

Similarity Measurement Technique of Two ECG Signals Analysis and Pattern Classification

¹Anil Kumar S. N, ²Dr. Sarika Tale

¹Student, VIAT, ²Associate Professor, VIAT, Karnataka, India

Abstract: Electrocardiogram (ECG) is the time varying signal represents the heart's electrical activity. The signal provides the information of coronary heart diseases, rhythm disorders etc. The classification of ECG signal is difficult since the morphological and temporal characteristics of ECG signal shows significant variations for different patients under different physical conditions. Many techniques available for analysis such as Neuro-Fuzzy, Self-organizing maps etc., This uses time-plane features ST, R, T segments thus increasing the computational complexity. In this paper, we use cross wavelet transform for the analysis and classification of electrocardiogram (ECG) signals. The cross-correlation between two time varying signals gives a measure of similarity between two waveforms. The application of the continuous wavelet transform to two time varying signals and the cross examination of the two decompositions reveal localized similarities in time and frequency. The application of the cross wavelet transform to signal yields wavelet cross spectrum (WCS) and wavelet coherence (WCOH). The proposed algorithm analyzes ECG data utilizing cross wavelet transform and explores the resulting spectral differences. A pathologically varying pattern from the normal pattern in the QT zone of the inferior leads shows the presence of inferior myocardial infarction. A normal beat ensemble is selected as the absolute normal ECG pattern template, and the coherence between various other normal and abnormal subjects is computed. The WCS and WCOH of various ECG patterns show distinguishing characteristics over two specific regions R1 and R2, where R1 is the QRS complex area and R2 is the T-wave region. The Physikalisch-Technische Bundesanstalt diagnostic ECG database is used for evaluation of the methods. A heuristically determined mathematical formula extracts the parameter from the WCS and WCOH. Empirical tests establish that the parameters are relevant for classification of normal and abnormal cardiac patterns. The overall accuracy, sensitivity and specificity after combining the three leads are obtained.

Keywords: Cross wavelet transform, fiducial points, interpolation, myocardial infarction, wavelet coherence (WCOH). *Software Required*—MATLAB, etc.

I. INTRODUCTION

An electrocardiogram (ECG) is a time varying signal that represents the activity of the heart. Each event has a distinctive waveform, the study of which can lead to greater insight into a patient's cardiac pathophysiology^{[1][2]}. An ECG can also be defined as a time varying signal reflecting the ionic current flow which causes the cardiac fibers to contract and subsequently relax. An Electrocardiogram signal can be used for detection of coronary artery disease, cardiomyopathies and left ventricular hypertrophy. It can also provide information for evaluation rhythm disorders. The ECG can be obtained by recording the potential difference between various electrodes placed on the surface of the skin, at specific locations. A single normal cycle of the ECG occurs with every heart beat. The objective of this paper is to find the similarities between the Normal and abnormal ECG signals using the measurement technique and pattern detection in ECG using Wavelet transform and therefore analyzing the ECG signal by using the pattern matching techniques for abnormality measure. This reduces the overall computational complexity in analyzing the signal. The similarities between the Normal and abnormal ECG signals are using the measurement technique and pattern detection in ECG using Wavelet transform and therefore analyzing the ECG signal by using the pattern matching techniques for abnormality measure. The uses of abnormal ECG pattern detection in ECG using Wavelet transform and pattern matching techniques for abnormality measure. Each individual heartbeat in the cardiac cycle of the recorded electrocardiogram (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 some underlying pathology, which could be detected by the analysis of the recorded ECG waveform^{[1][2]}. Coronary heart disease is one of the dominant health concerns all over the world. The analysis of individual ECG beat's characteristic shape, morphological features, and spectral properties can give significantly correlated clinical information for automatic detection of the ECG pattern. However, automated classification of ECG beats is a challenging problem because the morphological and temporal characteristics of ECG signals show significant variation for different patients under different physical conditions. Most of the classification methods of ECG beats use explicit time-plane features such as ST segment, R height, T height, etc. Apart from the measurement accuracy issues of the extracted features, a large feature set is obtained when time-plane features are used for classification. As a result, rule mining techniques are employed for feature set reduction, which increases the computational complexity. Good performance of any automatic ECG analyzing system depends upon the reliable and accurate detection of the basic characteristic features of the signal under study. QRS detection is necessary to determine the heart rate and is used as the reference point for beat alignment. Automatic delineation of the ECG has been widely studied, and many algorithms have been developed for QRS detection and fiducial point identification^{[3]-[6]}. ECG analysis algorithms operate on ECG data samples and generate automatic outputs, including morphology, time-interval measurements, and rhythm analysis^{[7]-[9]}. ECG signals are intrinsically nonstationary in nature. This makes wavelet transforms an effective tool for the analysis of ECG signals. Wavelet transforms have been applied to ECG signals for enhancing late potentials^[10], reducing noise^[12], QRS detection^[13], normal and abnormal beat recognition^[14], and delineation of ECG characteristic features^[11]. The methods used in these studies were conducted through continuous wavelet transform (CWT)^[15], multiresolution analysis, and dyadic wavelet transform^{[16][17]}.

Classification of ECG beats is a challenging problem because the morphological and temporal characteristics of ECG signals show significant variations for different patients. Many classification methods with distinguishing characteristics have been developed using neuro-fuzzy^[18] and self-organizing maps^{[8][20]-[22]}. Classification problems specific to myocardial infarction and ischemia are addressed in^{[9][19]} and^{[28]-[31]}, the time plane characteristics of the signal are employed for classification. A PC-based virtual instrument was used as a testing platform for acquiring, processing, presenting, and distributing ECG data^[36]. A very large scale integration implementation of a linear-phase digital filter for ECG signal processing has been designed in^[37], and the developed circuit is said to have very low computational complexity. A model based on the hidden Markov tree for ECG delineation technique is developed. An approach for human identification using standard 12-lead ECG recorded during rest is investigated in^[39] and^[40]. Most of the classification methods use explicit time-plane features such as ST segment, R height, T height, etc. Apart from the measurement accuracy issues of the extracted features, a large feature set is obtained when time-plane features are used for classification. As a result, rule mining techniques are employed for feature set reduction, which increases the computational complexity.

Myocardial infarction: Myocardial infarction (MI) or acute myocardial infarction (AMI), commonly known as a heart attack occurs when blood flow stops to part of the heart causing damage to the heart muscle. The most common symptom is chest pain or discomfort which may travel into the shoulder, arm, back, neck, or jaw. Often it is in the centre or left side of the chest and lasts for more than a few minutes. The discomfort may occasionally feel like heartburn. Other symptoms may include shortness of breath, nausea, feeling faint, a cold sweat, or feeling tired.^[42] About 30% of people have atypical symptoms,^[42] with women more likely than men to present atypically. Among those over 75 years old, about 5% have had an MI with little or no history of symptoms. An MI may cause heart failure, an irregular heartbeat, or cardiac arrest.^[42]

II. OVERVIEW OF WAVELET TRANSFORM

A wavelet is wave-like oscillation with amplitude that begins at zero, increases, and then decreases back to zero. It can typically be visualized as a "brief oscillation" like one might see recorded by a seismograph or heart monitor. Generally, wavelets are purposefully crafted to have specific properties that make them useful for signal processing. Wavelets can be combined, using a "reverse, shift, multiply and integrate" technique called convolution, with portions of a known signal to extract information from the unknown signal. **Application:** One of the most popular applications of wavelet transform is image compression. The advantage of using wavelet-based coding in image compression is that it provides significant improvements in picture quality at higher compression ratios over conventional techniques. Since wavelet transform has the ability to decompose complex information and patterns into elementary forms, it is commonly used in acoustics processing and pattern recognition. Moreover, wavelet transforms can be applied to the following scientific research areas: edge and corner detection, partial differential equation solving, transient detection, filter design, electrocardiogram (ECG) analysis, texture analysis, business information analysis.

A. Continuous Wavelet Transform:

Wavelet transform can be used to analyze time series that contain nonstationary power at many different frequencies. Assume that one has a time series x_n with equal time spacing dt and $n = 0 \dots N - 1$. A mother wavelet $\psi_0(\eta)$ is taken that depends on a nondimensional parameter time given by η . It qualifies as a wavelet if this function has zero mean and is localized in both the time and frequency space.^[27] CWT involves decomposing a signal x_n into a number of translated and dilated wavelets. The main idea behind this is to take a mother wavelet, translate and dilate it, convolve it with the function of interest, and map out the coefficients in wavelet space spanned by translation and dilation. Periodic behaviour then shows up as a pattern spanning all translations at a given dilation, and this redundancy in the wavelet space makes detection of periodic behaviour rather easy. The wavelet transform preserves temporal locality, which is an advantage over Fourier analysis. In definition, the continuous wavelet transform is a convolution of the input data sequence with a set of functions generated by the mother wavelet. The convolution can be computed by using the Fast Fourier Transform (FFT). Normally, the output $X_w(a, b)$ is a real valued function except when the mother wavelet is complex. A complex mother wavelet will convert the continuous wavelet transform to a complex valued function. The power spectrum of the continuous wavelet transform can be represented by $|X_w(a, b)|^2$. **Scale factor:** The scale factor either dilates or compresses a signal. When the scale factor is relatively low, the signal is more contracted which in turn results in a more detailed resulting graph. However, the drawback is that low scale factor does not last for the entire duration of the signal. On the other hand, when the scale factor is high, the signal is stretched out which means that the resulting graph will be presented in less detail. Nevertheless, it usually lasts the entire duration of the signal. **Time-Frequency Analysis:** In signal processing, time frequency analysis comprises those techniques that study a signal in both the time and frequency domains simultaneously, using various time–frequency representations. Rather than viewing a 1-dimensional signal (a function, real or complex-valued, whose domain is the real line) and some transform (another function whose domain is the real line, obtained from the original via some transform), time frequency analysis studies a two-dimensional signal a function whose domain is the two-dimensional real plane, obtained from the signal via a time frequency transform. **Morlet Wavelet:** The Morlet wavelet transform method presented offers an intuitive bridge between frequency and time information which can clarify interpretation of complex head trauma spectra obtained with Fourier transform. The Morlet wavelet transform, however, is not intended as a replacement for the Fourier transform, but rather a supplement that allows qualitative access to time related changes and takes advantage of the multiple dimensions available in a free induction decay analysis.^[45] The Morlet wavelet (or Gabor wavelet) is a wavelet composed of a complex exponential (carrier) multiplied by a Gaussian window (envelope). This wavelet is closely related to human perception, both hearing and vision.

The parameter in the Morlet wavelet allows trade between time and frequency resolutions. Conventionally, the restriction is used to avoid problems with the Morlet wavelet at low (high temporal resolution). For signals containing only slowly varying frequency and amplitude modulations (audio, for example) it is not necessary to use small values. In this case, becomes very small and is, therefore, often neglected. Under the restriction the frequency of the Morlet wavelet is conventionally taken. The wavelet exists as a complex version or a purely real-valued version. Some distinguish between the "real Morlet" vs. the "complex Morlet".^[45] Others consider the complex version to be the "Gabor wavelet", while the real-valued version is the "Morlet wavelet".^[45] The Morlet wavelet is the mother wavelet used in this paper (Fig.1) because it provides good balance between time and frequency localization.^{[15] [17]}

The Morlet wavelet is defined as

$$\psi_0(\eta) = \pi^{-1/4} e^{i\omega_0\eta} e^{-(1/2)\eta^2} \tag{1}$$

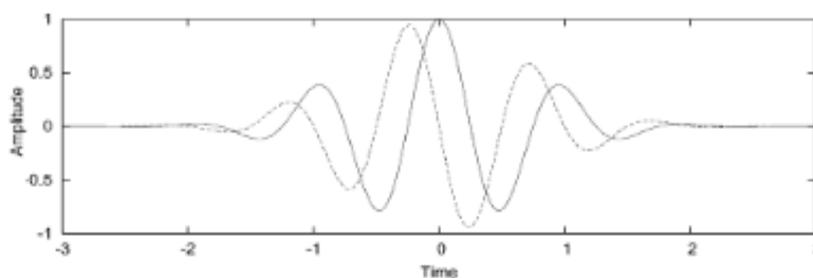


Fig.1 Real (solid lines) and imaginary (dashed line) part of the Morlet wavelet with $\omega_0 = 6$.

Where ω_0 is a dimensionless frequency and η is dimensionless time. The application of the Morlet wavelet analysis in the electrocardiogram (ECG) is mainly to discriminate the abnormal heartbeat behaviour. Since the variation of the abnormal heartbeat is a non-stationary signal, then this signal is suitable for wavelet-based analysis. [45]

B. Cross Wavelet transforms:

The study of interrelation between pairs of time varying signals can be performed by the application of Cross wavelet transform. The cross wavelet transform of two time series x_n and y_n is defined as

$$W^{XY} = W^X W^{Y*} \tag{2}$$

Where * denotes complex conjugation. We further defined the cross wavelet power as $|W^{XY}|$. The complex argument $\arg(W^{XY})$ can be interpreted as the local relative phase between x_n and y_n in the time–frequency space. The theoretical distribution of the cross wavelet power of two time series with background power spectra P_k^X and P_k^Y is given in [32]. Another useful measure is how coherent the cross wavelet transform is in time–frequency space. Following [33], the WCOH of two time series is defined as

$$R_n^2(s) = \frac{|S(s^{-1} W_n^{XY}(s))|^2}{S(s^{-1} |W_n^X(s)|^2) \cdot S(s^{-1} |W_n^Y(s)|^2)} \tag{3}$$

Where S is a smoothing operator and s is the scale. WCOH can be thought of as a localized correlation coefficient in the time–frequency space.

III. SOFTWARE DEVELOPMENT METHOD

The proposed method, cross wavelet transform is used for the analysis of ECG signals. R-peak registration followed by beat segmentation is performed for accurate time alignment of cardiac cycles and comparison of clinical signatures of two patterns under study. The heart rate is a variable quantity and it changes with beat duration. Therefore, each segmented beat is time normalized for proper alignment. These beats are subjected to further analysis for parameter selection. The cross wavelet analysis of the ECG beats reveals significant characteristic features of the waves under study. The proposed algorithm uses R-peak detection for beat segmentation, and extraction of any other explicit time-plane features is not required. The detailed description of the method is illustrated in Fig.2.

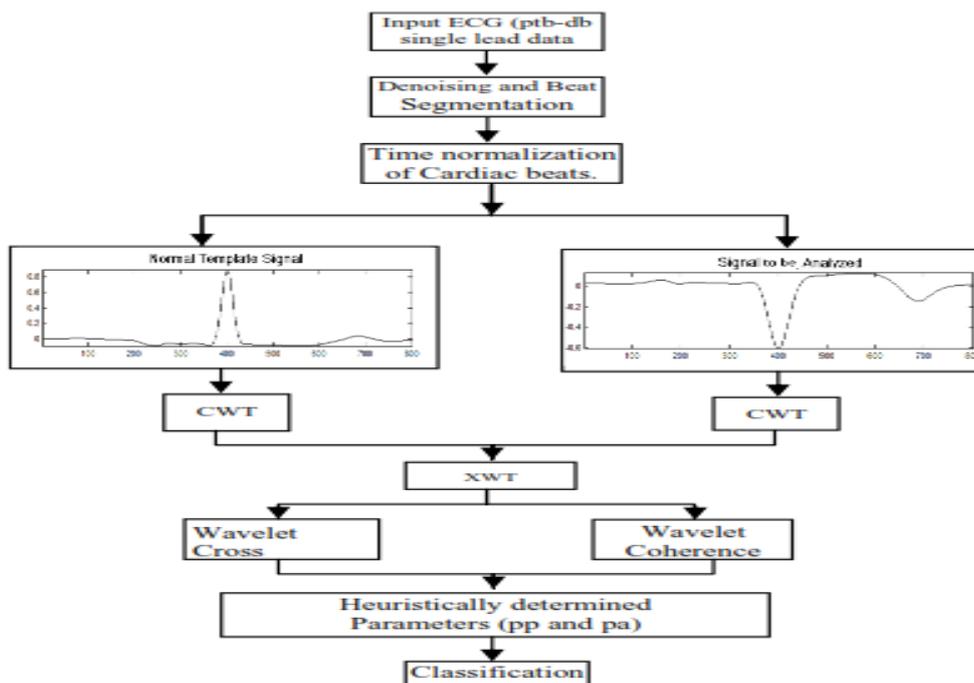


Fig.2 Flowchart for the analysis

B. ECG Data:

The National Metrology Institute of Germany has provided this compilation of digitized ECGs for research, algorithmic benchmarking or teaching purposes to the users of PhysioNet. The ECGs were collected from healthy volunteers and patients with different heart diseases. ^[23]The database contains 549 records from 290 subjects (aged 17 to 87, mean 57.2; 209 men, mean age 55.5, and 81 women, mean age 61.6; ages were not recorded for 1 female and 14 male subjects). ^[23] Each subject is represented by one to five records. There are no subjects numbered 124, 132, 134, or 161. Each record includes 15 simultaneously measured signals: the conventional 12 leads (i, ii, iii, avr, avl, avf, v1, v2, v3, v4, v5, v6) together with the 3 Frank lead ECGs (vx, vy, vz). Each signal is digitized at 1000 samples per second, with 16 bit resolution over a range of ± 16.384 mV. On special request to the contributors of the database, recordings may be available at sampling rates up to 10 KHz. ^[23] The classification method is developed in a manner where accuracy is measured over this beat ensemble thus obtained. If the entire beat ensembles obtained from a subject are normal, then the subject is classified as normal; otherwise they are assumed to be pathological.

Selection of a Standard Normal Template: A cardiac beat from a 25-year-old pathologically normal non-smoking male subject with a heart rate of 72 beats per min is considered as the normal template for analysis. A beat ensemble from patient Id: ptbdb/patient150/s0287ire is considered as the normal template. ^[23] This normal template is validated using standard textbooks ^{[1] [2]} and also by visual inspection by doctors.

All the ECG data in this paper have been selected from the Physikalisch-Technische Bundesanstalt diagnostic 12-lead ECG database (ptbdb) of Physionet. ^[23] The ptbdb ECG diagnostic database contains 549 records from 290 subjects with 52 healthy controls and 148 MI patients. The dataset is sampled at 1 kHz. The reason for choosing the ptbdb is that this database has abundant and well-classified ECG recordings related to MI and our main task in this paper are to detect MI from ECGs. In ptbdb, conventional simultaneously measured 12-lead data are presented. The inferior wall of the left ventricular cone is oriented to the standard leads II, III, and AVF. ^[23] IMI will therefore be reflected by the appearance of the classic features of hyperacute, fully evolved, and chronic stabilized phases in these leads. In the fully evolved phase, the standard lead III commonly reflects a QS complex and the standard leads II and AVF reflect QR complexes. Deviations of the ST segment and the T wave from normal pattern are also clinically significant. Type 1 IMI is non-Q type, with the presence of ST elevation and attenuated QRS complex, and Type 2 IMI is Q-type MI with deep Q and an inverted T. In this paper, the abnormal class consists of these two patterns. A total of 44 IMI and 220 normal records are available and are used for the evaluation of the system. ^[23] From these records, 18 489 cardiac beat ensembles are extracted for analysis and classification. One beat ensemble is an entity of three beats. Each beat has been taken from three individual inferior leads emerging at the same time generated by one cardiac cycle. Each beat from one ensemble is analyzed using cross wavelet transform, and the combined results are used for classification. Each record in the ptbdb is accompanied with a header file containing the present status of the patient. Description of cardiac ailment (if any) of the subject is also mentioned in the record, clearly stating the type of MI or other prevailing ailment. Since MI is a nonarrhythmic cardiac abnormality, the ECG trace will have homogeneity in the beat morphology throughout the record. An IMI patient's record contains beats with IMI morphology only, and this is true for normal subjects. ^{[1] [2]} This fact was also validated by the physicians. Moreover, we visually confirmed this fact for more than 50 records (both normal and abnormal) randomly selected from the database.

C. Denoising and R-Peak Registration:

Denoising of ECG data is an essential step before any form of analysis because this increases the efficiency of the algorithm. This paper uses Discrete Wavelet transform based decomposition and selective reconstructions of wavelet coefficients for Denoising and QRS detection. The Denoising and basic feature extraction technique used for this paper is the method developed. ^[11] In the stated technique, noise is eliminated by selective decomposition and identification of the noisy frequency band and thereby rejecting the corresponding coefficients. For, identification of the QRS frequency band, the detail coefficients D4 and D5 are selected, because together they contain most of the QRS information.

D. Normalization of ECG Signals:

Once the R peaks are registered, the R-R interval is computed and divided into 2:1 ratios (say, 2x: x points). One cardiac cycle gives the details of the pathological condition of the subject, and hence each beat needs to be segmented before subjecting it to cross wavelet analysis. Considering x points to the left and 2x points to the right of the R index, one cardiac beat is extracted. As the heart rate varies for each subject, the length of the cardiac cycle also varies. At a sampling rate m, if the heart rate is H beats/min, then the number of samples in one beat is given by

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

Therefore, n is a function of H.

The objective is to get equal number of samples in all beats so that point-to-point correlation analysis becomes possible. To get an equal number of samples, fast Fourier transform based interpolation technique has been used.^[24] After interpolation, all segmented beats have 1000 samples. This normalization^[19] is required for spectral pattern matching because point-to-point correlation will be analyzed through cross wavelet transform.

E. Cross Wavelet Transform:

Cross-correlation is a measure of similarity between two waveforms. Application of CWT to two time series and the cross examination of the two decompositions reveals localized similarities in time and scale (scale being nearly inverse of frequency) and divulge various characteristic information of the signal under study. As shown in Fig. 2, the result of cross wavelet transform of two signals generates the WCS and WCOH. WCS and WCOH are used for cross examination of a single normal and abnormal (IMI) beat with a standard normal template beat. Because of the morphological similarity with that of the QRS complex, the Morlet wavelet is selected as the mother wavelet for analysis. Cross wavelet transform gives a relationship between the two signals in timescale space. In this analysis, 512 scales are considered. The resultant WCS shows the spectral components of interest. Figs. 3–5 shows the distinguishing regions R1, the QRS complex region, and R2, the T-wave region. It is evident from the colour-coded plots that there exist distinct variations in the spectral and coherence components, revealing the nature of the analyzed signals. After analysis, parameter extraction formulas are developed for classification of normal and abnormal cardiac patterns. The procedure is discussed in the next section.

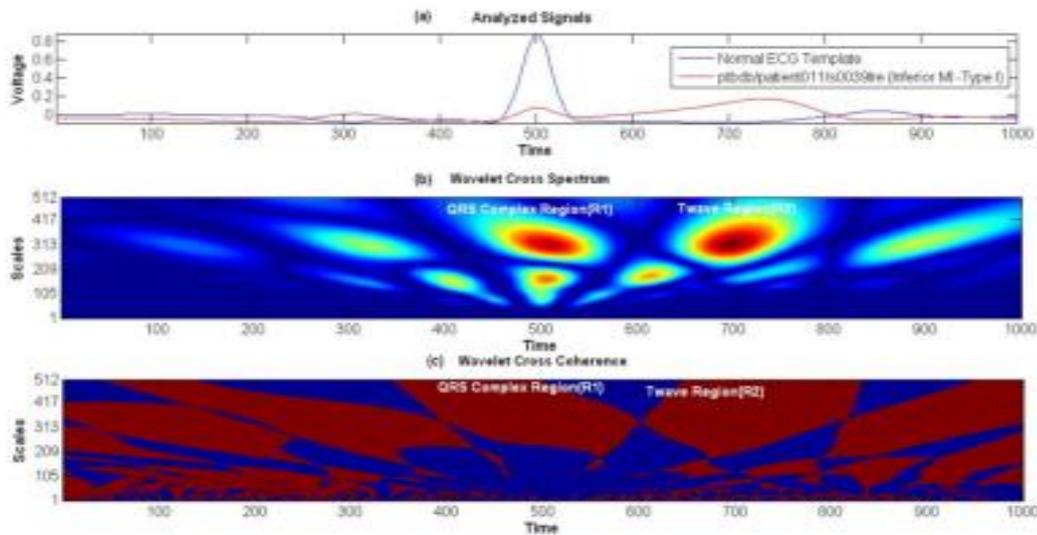


Fig.3 WCS and WCOH between the standard normal template and an abnormal ECG (IMI-Type1)

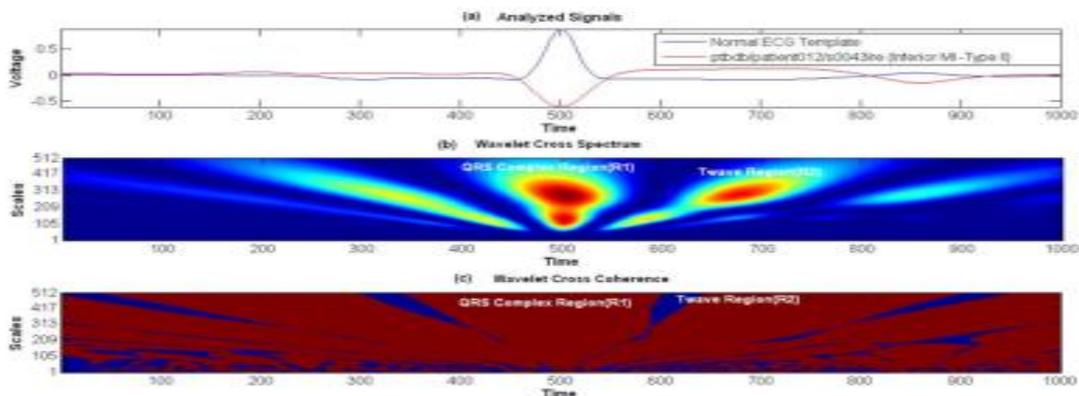


Fig.4 WCS and WCOH between the standard normal template and an abnormal ECG (IMI-Type2)

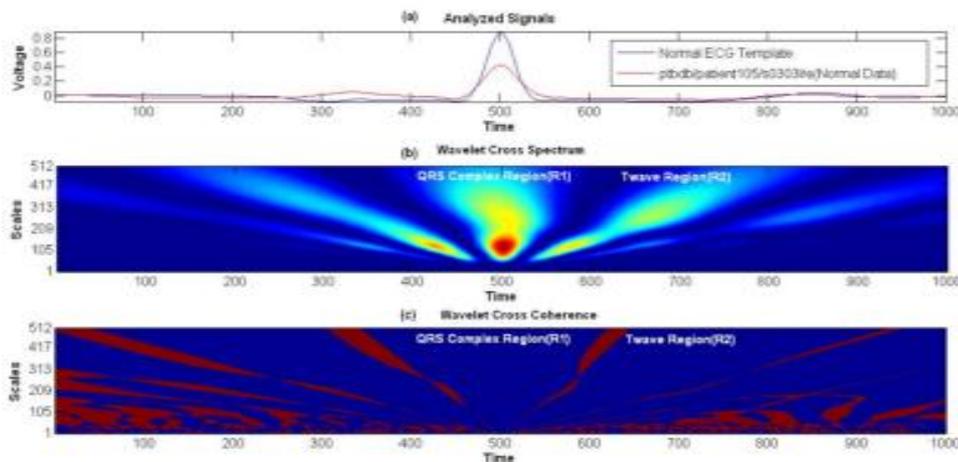


Fig.5 WCS and WCOH between the standard normal template and normal ECG data

F. Feature Extraction and Parameter Identification:

Feature extraction involves reducing the amount of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which overfits the training sample and generalizes poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. Cross wavelet transform generates WCS and WCOH, which are matrices containing the WCS and WCOH between two signals. In this paper, normal subjects and subjects with IMI are considered. There exists a pathologically varying QT pattern for normal and abnormal data. The QT zone, being the pathological region, is selected for parameter extraction and analysis. A span of 80 points from the left to 400 points to the right of the registered R peak is the QT zone [1] [2]. This start and end of the time zone are marked as t1 and t2, respectively. Visual inspection of the colour-coded spectrogram for WCS and WCOH in Figs. 3–5 reveals that the effect of the analysis is prominent over a particular scale range. Heuristically determined mathematical formulas are developed and tested over several datasets. These equations are for feature extraction from WCS and WCOH.

1) Scale Selection from WCS:

To find the actual span of scales that are significant for analysis, a timescale relation is established. From the WCS equation (4) is developed, where s signifies the scale, which varies from 1 to 512.

$$\text{sum_WCS}(s) = \sum_{t1}^{t2} \text{WCS}(s, t). \quad (4)$$

The variable $\text{sum_WCS}(s)$ contains the summation of the WCS value at each scale over the whole ECG beat. The $\text{sum_WCS}(s)$ values of several subjects both normal and abnormal are plotted in Fig. 2.4. By inspecting the graph, it is found that variation of the cross wavelet spectrum for normal and IMI subject is most prominent in the range marked with $s1 = 75$ to $s2 = 300$, where the normal is plotted with the solid lines and abnormal with the dotted line. Any parameter (or feature) extracted from this scale range over the QT zone from WCS and WCOH will produce a unique identification signature which can be used for classification. So this scale range is selected by visual inspection as shown in fig.6.

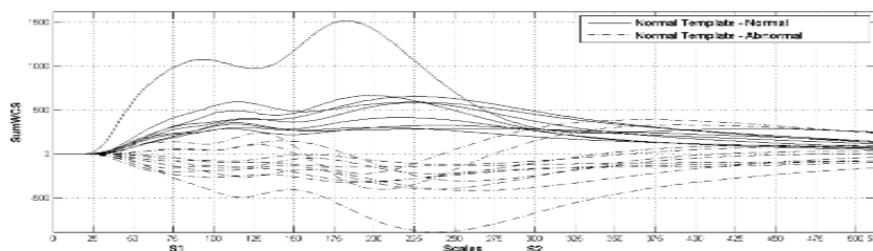


Fig.6 Plot for the identification of scale range

2) WCS- and WCOH-Based Parameter Extraction:

The parameters were extracted from the WCS and WCOH matrices over the selected scale range of s_1 to s_2 and over the QT segment t_1 to t_2 . Equations (5) and (6) are framed for parameter extraction which is used for classification

$$\text{sum_WCS_scale}(pp) = \sum_{s_1}^{s_2} \sum_{t_1}^{t_2} \text{WCS}(s, t) \quad (5)$$

$$\text{sum_WCOH_scale}(pa) = \sum_{s_1}^{s_2} \sum_{t_1}^{t_2} \text{WCOH}(s, t). \quad (6)$$

Accordingly, a threshold value (th) from (5) and (6) is set for normal and abnormal class identification. Three sets of parameters (pa_1, pp_1), (pa_2, pp_2), (pa_3, pp_3) are extracted from the standard leads II, III, and AVF. Separate threshold values are estimated for each of the leads and hence the dataset is subjected to threshold-based classification. The next section shows the results and empirical validation of the extracted parameter.

IV. PATTERN CLASSIFICATION

In the present scenario, the data classification reduces to a two-class problem creating a partition between normal and IMI class. As stated in Section III-A, standard leads II, III, and AVF are used for the identification of IMI ECG patterns. A heuristically determined mathematical formula extracts the parameter from the WCS and WCOH. Empirical tests establish that the parameter is relevant for classification of normal and abnormal cardiac patterns. However, many complex classification techniques exist in [25] and [26]. In this paper, a threshold-based classification is proposed because the extracted parameters have sufficiently distinct class of separable characteristics. Once the parameters are determined, particular threshold values of pp and pa are estimated for all the leads. After visual inspection of the normal and abnormal clusters, it was found that data points of the normal class lie in the upper right quadrant of the cluster plot. It was also seen that the cluster overlapping region has a parameter value greater than the centroid of the abnormal class and some of the data points of the abnormal class overlay with that of the normal cluster. So, points having pp and pa values greater than the centroid of the abnormal class are considered for finding the threshold and are sorted in ascending order. [41]

4.1 Threshold based classification:

The method employed for finding the threshold considers that 1% of the data points are dispersed and 99% of data coverage is given to the abnormal data class. The algorithm for threshold value calculation is as follows. Repeat Steps 1–3 for leads II, III, and AVF. [41]

Step 1: Find the centroid value of the abnormal class centroid (c_1, c_2).

Step 2: In the abnormal class, check the points having:

- (a) $pa > c_1$; mark them as dataset1.
- (b) $pp > c_2$; mark them as dataset2.

Step 3: Sort the obtained datasets in ascending order individually and give 99% coverage to the datasets.

Find the associated parameter value obtained at the margin and set those values as the thresholds (th) of pa and pp . Next, a score-value method is proposed, where a data point classified in normal class is given a score of 1 and when classified in the abnormal section is marked as 0. The total scores (say, lead total) for an absolute normal becomes 3 whenever each of the beats in the ensemble of leads is marked as normal and gets a score of 1. Any score less than 3 signifies the existence of pathology in one or more leads, and, then, based on the marked abnormal lead, further inspection is carried out accordingly.

The generated threshold-based classification rule is stated below.

- 1) If $pa_1 < th_{11}$ and $pp_1 < th_{12}$, then “abnormal” mark as “0” else “normal” mark as “1” end if
- 2) If $pa_2 < th_{21}$ and $pp_2 < th_{22}$, then “abnormal” mark as “0” else “normal” mark as “1” end if

3) If $pa3 < th31$ and $pp3 < th32$, then “abnormal” mark as “0” else “normal” mark as “1” end if

The score-value method is illustrated as follows.

If lead total = 3, then “absolute normal” else if lead total < 3,

Then “inspect the 0/1 score lead(s) for knowledge of infarction region”.

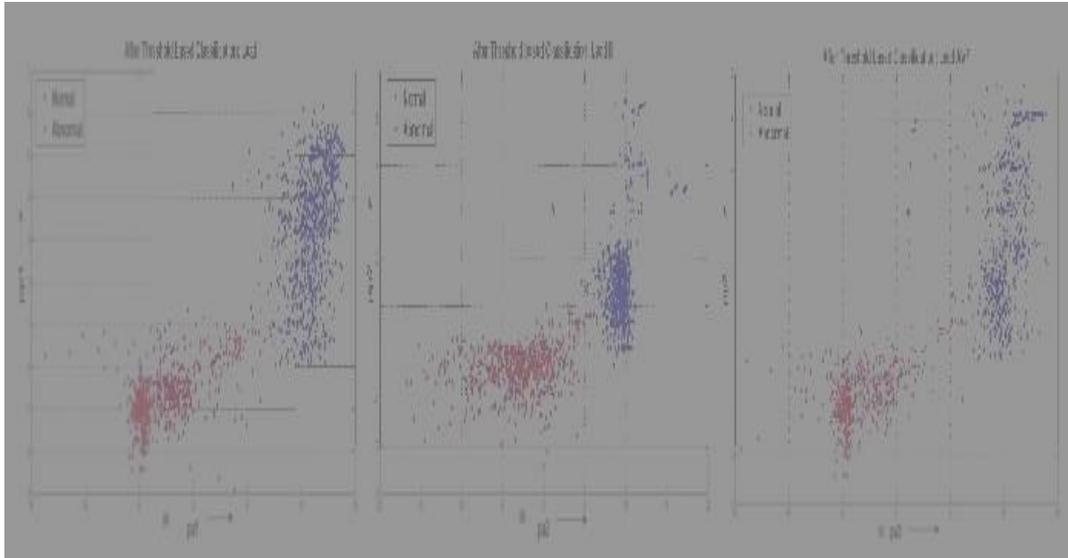


Fig.7 Threshold based classification over leads II, III and AVF

The Threshold based classification over leads II, III and AVF are shown in the above figure (fig.7).

4.2 Centroid based Classification:

The steps for the centroid based classification are as follows: Features extraction and parameter identification is done for the ECG signals by applying cross wavelet transform that resulting in wavelet cross spectrum and wavelet coherence. The summation of the data of the WCS and WCOH are calculated. The clusters of the datasets of the parameters are calculated and plotted. Using the rule of the centroid lying in which quadrant of the graph decides the normality and abnormality of the ECG signal. If the centroid lies in the both positive quadrant the signals are normal. If centroid lies in the both negative region the signal is abnormal myocardial infarction occurred. This method reduces the calculation of average threshold region of the cluster data and thus the computational complexity.

V. EXPERIMENTAL RESULTS

All the input data for this method has been selected from ptbdb of Physionet. The results are tested on three standard leads II, III, and AVF.

A. Variation Analysis of Morphologically Distinct ECG:

Data by WCS and WCOH reveals the differences in the WCS and WCOH when normal/ normal and normal/abnormal pairs are subjected to cross wavelet transform analysis. Two major regions of difference are marked as R1 and R2. R1 is the QRS complex region, and R2 is the T-wave region. Type 1 IMI, (non-Q type, with ST elevation and attenuated QRS complex), normal, and Type 2 IMI (Q-type MI with deep Q and inverted T). From the color-coded spectrogram for WCS and WCOH, it is evident that spectral and coherence variations exist in the regions R1 and R2. Two distinguishing parameters are determined for classification of normal and abnormal data. For concise presentation of data and because of space constraints, only a limited number of analyses are shown ^[41].

B. Performance Evaluation Metrics for Beat Classification:

In feature extraction and parameter identification ^[41] the equations (5) and (6) were developed and validated using test data. The spectrograms of sections variation Analysis of Morphologically Distinct ECG the results when applied over pairs of normal template/normal and normal template/abnormal beats. Parameter (or feature) extracted from this scale

range over the QT zone will produce a unique identification signature. In this paper, we have considered three statistical Indices: accuracy (Acc), sensitivity (Se), and specificity (Sp) are defined in following equations respectively. ^[41]

The most crucial metric for determining the overall system performance is usually accuracy. We can deduce the overall accuracy of the classification for each file as

$$Acc = \frac{NT - NE}{NT} \times 100.$$

In this equation, Acc is the accuracy, and the variables NE and NT represents the total number of classification errors and beats in the file, respectively.

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) are 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 nonevents, and is given by

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

Where, false positives (FP) are the number of falsely detected events.

VI. DISCUSSION AND FUTURE SCOPE

The method presented in this paper subjects the ECG beat patterns to wavelet transform and then explores the existing spectral variations between two signals. The ECG datasets are preprocessed for the removal of artifacts and other noise before subjecting them to cross wavelet transform-based analysis. If noisy beats are used as input to the system, the power Line interference noise will pose a problem during analysis because its frequency lies in the same band as that of the ECG signals. However, other noises will not affect the system because after CWT decomposition they will reside in a different scale range.

Unlike other existing methods, this algorithm does not require extraction of explicit time-plane features. The classification is based on only two attributes per lead. Therefore, the computational time required for classification reduces considerably.

The work reported here may be extended for 12-lead ECG data classification for differentiating different types of MI by examining various lead groups. Some preliminary investigations have shown that the parameter values *pp* and *pa* obtained from regions R1 and R2 can be used for ST-type and Q-type MI. We also propose to cover the arrhythmic disease class through our system. This also provides scope for further research work on these issues.

VII. CONCLUSION

In this paper, we presented a method for the analysis of ECG patterns using the cross wavelet transform technique. Cross-correlation is a measure of similarity between two waveforms. The application of CWT to two time varying signals and the cross examination of the two decomposition reveal localized similarities in time and scale. A Morlet wavelet was used as the mother wavelet. From the analysis, it was found that WCS and WCOH reveal the prominent dissimilarities between two time series. Region based differences were visible in WCS and WCOH spectrograms of normal/normal and normal/abnormal pairs. Visible dissimilarities in the regions R1 and R2 were marked, and after selection of scale appropriate classification parameters were established empirically. The classification accuracy for leads II, III, and AVF was obtained. This result further opens the possibility for extending the method for 12-lead ECG-based classification system and addressing other cardiac abnormality issues.

REFERENCES

- [1] L.Schamroth, anntroduction toElectrocardiography,7thed. New York, NY,USA:Wiley 2009.
- [2] A. L. Goldberg, Clinical Electrocardiography, 7th ed. Amsterdam, the Netherlands: Elsevier, 2010.
- [3] S. Mitra, M. Mitra, and B. B. Chaudhury, "Pattern defined heuristic rules and directional histogrambasedonlineECGparameterextraction,"*Measurement*,vol.42,no.1,pp.150-156, 2009.
- [4] H. K. Chatterjee, R. Gupta, and M. Mitra, "A statistical approach for determination of time plane features from digitized ECG," *Comput. Biol.Med.* vol. 41, no. 5, pp. 278–284, 2011.
- [5] P. Jiapu and W. J. Tompkins, "A real time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. 32, no. 3, pp. 230–236, Mar. 1985.
- [6] O. Pahlm and L. Sörnmo, "Software QRS detection in ambulatory monitoring—A review," *Med. Biol. Eng. Comput.*, vol. 22, no. 4, pp. 289–297, 1984.
- [7] M. C. MacLachlan, B. F. Nielsen, M. Lysaker, and A. Tveito, "Computing the size and location of myocardial ischemia using measurements of ST-segment shift," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 6, pp. 1024–1031, Jun. 2006.
- [8] S. Mitra, M. Mitra, and B. B. Chaudhury, "A rough-set-based inference engine for ECG classification," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 6, pp. 2198–2206, Dec. 2006.
- [9] T. H. Linh, S. Osowski, and M. Stodolski, "On-line heart beat recognition using hermite polynomials and neuro-fuzzy network," *IEEE Trans. Instrum.Meas.*,vol.52,no.4,pp.Aug. 2003.
- [10] P. Caminal and N. V. Thakor, "Detection of late potentials by means of wavelet transform," *IEEE Trans. Biomed. Eng.*, vol. 41, no. 7, pp. 625–634, Jul. 1994.
- [11] S. Banerjee, R. Gupta, and M. Mitra, "Delineation of ECG characteristic features using multiresolution wavelet analysis method," *Measurement*, vol.45,no.3,pp. 474–487, Apr. 2012.
- [12] J. R. Murray, S. Kadambe, and G. F. Boudreaux-Bartels, "Extensive analysis of a QRS detector based on the dyadic wavelet transform," in *Proc. IEEE-SP Int. Symp. Time-Freq. Time-Scale Anal.*, Oct. 1994, pp. 540–543.
- [13] J. L. Senhadji, J. J. Bellanger, G. Carrault, and J. L. Coatrieux, "Wavelet analysis of ECG signals," in *Proc. IEEE EMBS*, Mar. 1990, pp. 811–812.
- [14] P. Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1196–1206, Jul. 2004.
- [15] A. Grossmann and J. Morlet, "Decomposition of hardy functions into square integrable wavelets of constant shape," *SIAM J. Math. Anal.*, vol. 15, no. 4, pp. 723–736, 1984.
- [16] S. G. Mallat, "A theory of multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11,no7, pp. 674–693, Jul. 1989.
- [17] I. Daubechies, "Orthonormal basis of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, no. 7, pp. 909–996, 1988.
- [18] M. Engin, "ECG beat classification using neuro-fuzzy network," *Pattern Recognit. Lett.* vol. 25, no. 15, pp. 1715–1722, Nov. 2004.
- [19] M. S. Islam, N. Alajlan, Y.Bazi, and H.S.Hichri, "HBS: A novel biometric feature based on heartbeat morphology," *IEEETrans. Inf.Technol.Biomed.*, vol.16,no.3,pp.445–453, May 2012.
- [20] P. Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1196–1206, Jul. 2004.
- [21] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt, and L. Sornmo, "Clustering ECG complexes using Hermite functions and self-organizing maps," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 7, pp. 839–847, Jul. 2000.

- [22] P. de Chazal and R. B. Reilly, "A patient adapting heart beat classifier using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2535–2543, Dec. 2006.
- [23] (2012). PTB Diagnostic ECG Database Directory, Physiobank Archive Index, PTB Diagnostic ECG Database [Online]. Available: <<http://physionet.org/physiobank/database>>
- [24] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing*, Englewood Cliffs, NJ, and USA: Prentice-Hall, Apr. 2006.
- [25] R.O.Duda, P.E.Hart, and D.G.Stork, *Pattern Classification*, 2nd Ed, NY, USA: Wiley, 2007.
- [26] J. T. Tou and R. C. Gonzaleze, *Pattern Recognition Principles*, 4thed. Reading, MA, USA: Addison-Wesley, 1974.
- [27] I. Daubechis, "The wavelet transform time frequency localization and signal analysis," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1004, Sep. 1990.
- [28] P.-C. Chang, J.-J. Lin, J.-C. Hsieh, and J. Weng, "Myocardial infarction classification using multi-lead ECG using hidden Markov models and Gaussian mixture models," *Appl. Soft Comput.*, vol. 12, no. 10, pp. 3165–3175, 2012.
- [29] J. Fayn, "A classification tree approach for cardiac ischemia detection using spatiotemporal information from three standards ECG leads," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 95–102, Jan. 2011.
- [30] J.Fayn, P.Rubel, O.Pahlm, G. S. Wagner, "Improvement of the detection of myocardial ischemia thanks to information technologies," *Int.J.Cardiol.*vol.120,no.2,pp.187 2007.
- [31] C. Papaloukas, D. I. Fotiadis, A. Likas, and L. K. Michalis, "An ischemia detection method based on artificial neural networks," *Artif. Intell.Med.*,vol.24,no. 2, pp. 167–178, 2002.
- [32] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bull. Amer. Meteorol. Soc.*, vol. 79, no. 1, pp. 61–78, 1998.
- [33] C. Torrence and P. Webster, "Interdecadal changes in the ESNO– Monsoon system," *J. Clim.*, vol. 12, no. 8, pp. 2679–2690, Aug. 1999.
- [34] A. Grinsted, J. C. Moore, and S. Jevrejeva, "Application of the cross wavelet transform and wavelet coherence to geophysical time series," *Nonlinear Process Geophys.*, vol. 11, nos. 5–6, pp. 561–566, 2004.
- [35] D. Dey, B. Chatterjee, S. Chakravorti, and S. Munshi, "Cross-wavelet transform as a new paradigm for feature extraction from noisy partial discharge pulses," *IEEE Trans. Dielectr. Electr. Insul.* vol. 17, no. 1,pp. 157–166, Feb. 2010.
- [36] .A. Josko and R. J. Rak, "Effective simulation of signals for testing ECG analyzer," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 3, pp. 1019–1024, Jun. 2005.
- [37] T. Raita-Aho, T. Saramaki, and O. Vainio, "A digital filter chip for ECG signal processing," *IEEE Trans. Instrum. Meas.*, vol. 43, no. 4, pp. 644–649, Aug. 1994.
- [38] S. Graja and J.-M. Boucher, "Hidden Markov tree model applied to ECG delineation," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 6, pp. 2163–2168, Dec. 2005.
- [39] A. D. C. Chan, M. M. Hamdy, A. Badre, and V. Badee, "Wavelet distance measure for person identification using electrocardiograms," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 2, pp. 248–253, Feb. 2008.
- [40] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG analysis: A new approach in human identification," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 3, pp. 808–812, Jun. 2001.
- [41] Swathi Banerjee, "Application of cross wavelet transform between normal and abnormal ECG signals analysis and pattern classification", *IEEE Trans. App. Phy.* Feb 2014.
- [42] Brady WJ, Perron AD, Chan T (April 2001). "Electrocardiographic ST-segment elevation: correct identification of acute myocardial infarction (AMI) and non-AMI syndromes by emergency physicians". *Acad Emerg Med* 8 (4): 349–60.
- [43] Loyola University Chicago Stritch School of Medicine. > EKG Interpretive skills Retrieved on April 22, 2010

- [44] "Limb Leads, ECG Lead Placement, Normal Function of the Heart, Cardiology Teaching Package, Practice Learning, Division of Nursing, The University of Nottingham". Nottingham. ac.uk. Retrieved 15 August2009.
- [45] A Real-Time Gabor Primal Sketch for Visual Attention "The Gabor kernel satisfies the admissibility condition for wavelets, thus being suited for multi-resolution analysis. Apart from a scale factor, it is also known as the Morlet Wavelet." Time-Frequency Dictionaries, Mallat
- [46] Rolf G. Winter; Aephraim M. Steinberg (2008). "Coherence". Acce Science. McGraw-Hill.
- [47] J.A. Hartigan (1975). Clustering algorithms. John Wiley & Sons, Inc.
- [48] Vattani., A. (2011). "k-means requires exponentially many iterations even in the plane" (PDF). Discrete and Computational Geometry 45 (4): 596–616.doi:10.1007/s00454-011-9340-1.