, 1978.

[10] M.L. Hilton, "Wavelet and wavelet packet compression of electrocardiograms," IEEE Trans. Biomed. Eng., vol. 44, no. 5, pp. 394-402, May 1997.

[11] R.S.H. Istepanian and A.A.Petrosian, "Optimal Zonal Wavelet-based ECG data compression for mobile telecardiology sys," IEEE Trans. Inf. Technol. Biomed., vol. 4,no. 3,pp.200-211, Sep .2000.

[12] B.S. Kim, S.K. Yoo, and M.H Lee,"Wavelet-based low-delay ECG compression algorithm for continous ECG Transmission ,''IEEE Trans. Inf .Technol. Biomed., vol. 10 ,no. 1,pp.77-83, Jan .2006.

[13] H. Kim,Y. Kim, and H-J. Yoo," A low cost quardratic level ECG Compression algorithm and its hardware optimization for Body sensor network system," in Proc. IEEE EMBC, Aug. 2008, pp. 5490-5493.

[14] (1979).[online].Available: http://www.physionet.org/physiobank/database/mitdb/

[15] (1992).[online].Available: http://www.physionet.org/physiobank/database/nstdb/.

[16] Bousselijot, R. and Kreiseler, K. 1998, "ECG Signal Analysis by Pattern Comparison," Journal Computers in Cardiology, Vol. 25.

[17] Cheng, W. T. and Chan, K. L. 1998,"Classification of Electrocardiogram Using Hidden Markov Models", Proceedings of the 20th Annual conference of the IEEE Engineering in Medicine and Biology Society, vol. 20,no 1.

[18] Castro, B., Kogan, D. and Geva, A.B. 2000, "ECG Feature Extraction Using Optimal Mother Wavelet," Proceeding IEEE Convention of the Electrical and Electronic Engineers, Tel-Aviv, Isreal,pp. 346-350.

[19] Douglas, A.C., Stem, R.M., Cano, G.G. and Briller,S.A. 1990, 'An Approach to Cardiac Arrhythmia Analysis Using Hidden Markov Models', IEEE Transactions On Biomedical Engineering, vol. 37, no.9.

[20] Graja, S. and Boucher, J.M. 2005, "Hidden Markov Tree Model Applied to ECG Delineation', IEEE Transactions on Instrumentation and Measurement, vol. 54,no.6

[21] Hu, Y.H., Palreddy, S. and Tompkins, W. 1997, "A patient Adaptable ECG Beat Classifier Using A Mixture of Experts Approach', IEEE Transactions Biomedical Engineering, vol 44, pp.891-900.

[22] Li, C., Zheng C. and Tai, C. 1995, 'Detection of ECG Characteristic Points Using Wavelet Transforms', IEEE Transactions On Biomedical Engineering, vol. 42, n0.1.

[23] Lagerholm, M., Peterson, C., Braccini, G., Edendrant , L. and Sornmo, L. 2000, 'Cludtering ECG Complexes Using Hermite Functions and Self-Organizing Maps', IEEE Transactions On Biomedical Engineering, vol.47, no.7.

[24] Lin, K and Chang, W.H. 1989, "QRS Feature Extraction Using Linear Prediction,' In IEEE Transactions On Biomedical Engineering ,vol.36, no.10,pp. 1050-1055.

[25] Linh, T.H., Osowsaki, S. and Stodolski, M.2003, "On-Line Heart Beat Recoginition Using Hermite Polynomials And Neuro -fuzzy Network', IEEE Transcations On Biomedical Engineering, vol.48, no.11.

[26] Hyejung Kim, Refer Firat Yazicioglu, Patrick Merken,Chirs Van Hoof and Hoi-Jun Yoo, "ECG signal Compression and Classification Algorihm with Qua Level Vector for ECG Holter System", IEEE Trans. on IT In Biomedicine, vol.14.,no. 1,Jan 2010.

[27] Jager F.2002, 'Feature Extraction and Shape Representation of Ambulatory Electrocardiogram Using the Karhunen-Lo'eve Transform,' Electrotechnical reviewLjubljana, Slovenija.

[28] Papalouks, C., Fotiadis, D.I., Likas, A. and Michalis, L.K. 2003, 'Automated Methods for Ischemia Detection in Long-Duration ECGs,' Cardiovascular Reviews and Reports, vol.24, no.6,pp. 313-319.