Analysis of ECG Signals

Apurva Kulkarni, Snehal Lale, Pranali Ingole, Sayali Gengaje
Student, Electronics and Telecommunication Department,
Vishwakarma Institute of Technology
Pune, Maharashtra, India.

Abstract :
Electrocardiogram (ECG) Feature Extraction plays a significant role in diagnosing most of the cardiac diseases. ECG is a non-linear, non-stationary signal. It is used for the primary diagnosis of heart abnormalities. The ECG signals were taken from MIT-BIH arrhythmias database for analysis. Noises in the ECG signal such as powerline interference, baseline wandering and muscle noises were removed using bandpass filter. Different statistical and morphological features were extracted for both normal as well as abnormal cases. These features include R-R interval, heart rate arithmetic mean, median, variance, skewness, kurtosis etc. The values of the feature vector reveal the information regarding status of cardiac health. Platform used for implementation is MATLAB. This paper includes comparative analysis of various transform-based techniques. For differentiating normal and abnormal signals, we have used KNN classifier. We have achieved an accuracy of 86.95%, sensitivity of 87.09% and specificity of 86.66% for 60% training dataset using this classifier.

Keywords - *Electrocardiogram* **(***ECG), Discrete Wavelet Transform (DWT), KNN Classifier*

I. INTRODUCTION

Electrocardiogram (ECG) is the graphical recording of the electrical activity of the heart, and is the most important physiological parameter that gives the correct assessment regarding the functioning of the heart. In general, the cardiologist looks critically at different time intervals, polarities and amplitudes of the ECG to arrive at the diagnosis.

A large amount of data produced by ECG monitoring and recording facilitates the need for analysis and classification of normal and abnormal conditions.

The methods used for diagnosis generally involve four essential processes to arrive at the accurate and quick decisions about the kind of heart disease a patient suffers from. It includes:
1. Data Acquisition
2. Denoising
3. Feature Extraction
4. Classification.

A large number of techniques exist in the literature for the analysis of ECG signals. Prashanth Shetty et.al. presented a comparative study of various analog filtering methods for noise removal [3]. Pan and Tompkins designed a special digital bandpass filter for noise removal and developed real time QRS complex detection [6]. Dangi et.al. compared various transform based techniques for compression of ECG signal[2]. Saini et.al. presented an application of KNN algorithm as a classifier for QRS complex detection[7].

The database of 46 patients has been collected from MIT-BIH arrhythmias database[1] with lead II (ML II),sampled at 360 samples per second with 11 bit resolution recorded at Beth Israel Laboratory.

II. NOISES AND ARTIFACTS IN ECG WAVEFORM

Raw ECG data contains some noise and artifact components that alter the shape of the ECG trace from the ideal structure, which render the clinical interpretation inaccurate and misleading; consequently, a pre-processing step for improving the signal quality is a necessity. The types of noises contaminating ECG signal include Power line Interference, Baseline Wandering, Muscle Tremor noise, Electrosurgical noise, Instrumentation noise and other less significant source of noise [3]. Among these noises, the power line interference and the baseline wandering are the most significant and can strongly affect ECG signal analysis. It is essential to reduce disturbances in ECG signal and improve the accuracy and reliability for better diagnosis. We have used Pan Tompkins algorithm, which removes baseline wander, power line noise and muscle noise in linear filtering stage using bandpass filter having passband of 5-12 Hz [6].

III. FEATURE EXTRACTION

A. Statistical Features

Mean, mode, median, standard deviation are first order statistical features. Variance, kurtosis, skewness are higher order statistical features. Standard Deviation gives the measure to quantify the amount of variation or dispersion of a set of data values. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. Data with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly and have heavy tails. Skewness indicates asymmetry and deviation in distribution analysis from a normal distribution.

B. Morphological Features

The Feature Extraction stage extracts diagnostic information from the ECG signal. The detection of R peak is the first step of feature extraction. The R peak in the signal has the largest amplitude among all the waves compared to other leads. Pan Tompkins algorithm was used to detect the R-R interval.

C. Transform Based Features

ECG signals are in time-domain in their raw format. The pathological conditions, however, may not always be quite obvious in the original time-domain signal. These conditions can sometimes be diagnosed more easily when the frequency contents of the signal are analyzed.

DCT, DST, FFT and Walsh-Hadamard transforms give frequency-amplitude representation of the signal. Wavelet transform is capable of providing the time and frequency information simultaneously, hence giving a time-frequency representation of the signal [8].The discrete wavelet transform (DWT), provides sufficient information for both analysis and synthesis of the original signal, with a significant reduction in the computation time. DWT can be used in extracting features of ECG signal. It also helps in denoising and compression of ECG signal without any loss in clinical information.
We have decomposed the ECG signal up to 4 levels using biorthogonal 4.4 wavelet family.
In the feature extraction stage, a feature vector of 46x32 dimension is obtained. Feature set includes morphological as well as statistical features.

IV. KNN CLASSIFIER

K-Nearest Neighbor (KNN) classification is a very simple, yet powerful classification method. It is also called as memory based classification as the training samples need to be in the memory at run time. It became popular due to its simplicity and relatively high convergence speed. The key idea behind KNN classification is that similar observations belong to similar classes. Thus, one simply has to look for the class designators of a certain number of the nearest neighbors and weigh their class numbers to assign a class number to the unknown.

When dealing with continuous attributes the difference is calculated as the Euclidean distance. If the first instance is (a_1, a_2, a_3… a_n) and the second instance is (b_1, b_2, b_3…b_n) the distance between them is calculated by the following formula:

$$D = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + ..(a_n - b_n)^2}$$

Continuous attributes are normalized to have the same influence on the distance measured between instances. The weighing scheme of the class numbers is often a majority rule, but other schemes are conceivable. The number of the nearest neighbors, k, should be odd in order to avoid ties, and it should be kept small, since a large k tends to create misclassifications unless the individual classes are well-separated[7]. The performance of a KNN classifier is always at least half of the best possible classifier for a given problem. One of the major drawbacks of KNN classifiers is that the classifier needs all available data. This may lead to considerable overhead, if the training data set is large.

We have compared the classified database by taking different percentage of signals in training dataset and calculated the performance parameters.

V. RESULTS

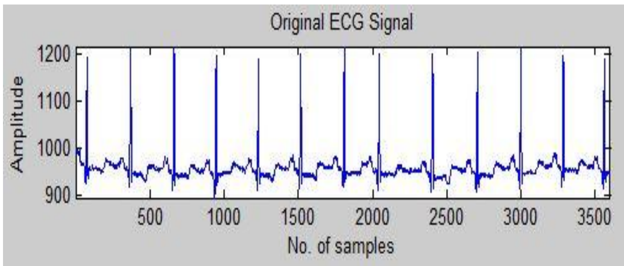

Fig. 1 Original ECG Signal (100m)

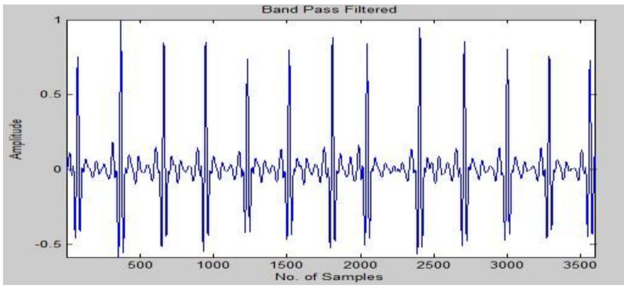

Fig. 2 Filtered Signal

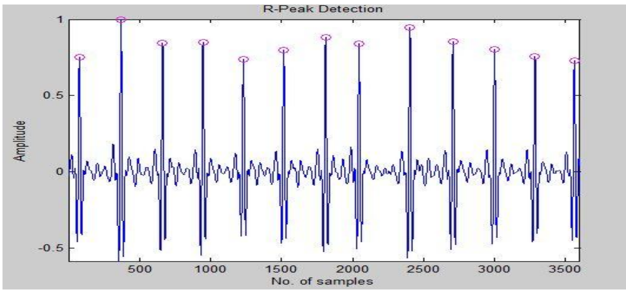

Fig. 3 R Peak Detection

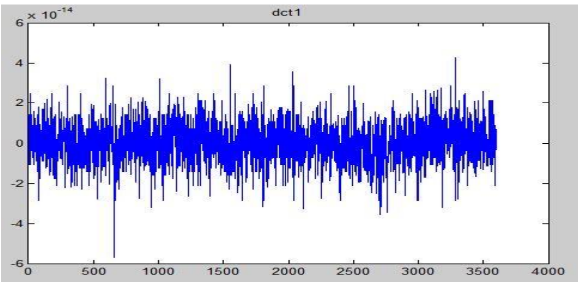

Fig. 4 Reconstruction error for DCT-I

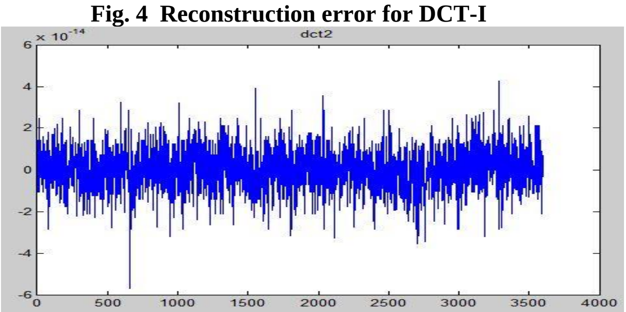

Fig. 5 Reconstruction Error for DCT-II

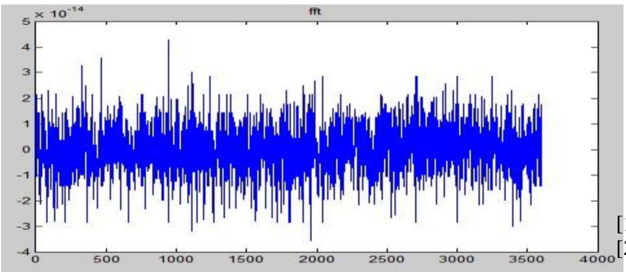

Fig. 6 Reconstruction Error for FFT

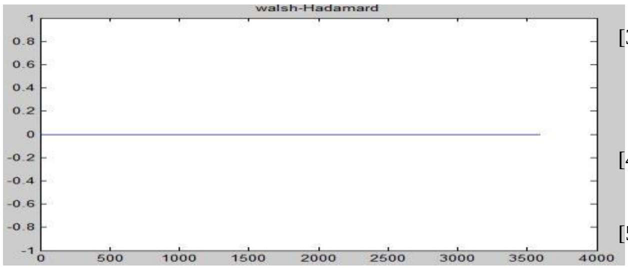

Fig. 7 Reconstruction Error for Walsh-Hadamard

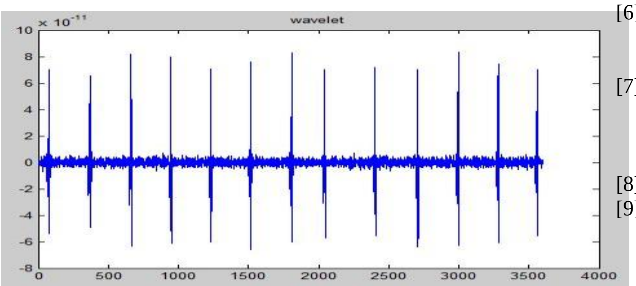

Fig. 8 Reconstruction Error for Wavelet

Table I : Comparison of Performance Parameters of Classifier

Training dataset (%)	Accuracy	Sensitivity	Specificity
40	78.26	85.18	68.42
50	84.78	86.66	81.25
60	**86.95**	**87.09**	**86.66**
70	89.13	85.29	100

VI. CONCLUSION

In this paper, R peak detection was performed using Pan Tompkins algorithm. After comparative analysis of transforms, we found that Walsh-hadamard method gives zero distortion and Wavelet transform gives distortion in the range of 10^{-11} after reconstruction of signal. The signals were decomposed up to 4 levels using biorthogonal 4.4 wavelet family.

KNN classifier was applied for classification of ECG arrhythmias. We have used 46 recordings from the MIT-BIH arrhythmias database for testing in the classifier. Highest sensitivity of 87.09% is achieved for 60% training dataset.

REFERENCES

[1] www.physionet.org
[2] Dangi, Vikramaditya, Amol Parab, and Kshitij Pawar. "Analysis of Transform Based ECG Compression Techniques." Proc. of the International Conference on Advances in Electronics and Electrical Engineering–AEEE. 2012.
[3] Shetty, Prashanth, and Somashekhara Bhat. "Analysis of Various Filter Configurations on Noise Reduction in ECG Waveform." International Journal of Computing, Communications and Instrumentation Engg., Vol. 1, Issue 1 (2014).
[4] Saritha, C., V. Sukanya, and Y. Narasimha Murthy. "ECG signal analysis using wavelet transforms." Bulg. J. Phys 35.1 (2008): 68-77.
[5] Rai, Hari Mohan, and Anurag Trivedi. "De-noising of ECG Waveforms based on Multi-resolution Wavelet Transform." International Journal of Computer Applications (0975–8887) Volume (2012).
[6] Pan, Jiapu, and Willis J. Tompkins. "A real-time QRS detection algorithm" Biomedical Engineering, IEEE Transactions on 3 (1985): 230-236.
[7] Saini, Indu, Dilbag Singh, and Arun Khosla. "QRS detection using K-Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases" Journal of advanced research 4.4 (2013): 331-344.
[8] Robi Polikar, "The Wavelet Tutorial"
[9] Minakshi Chaudhary, D.V.Saini, Sukhvinder Kaur. "Entropic method of measurement of heart rate in 12 lead ECG using Artificial Neural Network" SSRG International Journal of Electronics and Communication Engineeering – EFES April 2015