Features Extraction of ECG signal for Detection of Cardiac Arrhythmias

P.D.Khandait
Department of Etrx.
K.D.K.C.E.,Nagpur,
Maharashtra, India

N.G. Bawane
Principal,
SBJIET, Nagpur
Maharashtra, India

S.S.Limaye
Principal
JIT,Nagpur,
Maharashtra,India

ABSTRACT
Electrocardiogram (ECG) is the record of the heart muscle electric impulses. Received and processed ECG signal could be analyzed, and results could be used for detection and diagnostics of cardiovascular diseases (CVD).One of the important cardiovascular diseases is arrhythmia.This paper deals with improved ECG signal features Extraction using Wavelet Transform Techniques which may be employed for Arrhythmia detection. This improvement is based on suitable choice of features in evaluating and predicting life threatening Ventricular Arrhythmia. Analyzing electrocardiographic signals (ECG) includes not only inspection of P, QRS and T waves, but also the causal relations they have and the temporal sequences they build within long observation periods. Wavelet-transform is used for effective feature extraction which may be considered for the classifier model. In a first step, QRS complexes are detected. Then, each QRS is delineated by detecting and identifying the peaks of the individual waves, as well as the complex onset and end. Finally, the determination of P and T wave peaks, onsets and ends is performed. Analysis is carried out using MATLAB Software. We evaluated the algorithm on MIT-BIH Arrhythmia Database which is manually annotated and developed for validation purposes. Features based on the ECG waveform shape and heart beat intervals may be used as inputs to the classifiers. A correct beat classification accuracy of 98.17% is achieved which is a significant improvement.

General Terms
Biomedical Signal analysis, bioelectric potentials, cardiovascular diseases.

Keywords
ECG, Wavelet, QRS Complex, median filter, Cardiac Arrhythmia

1. INTRODUCTION
The Electrocardiogram (ECG) is the record of variation of the biopotential signal of the human heartbeats. Electrodes are placed on the users skin to detect the bioelectric potentials given off by the heart that reach the skins surface. The ECG detection which shows the information of the heart and cardiovascular condition is essential to enhance the patient living quality and appropriate treatment. It is valuable and an important tool in diagnosing the condition of the heart diseases. Each individual heartbeat in the cardiac cycle of the recorded ECG waveform shows the time evolution of the heart's electrical activity, which is made of distinct electrical depolarization–repolarization patterns of the heart. Any disorder of heart rate or rhythm, or change in the morphological pattern, is an indication of an arrhythmia, which could be detected manually by the Medical Practitioner. Manual observation for analyzing the recorded ECG waveform takes longer time for the decision making. Hence Artificial Intelligence (A.I.) based detection & classification system is employed [1]

Arrhythmia can be broadly classified into many categories including:
(i) Left bundle branch block (LBBB).
(ii) Normal sinus rhythm (NSR).
(iii) Pre-ventricular contraction (PVC).
(iv) Atrial premature contraction (APC)
(v) Supraventricular tachycardia (SVT)
(vi) Ventricular tachycardia (VT)
(vii) Atrial fibrillation (AF).
(viii)Ventricular fibrillation (VF).
(ix) Complete heart block (CHB) etc.

Out of these arrhythmias, Premature ventricular contraction (PVC) arrhythmia, also known as ventricular premature beat (VPB) or extra systole, result from irritated ectopic foci in the ventricular area of the heart. These foci cause premature contractions of the ventricles that are independent of the pace set by the sinoatrial node. Many studies have shown that PVCs, when associated with myocardial infarction, can be linked to mortality. Consequently, their immediate detection and treatment is essential for patients with heart disease. So is the focus of attention in this paper. Real-time automated ECG analysis in clinical settings is of great assistance to clinicians in detecting cardiac arrhythmias, which often arise as a consequence of a cardiac disease, and may be life-threatening and require immediate therapy. However, automated classification of ECG beats is a challenging problem as the morphological and temporal characteristics of ECG signals show significant variations for different patients and under different temporal and physical conditions [2]. For efficient automatic detection and classification of ECG heartbeat patterns robust feature extraction method is essential.QRS detection is one of the fundamental issues in the analysis of Electrocardiographic signal. The QRS complex consists of three characteristic points within one cardiac cycle denoted as Q, R and S. The QRS complex is considered as the most striking waveform of the electrocardiogram and hence used as a starting point for further analysis or compression schemes. The detection of a QRS complex seems not to be a very difficult problem. However, in case of noisy or pathological signals or in case of strong amplitude level variations, the detection quality and accuracy may decrease significantly. Numerous QRS detection algorithms such as derivative based algorithms [3-6], wavelet transform [7], Filtering Techniques [8] artificial neural networks [9-11], genetic algorithms [12], syntactic methods [13], Hilbert transform [14], Markov models [15] etc. are reported in literature. Kohler et al [16] described and compared the performance of all these QRS detectors. Recently few other methods based on pattern recognition [17], moving- averaging [18] etc are proposed for the detection of QRS complex. Once the position of the QRS complex is obtained, the location of other components of ECG like P, T waves and ST segment etc. are found relative to the position of QRS, in order to analyze the complete cardiac period. The performance of ECG pattern classification strongly depends on the characterization power of the features extracted from the ECG data and the design of the classifier. Due to its time–frequency localization properties, the wavelet transform is an efficient tool for analyzing nonstationary ECG signals .The wavelet transform can be used to decompose an ECG signal according to scale, thus allowing separation of the relevant ECG waveform morphology descriptors from the noise, interference, baseline drift, and amplitude variation of the original signal. Several researchers have previously used the wavelet transform coefficients at the appropriate scales as morphological feature vectors rather than the original signal time series and achieved good classification performance [19][20]. Accordingly, in the current paper, the proposed feature extraction technique employs the suitable wavelet transform in order to effectively extract the morphological and temporal information from ECG data and the extracted features are given to classifier. The overview of the proposed system is shown in Fig. 1.

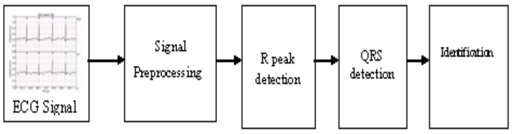


Fig.1 Proposed System for Feature Extraction

The main focus of the paper is to utilize wavelet analysis methods for extracting various features like R peak, QRS detection etc from ECG using wavelet analysis technique. Next section deals with ECG data analysis and fundamentals of wavelet analysis.

2. ECG DATA PROCESSING

2.1 ECG Data

In this study, the MIT/BIH arrhythmia database is used for training and performance evaluation of the proposed ECG classifier. The database contains 48 records, each containing two-channel ECG signals for 30 min duration selected from 24-hr recordings of 47 individuals. Continuous ECG signals are band pass-filtered at 0.1–100 Hz and then digitized at 360 Hz. The database contains annotation for both timing information and beat class information verified by independent experts. The first 20 records (numbered in the range of 100–124), which include representative samples of routine clinical recordings, are used to select representative beats to be included in the common training data. The remaining 24 records (numbered in the range of 200–234) contain ventricular, junctional, and supraventricular arrhythmias. It is recommended that each ECG beat be classified into the following three heartbeat types: N (beats originating in the sinus mode), V (ventricular ectopic beats (VEBs/PVC)), and Other (unclassified beats). For all records, we used the modified-lead II signals and utilized the labels to locate beats in ECG data. Fig. 2 shows the sample ECG taken from the record number 208 of MIT-BIH arrhythmia database and the normal ECG beats showing P wave, QRS complex and T wave. In Fig.2(a) label of X-axis is number of samples and Y-axis indicates the amplitude in mV.

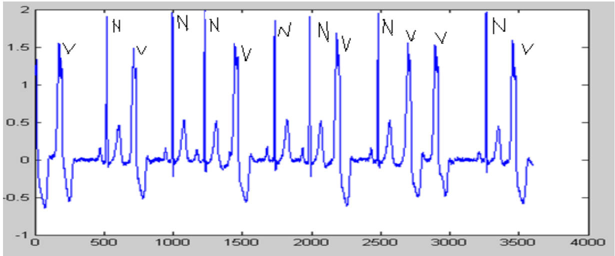


Fig. 2 (a) Sample beats from record No 208

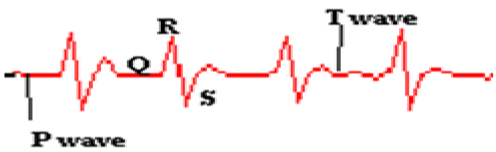


Fig. 2 (b) Normal beats

2.2 Preprocessing of ECG signals

ECG signal mainly contains different types of noises, like frequency interference, baseline drift, electrode contact noise, polarization noise, muscle noise, the internal amplifier noise and motor artifacts. Artifacts are the noise induced to ECG signals that result from movements of electrodes. One of the common problems in ECG signal processing is baseline wander removal. Removal of baseline wander is therefore required in the analysis of the ECG signal to minimize the changes in beat morphology. Respiration and electrode impedance changes due to respiration are important sources of baseline wander in most types of ECG recordings. The frequency content of the baseline wander is usually in a range well below 0.5Hz. This baseline drift can be eliminated without changing or disturbing the characteristics of the waveform. We use the median filters (200-ms and 600-ms) [21] to eliminate baseline drift of ECG signal. The process is as follows-

a) The original ECG signal is processed with a median filter of 200-ms width to remove QRS complexes and P waves.
b) The resulting signal is then processed with a median filter of 600-ms width to remove T waves. The signal resulting from the second filter operation contains the baseline of the ECG signal.
c) By subtracting the filtered signal from the original signal, a signal with baseline drift elimination can be obtained.

The sample beats before and after removal of baseline wander from record No. 106 of MITBIH Arrhythmia database are shown in Fig. 3, where X-axis represents number of samples and Y-axis indicates the amplitude of signal in mV.

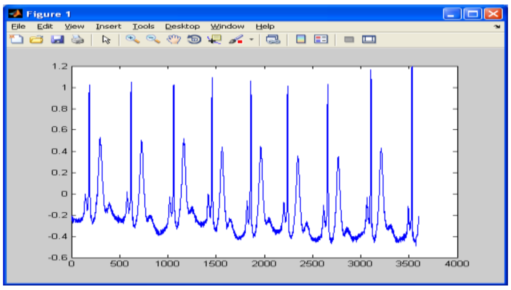


Fig. 3(a) Sample beats from record No. 106

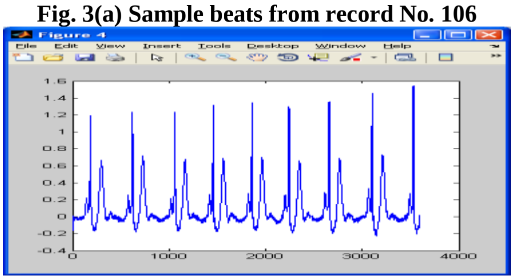


Fig. 3(b) Beats after removal of baseline wander

2.3 Wavelet Technique

In contrast to conventional techniques, the wavelet transform (WT) provides a new dimension to signal processing and event detection. Due to its time-frequency localization properties, the wavelet transform is an efficient tool for analyzing non-stationary ECG signals. The wavelet technique (WT) provides a description of a signal in a timescale domain, analogous to a time-frequency domain, allowing the representation of temporal features at multiple resolutions. This is achieved by the decomposition of the signal over dilated (scale) and translated (time) versions of a prototype wavelet. In its continuous form, the WT of a signal $x(t)\epsilon L^2$ **R** is defined as below-

$$CWT_x(b) = \int_{-\infty}^{+\infty} x(t)\frac{1}{\sqrt{a}}\psi^*\left(\frac{t-b}{a}\right)dt \qquad (1)$$

where $\psi(t)$ is the prototype (or mother) wavelet ($\psi\epsilon L^2$ **R**), a the scale factor ($a\epsilon R^+$), b the translation factor ($b\epsilon R$), CWT the continuous wavelet transform operator and * the complex conjugate operator. The wavelet transform in equation (1), can be rewritten as a convolution product between the signal and the scaled wavelets and can also be interpreted as a filtering of the signal by band pass filters whose center frequencies and bandwidths depend on the scaling factor. High scales translate into long, slow wavelets equivalent to narrow, low-frequency filters, while lower scales produce shorter, faster wavelets equivalent to wider, higher-frequency filters. With these properties, WT achieves an ideal balance of time and frequency resolution: slow trends are represented with a high frequency resolution and a low time resolution, while fast components are well defined in time but less in frequency. Such inherent multiresolution capabilities make WTs very effective at detecting and representing singularities, and WTs have been applied in many occasions to ECG analysis. The ECG signals are considered as representative signals of cardiac physiology, useful in diagnosing cardiac disorders. The most complete way to display this information is to perform spectral analysis. The ECG signal, consisting of many data points, can be compressed into a few parameters by the WT. These parameters characterize the behavior of the ECG signal and they can be used for recognition and diagnostic purposes. The WT can be thought of as an extension of the classic Fourier transform, except that, instead of working on a single scale (time or frequency), it works on a multi-scale basis. This multi-scale feature of the WT allows the decomposition of a signal into a number of scales, each scale representing a particular coarseness of the signal under study. The procedure of multiresolution decomposition of a signal x[n] is schematically shown in Fig.4. Each stage of this scheme consists of two digital filters and two down samplers by 2. The first filter, g(n) is the discrete mother wavelet, high pass in nature, and the second, h(n) is its mirror version, low-pass in nature. The down sampled outputs of first high-pass and low-pass filters provide the detail d_1, and the approximation a_1, respectively. The first approximation, a_1 is further decomposed and this process is continued as shown in Fig.4.

The reconstruction process is the reverse of decomposition, where the approximation and detail coefficients at every level are up-sampled by 2 and passed through low-pass g(n) and high pass h(n) synthesis filters and finally added as shown in Fig. 5. The same number of levels is taken as in the case of decomposition.

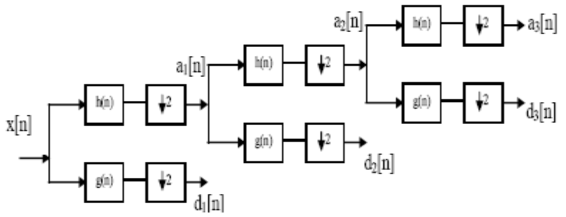


Fig.4 Three Level Wavelet decomposition tree

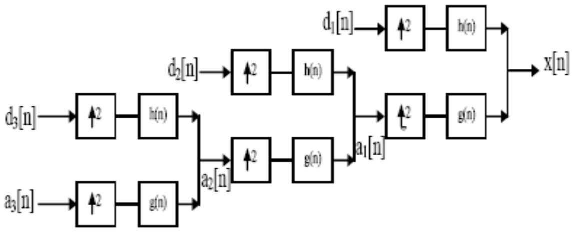


Fig. 5 Three Level Wavelet reconstruction tree

2.4 Selection of suitable Wavelet

The selection of relevant wavelet is an important task before starting the detection procedure. The choice of wavelet depends upon the type of signal to be analyzed. The wavelet similar to the shape of signal is selected. There are several wavelet families like Harr, Daubechies, Biorthogonal, Coiflets, Symlets, Morlet, Mexican Hat, Meyer etc. and several other Real and Complex wavelets. We have tried for Daubechies 4 (Db4),Db6, Db8,rbio6.8,bior 5.5 etc. However, Daubechies (Db4) Wavelet has been found to give details more accurately, which is also mentioned in the already published paper [22]. Moreover, this Wavelet shows similarity with QRS complexes and energy spectrum is concentrated around low frequencies. Therefore, we have chosen Daubechies 4 (Db4) Wavelet for extracting ECG features in our application. The Daubechies 4 Wavelet is shown in Fig. 6.

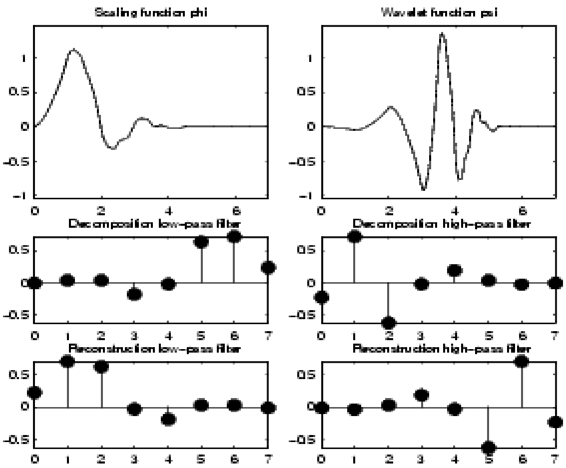


Fig. 6 Daubechies 4 wavelet

3. FEATURE EXTRACTION: DETECTION OF R PEAK & QRS COMPLEX

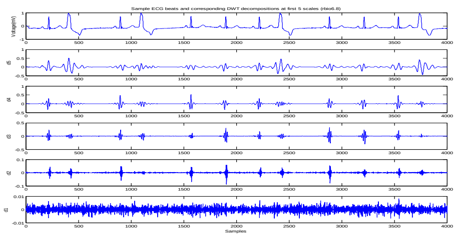
The purpose of the feature extraction process is to select and retain relevant information from original signal. The Feature Extraction stage extracts diagnostic information from the ECG signal. In order to detect the peaks, specific details of the signal are selected. The detection of R peak is the first step of feature extraction. The R peak in the signal from the Modified Lead II (MLII) lead has the largest amplitude among all the waves compared to other leads. The QRS complex detection consists of determining the R point of the heartbeat, which is in general the point where the heartbeat has the highest amplitude. A normal QRS complex indicates that the electrical impulse has progressed normally from the bundle of His to the Purkinje network through the right and left bundle branches and that normal depolarization of the right and left ventricles has occurred. Most of the energy of the QRS complex lies between 3 Hz and 40 Hz [23]. The 3-dB frequencies of the Fourier Transform of the wavelets indicate that most of the energy of the QRS complex lies between scales of 2^3 and 2^4, with the largest at 2^5. The energy decreases if the scale is larger than 2^5. The energy of motion artifacts and baseline wander (i.e., noise) increases for scales greater than 2^5. Therefore, we choose to use characteristic scales of 2^1 to 2^5 for the wavelet. In the proposed algorithm ECG signal is squared after removing noise (e. g .baseline wander) and decomposed up to level 5 using Db 4 wavelet thus separating approximate and detail coefficients. Then inverse Discrete Wavelet transform is applied to reconstruct the signal approximately. Then number of QRS complex wavelet transform features was extracted by selecting a window of -300ms to +400ms around the R wave as found in the database annotation. The 252-sample vectors were down sampled to 21, 25, 31, 42 or 63 samples (corresponding to $12x$, $10x$, $8x$, $6x$, $4x$ decimation, respectively), and normalized to a mean of zero and standard deviation of unity. This reduced the DC offset and eliminated the amplitude variance from file to file.QRS width is calculated from the onset and the offset of the QRS complex. The onset is the beginning of the Q wave and the offset is the ending of the S wave. Normally, the onset of the QRS complex contains the high-frequency components, which are detected at finer scales. To identify the onset and offset of the wave, the wave is made to zero base. The onset is the beginning and the offset is the ending of the complex.. Once this QRS complex is located the next step is to determine the onset and offset points for each QRS complex and to identify the component waves of the QRS complex. The Daubechies (db4) at level

three is found more suitable for R peak detection. The Discrete wavelet decomposition of sample signal for first five scales using db4 and rbio6.8 wavelet are shown in Fig. 7 and 8 respectively.

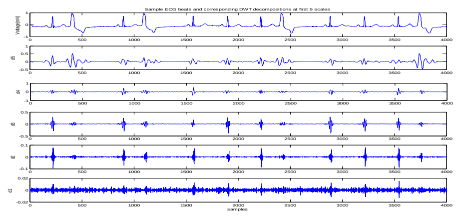
Fig. 7 ECG signal and its five level wavelet decomposition using db4

Fig. 8 ECG signal and its five level wavelet decomposition using rbio6.8

4. RESULTS AND DISCUSSION



An algorithm for R Peak and QRS complex detection using Wavelet Transform technique has been developed. Table 1 shows the detection results on the selected



files of MITBIH database. The information about the R Peak and QRS complex obtained is very useful for ECG Classification, Analysis, Diagnosis, Authentication and Identification performance. The QRS complex is also used for beat detection and the determination of heart rate through R-R interval estimation. This information can also serve as an input to a system that allows automatic cardiac diagnosis. The main advantage of this kind of detection is less time consumption for long time ECG signal. Further, ECG signal is a life indicator, and can be used as a tool for liveness detection. The physiological and geometrical differences of the heart in different individuals display certain uniqueness in their ECG signals. Hence ECG can be used as a tool for Identification and Verification of Individuals. Testing classification accuracy of 98.17% is achieved which is a promising result.

Table 1. Detection Results On The Selected Files Of MIT-BIH Database.

Record Number	Total beats tested	Beats correctly recognized	Recognition accuracy in %	Recognition time in seconds
100	2271	2238	98.55	80.526727
101	1866	1858	99.57	65.058209
103	2087	2083	99.81	72.523010
113	1793	1766	98.49	62.209119
115	1956	1951	99.74	68.260953
116	2357	2327	98.73	82.644785
123	1517	1516	99.93	52.773334
202	2139	2018	94.34	74.735331
220	2061	1964	95.29	71.920176
221	2440	2374	97.29	85.735838
234	2758	2705	98.08	96.774988
Total	23245	22800	1079.82	813.16247
Average	2113.18	2072.73	98.17	73.92

5. ACKNOWLEDGMENTS

Our thanks to the experts who have contributed in development of MIT-BIH database.

REFERENCES

[1] Prabhakar D. Khandait, N.G. Bawane and S.S. Limaye, " Wavelet Transformation , Artificial Neural Network and Neuro-Fuzzy Approach for CVD Detection and Classification – An Overview", First IEEE International Conference on Emerging Trends in Engineering and Technology(ICETET-08),pp. 612-617, DOI 10.1109/ICETET.2008.228.

[2] R. Hoekema, G. J. H. Uijen, and A. V. Oosterom,"Geometrical aspects of the interindividual variability of multilead ECG recordings," *IEEE Trans.Biomed. Eng.*, vol. 48, no. 5, pp. 551–559, May 2001.

[3] J. Pan and W.J. Tompkins, "A real-time QRS detection algorithm", *IEEE Trans. Biomed. Eng.*, vol. 32, pp. 230–236, 1985.

[4] V.X. Afonso, W.J. Tompkins, T.Q. Nguyen, and S. Luo, "ECG beat detection using filter banks", *IEEE Trans. Biomed. Eng.*, vol. 46, pp.192–202, 1999.

[5] J. Fraden and M.R. Neumann, "QRS wave detection", *Med. Biol. Eng. Comput.*, vol. 18, pp. 125–132, 1980.

[6] W.P. Holsinger, K.M. Kempner, and M.H. Miller, "A QRS preprocessor based on digital differentiation", *IEEE Trans. Biomed. Eng.*, vol. 18, pp. 121–217, May 1971.

[7] M. Bahoura, M. Hassani, and M. Hubin, "DSP implementation of wavelet transform for real time ECG wave forms detection and heart rate analysis", *Comput. Methods Programs Biomed.* vol. 52, no. 1, pp. 35– 44, 1997.

[8] S. A. Israel et al. "ECG to identify individuals". Pat. Rec. 38:133-142, 2005.

[9] Y.H. Hu, W.J. Tompkins, J.L. Urrusti, and V.X. Afonso, "Applications of artificial neural networks for ECG signal detection and classification", J.*Electrocardiology*, vol. 26 (Suppl.), pp. 66-73, 1993.

[10] G. Vijaya, V. Kumar, and H.K. Verma, "ANN-based QRS-complex analysis of ECG," *J. Med. Eng. Technol.*, vol. 22, no. 4, pp. 160-167

[11] Q. Xue, Y. H. Hu, and W. J. Tompkins, "Neural-network-based adaptive matched filtering for QRS detection", *IEEE Trans. Biomed. Eng.*, vol. 39, pp. 317-329, 1992.

[12] R. Poli, S. Cagnoni, and G. Valli, "Genetic design of optimum linear and nonlinear QRS detectors", *IEEE Trans. Biomed. Eng.*, vol. 42, pp. 1137- 1141, 1995.

[13] E. Skordalakis, "Syntactic ECG processing: a review", *Pattern Recognition.*, vol. 19, no. 4, pp. 305–313, 1986.

[14] S. K. Zhou, J.-T. Wang and J.-R. Xu, "The real-time detection of QRScomplex using the envelop of ECG", in *Proc. 10th Annu. Int. Conf.,IEEE Engineering in Medicine and Biology Society,* New Orleans, LA,1988, p. 38.

[15] R. V. Andreao, B. Dorizzi, and J. Boudy, "ECG signal analysis through hidden Markov models", *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 8, pp. 1541–1549, 2006.

[16] B.U. Kohler, C. Hennig, and R. Orglmeister, "The principles of software QRS detection", Engineering *in Medicine and Biology Magazine, IEEE*, vol. 21, pp. 42 – 57, Jan.-Feb. 2002

[17] S. S. Mehta et al. "Computer-aided interpretation of ECG for diagnostics", Int. Journal of System Science, 43-58, 1996.

[18] S. W. Chen et al. "A real time QRS detection method based on moving averaging incorporating with wavelet denoising", Comp. Methods and Progs. in Biomed. 82:187-195, 2006.

[19] L. Y. Shyu, Y. H. Wu, and W. C. Hu, "Using wavelet transform and fuzzy neural network for VPC detection from the holter ECG," *IEEE Trans.Biomed. Eng.*, vol. 51, no. 7, pp. 1269–1273, Jul. 2004.

[20] O. T. Inan, L. Giovangrandi, and G. T. A.Kovacs, "Robust neural-networkbased classification of premature ventricularcontractions using wavelet transform and timing interval features," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2507–2515, Dec. 2006.

[21] P. de Chazal, C. Heneghan, E. Sheridan, R.Reilly, P. Nolan, M. O'Malley, "Automated Processing of the Single-Lead Electrocardiogram for the Detection of Obstructive Sleep Apnoea", *IEEE Trans. Biomed.Eng.*, 50(6): 686-689, 2003.

[22] S. Z. Mahmoodabadi, A. Ahmadian, and M. D. Abolhasani, "ECG Feature Extraction using Daubechies Wavelets", Proceedings of the fifth IASTED International conference on Visualization, Imaging and Image Processing, pp. 343-348, 2005.

[23] N.V.Thakor, J.G.Webster and W.J.Tompkins, "Estimation of QRS complex power spectra for design of a QRS filter", IEEE Transactions on Biomedical Engineering, vol. BME-31, no. 11,pp. 702-706, 1986, pp. 702-706, Nov,1986.