2008 International Conference on Emerging Technologies
IEEE-ICET 2008
Rawalpindi, Pakistan, 18-19 October, 2008

Identification of Sudden Cardiac Death Using Spectral Domain Analysis of Electrocardiogram(ECG)

Usman Rashed
Electrical Engineering Department,
University of Engineering and Technology
Taxila, Pakistan
Email: usman_rashed@yahoo.com

Muhammad Javed Mirza
Riphah International University
Rawalpindi, Pakistan

Abstract— **Sudden death from cardiac arrest is a major health problem and is responsible for almost half of all heart disease deaths [1]. This paper introduces work that has been done to distinguish the Electrocardiogram (ECG) of a normal healthy human from that of a patient who may suffer from Sudden Cardiac Death (SCD), but this condition has not been detected. In SCD, the cardiac arrest occurs for a very short time which is preceded and followed by normal ECG (Fig 1). In time domain, detection of such condition would involve monitoring the ECG for over 24 hrs which is not at all feasible. Therefore we worked on normal portion of SCD ECG and compared its parameters with those of a healthy person's ECG. The intention is to design an algorithm that may enable doctors to detect chances of myocardial infarction beforehand on the basis of spectral analysis of an ECG. Fast Fourier Transform (FFT) on QRS complex was used to extract information from the ECG signals providing the basis with which a signal suggesting predisposition of the patient to suffer a cardiac arrest can be differentiated from a normal signal. In this way, instead of waiting for over 24 hrs, 4-5 min. of ECG of any patient is enough to detect possibility of SCD. The algorithm was tested on MIT-BIH (Massachusetts Institute of Technology- Beth Israel Hospital) Databases and the results verified our hypothesis that given an individual's ECG signal during normal function of the heart, it is possible to analyze it and predict whether he is susceptible to cardiac arrest. Further research is being carried out by utilizing the concept for analyzing an ECG signal to identify predisposition to other diseases**

Keywords-ECG; Sudden Cardiac Death; QRS complex

I. INTRODUCTION

An ECG signal is the rhythm generated by heart during its action of pumping blood into the body and is the characteristic of the working condition of heart. Significant work has been done on time domain analysis of ECG [2], however the spectral domain gives a different look into the signal, and parameters here give a unique representation of signal that helps to understand the activity of the heart. Fourier stated that any periodic signal can be represented as sum of sinusoids, each representing a different frequency.

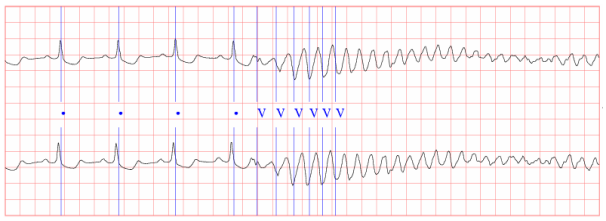

Figure 1 ECG of Ventricular Tachycardia

Since ECG is a periodic signal, its spectrum must contain valuable data. Thus, apart from having effect on time domain portion, every disease should have its effect on the spectrum of the ECG as well.

The above explained concept provided the motivation of performing this work in frequency domain to help enable physicians in future to identify diseases based not only on time domain analysis but spectral analysis as well. Furthermore, future aim is to identify diseases based on spectral domain that are not at all detected in time domain.

ECG signal consists of 3 significant portions, P-wave, QRS complex and T-wave (Fig 2)**.** It has been shown that significant information is present in QRS complex [3] and QRS complex plays a significant role in heart diseases [4], so it was obvious that implementation of the algorithm, for a start, should be on QRS complexes.

The Fast Fourier Transform (FFT) Model of the signal utilizing the Discrete Fourier Transform (DFT) technique was considered for analysis. This technique is most widely used for determining the Fourier Transform of a signal and is efficiently calculated using Matlab.

Sudden Cardiac Death (SCD), or cardiac arrest, is the sudden, abrupt loss of heart function in a person who may or may not have a diagnosed heart disease. It occurs instantly or shortly after symptoms appear. These events are most often initiated with symptoms including *ventricular tachyarrhythmia*. In this, the heart starts to beat too rapidly

978-1-4244-2211-1/08/$25.00 ©2008 IEEE

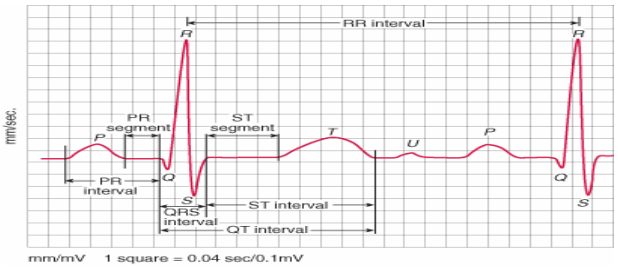

Figure 2: P, QRS and T wavelets

then returns back to normal sinus rhythm of ECG. Our aim is to analyze this normal rhythm of such an ECG of a patient who is prone to cardiac death if this condition is sustained. The ECG samples from MIT-BIH Database (DB) are used for this purpose. Fig 1 shows the onset of ventricular tachycardia (annotated by v).

Proposed hypothesis states that the activity of the heart should be affected by symptoms such as ventricular tachycardia and apart from having its effects during irregular beat; it should have its after effects during normal sinus rhythm as well. Since frequency is defined as rate of change, a change in activity of heart means that the significant spectrum of the signal should be affected as well.

By calculating the Fourier Transform and threshold the signal for the significant lobe of spectrum, different parameters were calculated that represent the signal. Signal was normalized in frequency domain. The parameters were also plotted using two dimensional plots to show significant difference in the desired and pathological parameters. A number of signals were tested upon to validate our hypothesis.

II. METHODOLOGY

The ECG signals were obtained from the MIT-BIH DB via the Physionet [5] web site. The MIT-BIH DB contains many data sets of electrocardiogram signals, mostly abnormal or unhealthy electrocardiograms but some are normal that can be used as a reference. Two databases were selected for analysis:

Normal Sinus Rhythm (NSR) Database

Sudden Cardiac Death (SCD) Database

Signals from each of the above databases were selected and divided into two groups of four signals each. One group was used for analysis and the other for test. The distribution is given in Tables 1 & 2.

The ECG signals in the databases were long, some up to 30 minutes in length, so small sections of these signals were obtained and QRS complexes separated for analysis.

TABLE 1 SUDDEN CARDIAC DEATH

Group	Patient 1	Patient 2	Patient3	Patient 4
1 (Analysis)	Pt: 33	Pt: 39	Pt: 41	Pt:52
2 (Test)	Pt:30	Pt:31	Pt: 35	Pt:46

TABLE 2 NORMAL SINUS RHYTHM

Group	Patient 1	Patient 2	Patient 3	Patient 4
1(Analysis)	Pt:16272	Pt:16273	Pt:16773	Pt:16539
2 (Test)	Pt:16483	Pt:16795	Pt:16786	Pt:16420

A. *Procedure*

Each signal was pre-processed before analysis. Sinus Rhythm signals were originally sampled at 128Hz and Cardiac Death signals at 256Hz. In order to apply the signal processing techniques equally to each record signal and preserve most information, each signal was re-sampled to 500Hz [6]. After re-sampling, the mean was removed from each ECG signal and these were passed through a Butterworth, 3rd-order low-pass filter to remove any signal components beyond 32 Hz.

Since there were slight variations in QRS complexes obtained from the same Database files, it was decided to average out the samples to obtain one representative QRS complex for each file. For this, three QRS samples were taken each from of the following intervals: 0-20 seconds, 500-520 seconds and 1000-1020 seconds of the signals to obtain, in all, 9 samples of QRS from each ECG data file. The lengths of these samples were adjusted using zero padding. These 9 samples were averaged out to determine a single QRS complex representative of that ECG data file.

B. *Discrete Fourier Transform*

Discrete Fourier Transform (DFT) is one of the specific forms of Fourier Analysis. As such, it transforms one function into another, which is called the frequency domain representation of the original function (usually time domain). But DFT requires an input function that is discrete. Such inputs are often created by sampling a continuous function.

A sequence of N complex numbers x_0, x_{N-1} can be transformed into a sequence of N complex numbers $X_0,\ldots X_{N-1}$ by the DFT according to the formula [7]:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} \qquad k = 0, \ldots, N-1 \qquad (1)$$

where $e^{\frac{2\pi i}{N}}$ is a primitive N'th root of unity.

The complex numbers X_k represent the amplitude and phase of the different sinusoidal components of the input "signal" x_n. The DFT computes the X_k from the x_n. The data samples available on MIT database are already in a sampled form so no sampling was needed. However, as stated above, re-sampling was done to apply techniques equally on all signals.

The indexes 'k' in DFT were converted to analog frequency via the following transformation:

F = (1:L_Y) * Sfecg/N; Where

F = Analog frequency

L_Y = Length of frequency domain signal

Sfecg = Sampling Frequency

N^* = No. of points used for DFT computations

** Should be a power of 2*

A minimum sufficient value of N=2048 was used in analysis. Since DFT spectra repeat, only first half of coefficients were used for further analysis.

C. *Frequency Domain Segmentation*

The Time and Frequency Domain QRS complex is plotted in Figure 3. The spectrum was divided into following regions for parameter comparison.

Region 1 (R1)	0-2Hz
Region 2 (R2)	2-8Hz
Region 3 (R3)	8-16Hz
Region 4 (R4)	16-22Hz
Region 5 (R5)	22-28Hz

Spectra in these regions were plotted to gain knowledge regarding the variation and energy localization. Then parameters were selected to represent the spectra in the respective region including Energy, Power, Standard Deviation (S.D), Mean and Median. The parameters in each of the regions were observed and comparisons were made between the NSR and SCD DB. Tables 3-10 give the comparison between NSR and SCD spectra parameters. An algorithm was developed to classify ECG files on the basis of information obtained from above region parameters. Test files were used to verify the proposed classification algorithm.

From the normalized spectra plot (Fig 3), it is visible that the significant values tend to lie in below 50 Hz.

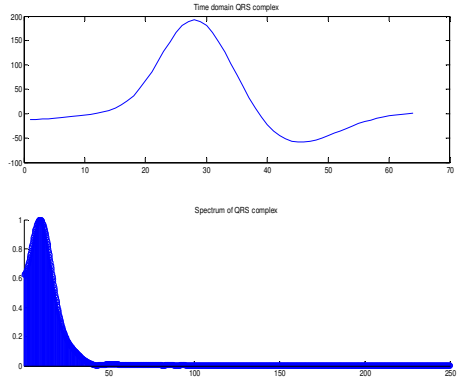

Figure 3: Time and Frequency Domain QRS Complex of Normal Sinus Rhythm (16272)

D. *First Lobe Analysis*

Careful observation of the spectra indicates that most energy of the signal lies in the first lobe of the plotted spectra within a threshold of 0.1 in the normalized plotted figure (Fig 3). So this portion was separated for further analysis. Once the first lobe coefficients were obtained, various parameters were used to represent the signal in that lobe. The parameters in C, along with some other parameters, were calculated. Most significant of these was the frequency range of the First Lobe of spectra.

From the parameters extracted from the signal, those showing significant variations between the two signals (called feature values) were selected. These values were then placed inside a vector called the feature vector, F.

$F_i = [p_1, p_2 \ldots . p_n]$; Where

F_i= feature vector of ith signal

p_n =nth features of ith signal

Since more than one signal was used for analysis, a mean feature vector was calculated. This vector completely represented one type of ECG signal. Similarly, a feature vector is calculated for each unknown signal that we wish to classify by determining the appropriate parameters of that instance. In order to classify a signal, the distance between the two feature vectors is calculated. This is accomplished by computing the Euclidean distance between mean feature vector, u, and the unknown signal feature signal, v, as follows:

$$d(u,v) = \sqrt{\sum_{i=1}^{N}(v_i - u_i)^2} = \sqrt{(v_1-u_1)^2 + (v_2-u_2)^2 + \cdots + (v_N-u_N)^2}. \quad (2)$$

Where

N = No. Of feature values

d (u, v) = Euclidean Distance

Euclidean distance of unknown signal from both NSR feature vector and SCD feature vector is calculated and classification done according to the nearness of signal to a particular signal.

III. RESULTS

According to the analysis techniques discussed in Section C, different spectral features were determined for both types of signals. The results obtained are given in Tables 3-10.

Normal Sinus Rhythm (NSR) Parameters

TABLE 3 : PATIENT 16272

Region	Power	Energy	S.D	Mean	Median
R1	0.2628	2.1026	0.0421	0.5133	0.5084
R2	0.3635	9.0874	0.4134	0.6860	0.7987
R3	0.1758	5.7998	0.6376	0.6308	0.9251
R4	0.1024	**2.5606**	0.2534	0.3354	0.3851
R5	0.0061	0.1525	0.0813	0.0720	0.0872

TABLE 4: PATIENT 16273

Region	Power	Energy	S.D	Mean	Median
R1	0.3366	2.6932	0.0922	0.5862	0.5873
R2	0.2795	6.9872	0.6279	0.6678	0.8312
R3	0.1522	5.0213	0.6376	0.5918	0.8824
R4	0.0930	**2.3262**	0.2071	0.3227	0.3558
R5	0.0095	0.2374	0.0702	0.0901	0.0913

TABLE 5: PATIENT 16773

Region	Power	Energy	S.D	Mean	Median
R1	0.0085	0.0677	0.0604	0.1065	0.1195
R2	0.0939	2.3486	0.3157	0.2414	0.3189
R3	0.1914	6.3163	0.6514	0.6643	0.9452
R4	0.2529	**2.3872**	0.3600	0.5537	0.6394
R5	0.0357	0.8922	0.1383	0.1919	0.2071

TABLE 6: PATIENT 16539

Region	Power	Energy	S.D	Mean	Median
R1	0.2628	2.1026	0.0345	0.5133	0.5084
R2	0.3635	9.0874	0.3887	0.6860	0.7987
R3	0.1758	5.7998	0.6723	0.6308	0.9251
R4	0.1024	**2.5606**	0.2964	0.3354	0.3851
R5	0.0061	0.1525	0.1200	0.0720	0.0872

Sudden Cardiac Death (SCD) Parameters

TABLE 7: PATIENT 39

Region	Power	Energy	S.D	Mean	Median
R1	0.1093	0.8742	0.0719	0.3246	0.3114
R2	0.2090	5.2250	0.5843	0.6014	0.8659
R3	0.1419	4.6811	0.5635	0.3978	0.6546
R4	0.0066	**0.1649**	0.0426	0.0743	0.0649
R5	0.0011	0.0286	0.0172	0.0348	0.0313

TABLE 8: PATIENT 52

Region	Power	Energy	S.D	Mean	Median
R1	0.4895	3.9158	0.2316	0.7311	0.7548
R2	0.0298	0.7462	0.7450	0.5682	0.9319
R3	0.0819	2.7035	0.2689	0.2212	0.1052
R4	0.0027	**0.0678**	0.0228	0.0475	0.0430
R5	0.0008	0.0201	0.0045	0.0283	0.0282

TABLE 9: PATIENT 41

Region	Power	Energy	S.D	Mean	Median
R1	0.8585	0.8677	0.2744	0.9604	0.9945
R2	0.1611	4.0269	0.6332	0.4658	0.7905
R3	0.0131	0.4321	0.1654	0.0527	0.1033
R4	0.0003	**0.0068**	0.0282	0.0053	0.0267
R5	0.0001	0.0025	0.0033	0.0100	0.0104

TABLE 10: PATIENT 33

Region	Power	Energy	S.D	Mean	Median
R1	0.8477	6.7814	0.2350	0.9449	0.9762
R2	0.2154	5.3852	0.4271	0.5140	0.6484
R3	0.0097	0.3208	0.1018	0.0421	0.0605
R4	0.0016	**0.0401**	0.0197	0.0395	0.0440
R5	0.0006	0.0145	0.0224	0.0238	0.0281

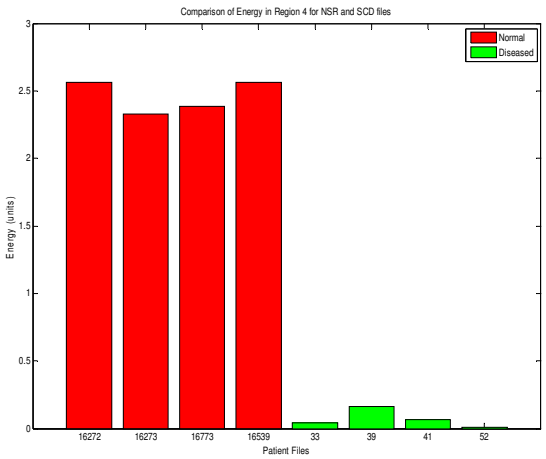

Figure 4: Comparison of Energies of NSR and SCD files in Region 4

Fig 4 shows the plot, in the form of bar chart, of the energies in Region 4 for each of the analysis files. The red colored bars indicate the energies for NSR files while green bars indicate energies for SCD files. Clear distinction between the two sets of files is visible.

Similarly, the First Lobe technique discussed in section D was used on the spectra of the two diseases and the resulting plot is shown in Fig 5 and 6. Clear distinction in frequency range covered is observed between the plots on the basis of which classification can be done.

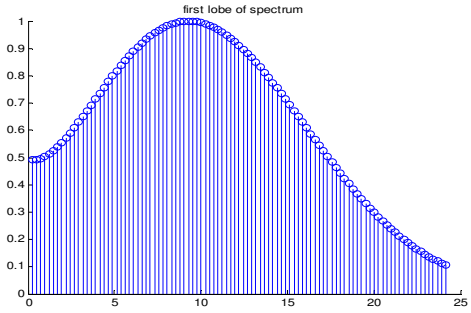

Figure 5 First Lobe Plot of Normal Sinus Rhythm Signal (16272)

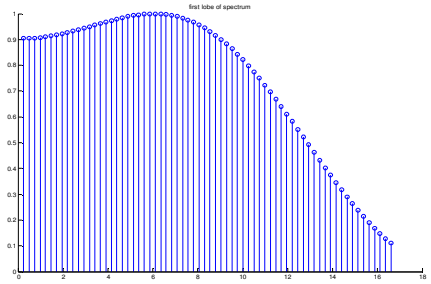

Figure 6: First Lobe Plot of Sudden Cardiac Death Signal (52)

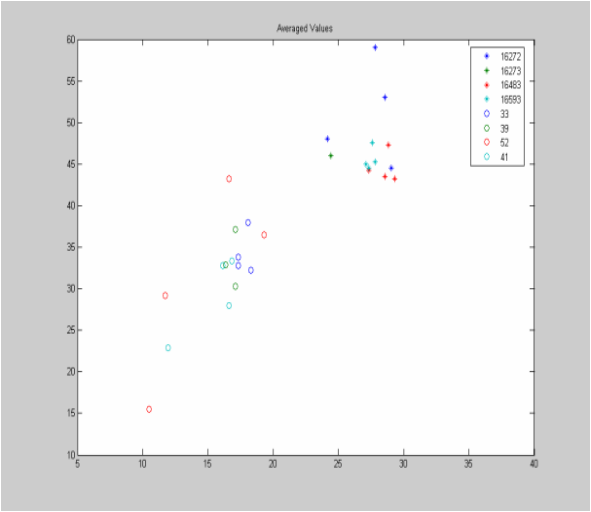

Figure 7: Accumulated plot of Signal Features of NSR and SCD signals

The parameters obtained in First Lobe were studied and two parameters, Energy and Frequency Range, were plotted against each other for each of the eight test files, in Fig 7. The asterisks (*) show the accumulation of NSR features while circles indicate the features obtained from SCD.

TABLE 11: RESULTS OF SCD TEST FILES

Patient ECG file #	Euclidean Distance from Normal Sinus Rhythm Feature Vector (D_N)	Euclidean Distance from Sudden Cardiac Death Feature Vector (D_SDC)
30	25.2	9.6
31	11.6	8.6
35	26.6	6.6
46	12.8	7.3

TABLE 12: RESULTS OF NSR TEST FILES

Patient ECG file #	Euclidean Distance from Normal Sinus Rhythm Feature Vector (D_N)	Euclidean Distance from Sudden Cardiac Death Feature Vector (D_SDC)
16483	2.5	17.7
16795	5.2	15.4
16786	3.2	22.1
16420	6.5	26.4

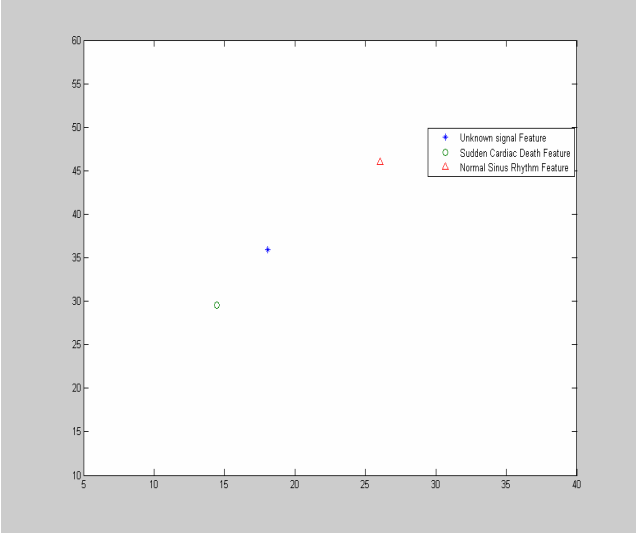

Figure 8: Test results of a specimen file

Next, the test signals were tested using algorithm developed involving calculation of Euclidean distance of the feature vector. The results of these are tabulated in Tables 11 & 12.

Figure 8 shows plot of one of the results.

IV. CONCLUSION

The Data in Tables 3-10 indicates that there is significant energy difference between the two files in Region 4 (R4). The highlighted values in the tables indicate that for all SCD files, the energy is significantly below 1 unit in R4. However significant high energy (above 1 unit) is present in R4 region for the NSR database files and this feature was used for classification of the files. This hypothesis was put to the test and all the test files for SCD and all but one for NSR tested positive for this hypothesis. File 16795 had energy less than 1 unit (0.98) but that can be approximated to 1.These results are further verified in bar chart of Fig 4 that shows clear discrimination between the two sets of files. So this feature can be utilized to discriminate between the two files, and hence identify the SCD case.

In First Lobe analysis, Energy and Frequency Range were selected as feature values and were plotted in Fig 7. Clear distinction is shown by the plot. The Euclidean Distance from both feature vectors was calculated for all the four test files of SCD and NSR and the results in Table 11&12 and Fig 8 verifies the classification technique. All the SCD test files lie nearer to the mean feature vector calculated from SCD analysis files (Table 11). Same holds true for NSR files in Table 12. Therefore, a simple comparison of the distances calculated will classify the signals.

Hence it was discovered, that the normal ECG rhythm of patient suffering from Sudden Cardiac Death has:

1) Lower Spectral Energy.

2) Low Frequency Range of First Lobe.

3) Negligible Energy in Region 4 (16-22 Hz).

Classification can be done on basis of these three findings.

As mentioned above, manual segmentation of QRS complexes was done, which may not be efficient. The process of devising an algorithm for automation is currently under progress.

REFERENCES

[1] http://www.webmd.com/heart-disease/guide/sudden-cardiac-death

[2] Morlet, D., et al. "Time-scale analysis of high-resolution signal-averaged surface ECG using wavelet transformation."., Proceedings of Computers in Cardiology , pp. 393-6, Italy, 23-26 Sep 1991

[3] Waujgh RA, Wagner GS, Haney TL, Rosati RA, Morres JJ: "Immediate and remote prognostic significance of fascicular block during acute myocardial infarction". American Hearts Association Circulation 47: 765, 1973

[4] Amir Kashani and S. Serge Barold**,** "Significance of QRS Complex Duration in Patients With Heart Failure", Journal of The American College of Cardiology, 2005

[5] http://physionet.org/physiobank/database/mitdb

[6] T. Srikanth, S. Napper, and H. Gu, "Bottom-Up Approach to Uniform Feature Extraction in Time and Frequency Domains for Single Lead ECG Signal," Journal of the International Society for Biolectromagnetism,vol. 4, no. 1, 2002.

[7] Oppenhiem & Wilsky, " Signals & Systems", 2nd Ed.