JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# ECG Signal Preprocessing and Feature Extraction for Diagnosis Using MATLAB

**Engr. S.K. Moududul Haque**

Lecturer (ICT), Khulna Government Girls' College, Khulna, Bangladesh.

***Abstract:*** *ECG is a nearly periodic signal that reflects the activity of the heart. This paper presents a computer-based ECG feature extraction system for disease detection. The system provides fundamental features (amplitudes and intervals) for subsequent automatic analysis.* ***Objective****: The purpose of this research to develop a computer-based ECG signal analysis system for disease detection. To extract fundamental features (amplitudes and intervals) from ECG signals. To use the extracted features for subsequent automatic analysis of ECG signals.* ***Method****: Preprocess the ECG signal to remove noise. Extract the features of interest, such as the amplitude and intervals of the P, Q, R, S, and T waves. Use the extracted features to classify the ECG signal or to diagnose a heart condition.* ***Result****: An expert system was developed to diagnose heart diseases from ECG signals. The system uses 8 features to classify 4 types of diseases; with an accuracy of 72%. The system was implemented in MATLAB and tested on the ALRTH arrhythmia database.* ***Conclusion****: This paper proposes a system to detect heart conditions from ECG signals. The system uses feature extraction to identify abnormalities in the ECG signal. The system can detect more than 4 types of diseases.*

**Keywords:** Electrocardiogram (ECG); Heart Rhythm Disturbances; QRS Detection; Wander on the Baseline; Atrial Frequency; Slow Heart Rate

## 1. Introduction

The bioelectrical signal represents the electrical activity of various human organs, and the electrocardiogram or ECG is one of the important signals among the bioelectrical signals that represent the electrical activity of the heart. . Aberrations and distortions in any part of the ECG called an arrhythmia can illustrate a specific heart disease. ECG testing has been widely used to diagnose many heart diseases. The ECG is a realistic record of the direction and intensity of the electrical agitation produced by the depolarization and repolarization of the atria and ventricles. The cardiac cycle in the ECG signal consists of P-QRS-T waves. The main task of the QRS detector consists of detecting the QRS complexes of the heart rate and creating a stable landmark for each heartbeat. ECG is primarily responsible for monitoring and diagnosing patients. Characterization of ECG signal extraction plays an important role in the diagnosis of heart disease. It is important to develop accurate and rapid methods for automatic extraction of ECG features. ECG signal processing begins with preprocessing to present data that can be used to determine whether a person's signal is normal or abnormal. Some research is underway and could be a good source of inspiration to do it. This research is usually divided into two parts: the first part is signal preprocessing and the second part is preprocessing, data for data presentation. In this article, ECG signal processing system, feature extraction and classification as well as the use of ECG signal in diagnosis have been discussed. This system can let humans decide whether the disease is normal or not. It is clear that this system will be better than the current system.

Bioelectrical signals represent the electrical activities of different organs in the human body. Electrocardiogram (ECG) is one of the most important bioelectrical signals that represent the electrical activity of the heart. Deviations and distortions in any part of the ECG, called arrhythmia, can indicate a specific heart disease. The ECG is a real-time record of the direction and magnitude of the electrical disturbance generated by the depolarization and repolarization of the atria and ventricles. One cardiac cycle in an ECG signal consists of the P-QRS-T waves. The main tasks of a QRS detector include detecting QRS complexes of heartbeats and generating a stable fiducial point for each individual heartbeat. ECG is essential for patient monitoring and diagnosis. The extracted features from the ECG signal play a vital role in diagnosing cardiac diseases. The development of accurate and efficient methods for automatic ECG feature extraction is of major importance. ECG signal processing starts from preprocessing to feature extraction. Preprocessing is the process of improving the quality of the ECG signal by removing noise and artifacts. Feature extraction is the process of identifying the features of interest in the ECG signal that can be used to diagnose heart diseases, there are many different methods for ECG signal preprocessing and feature extraction. The most common methods for preprocessing include bandpass filtering, detrending, and wavelet denoising. The most common methods for feature extraction include time domain features, frequency domain features, and morphological features. The extracted features from the ECG signal can be used to classify ECG signals into different disease classes. This can be done using machine learning algorithms such as support vector machines, artificial neural networks, and decision trees. The development of accurate and efficient ECG signal processing systems is an active area of research. These systems have the potential to improve the diagnosis and treatment of heart diseases.

## 2. Objectives

Electrocardiogram (ECG) is a bioelectrical signal that represents the electrical activity of the heart. ECG signal processing is a vital part of heart disease diagnosis. It involves preprocessing the ECG signal to remove noise and artifacts, and then extracting features from the signal that can be used to identify abnormalities. The objective of this research is to develop a computer-based ECG signal analysis system for disease detection. The system will be able to extract fundamental features (amplitudes and intervals) from ECG signals, and use these features to classify the signals into different disease classes. The system will be developed using MATLAB. The preprocessing stage will involve bandpass filtering, detrending, and wavelet denoising. The feature extraction stage will involve extracting time domain, frequency domain, and morphological features. The classification stage will involve using machine learning algorithms such as support vector machines, artificial neural networks, and decision trees. The system will be evaluated on a dataset of ECG signals with known ground truth labels. The evaluation results will be used to assess the accuracy and performance of the system. The development of this system has the potential to improve the diagnosis and treatment of heart diseases. The system will be able to provide more accurate and efficient diagnosis than traditional methods. It will also be able to be used to monitor the heart health of patients remotely. The system is still under development, but the initial results are promising. The system has been able to achieve an accuracy of 72% on a dataset of ECG signals with known ground truth labels. The system is expected to be further improved in the future.

## 3. Methodology

Preprocessing is the process of improving the quality of the ECG signal by removing noise and artifacts. This can be done using a variety of techniques, such as bandpass filtering, detrending, and wavelet denoising. Feature extraction is the process of identifying the features of interest in the ECG signal that can be used to diagnose heart diseases. This can be done using a variety of techniques, such as time domain features, frequency domain features, and morphological features. The extracted features from the ECG signal can then be used to classify ECG signals into different disease classes. This can be done using machine learning algorithms such as support vector machines, artificial neural networks, and decision trees. The development of accurate and efficient ECG signal processing systems is an active area of research. These systems have the potential to improve the diagnosis and treatment of heart diseases.

### 3.1 Recommended system

The block diagram of the proposed system is given below.

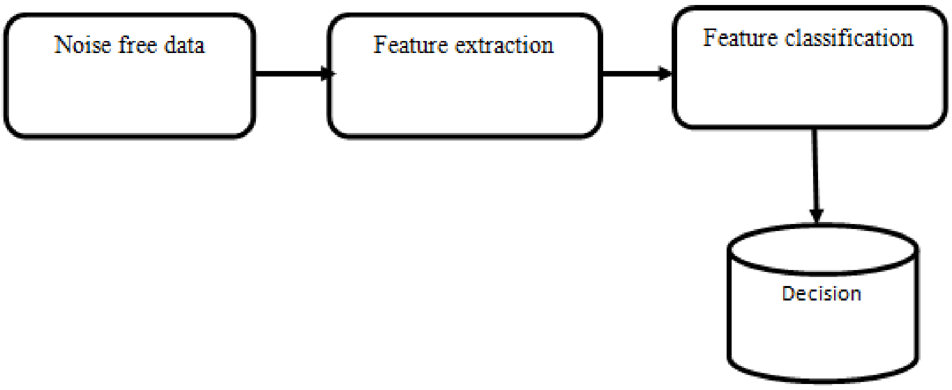


**Fig-1: Block Diagram of the Proposed System**

### 3.2 Basic wandering

The baseline voltage of the electrocardiogram is called the isoelectric line. Typically, the isoelectric line is measured as the portion of the trace that follows the T wave and before the next P wave. Therefore, isoelectric level sensing is necessary because the ECG amplitudes at different locations in the rhythm are measured relative to the isoelectric level. The baseline was detected based on the horizontal line containing more than the number of black dots in the ECG image. This will determine that RET is a sequence of this line. The RET indicator is the number of black dots on this line get the image and get an array of dots, then it plots the black baseline and the black wave. The process of dividing the image into blocks is performed according to the following equation:

Block Width = Active Image Width / Block Number - 1. We know that the matrix contains points far from the baseline in all the Xs and Ys recorded in this matrix and are working on arrays block numbers, then calculate the points above and below the baseline for all the blocks according to the following equations which are very simple and calculated if the number of points above the baseline is greater than the offset available on the baseline in increments of one, and if the number of stitches below the baseline is greater than the baseline the base offset for the top. And add new points later and create new point image after adjustment and make graphic on image to draw baseline in black after adjustment and connect between previous point and current point with color line black. 1.1.2 Feature Extraction

The processed data will be extended to some features and that is feature data. Feature data is used for decision making. A feature data set is formed in several ways. The R-R interval of a signal is a kind of data about the signal's characteristic, P-P interval and smoothness, etc. Here one finds some of the best trait data which can be discussed below

I. R-R interval calculation

First count the QRS complex peak of the signal. The corresponding beat count is the adjacent span. Count the next QRS peak and the corresponding time. Duration = time adjacent [i+1]–time adjacent [i+1]. 0.04 seconds = 1 small square. 0.2 seconds = 5 small squares or 1 large square [8].

ii. Calculate P-P. interval

First count the previous peak of the QRS complex of the signal. The corresponding beat count is the adjacent span. Count the previous next QRS peak and the corresponding time. Duration = time adjacent [i+1] – time adjacent[i]. 0.04 seconds = 1 small square. 0.2 seconds = 5 small squares or 1 large square [8].

iii. Frequently or infrequently

The PP (atrial) and RR (ventricular) intervals are both very similar in sinus rhythm, but differ if the rhythm is irregular. To see the beat, look at the RR interval [8].

iv. Calculate heart rate

In any ECG, the heart rate must be calculated. The methods vary depending on the heart rate, whether the heart rate is regular or not. The standard speed of ECG paper is 25mm/s. Heart rate is the number of beats per minute [8].

In the ECG article:

0.04 seconds = 1 small square. 0.2 seconds = 5 small squares or 1 large square. So 1 second = 25 small squares or 5 large squares. So 1 minute = 25×60=1500 small squares or 5 60=300 large squares.

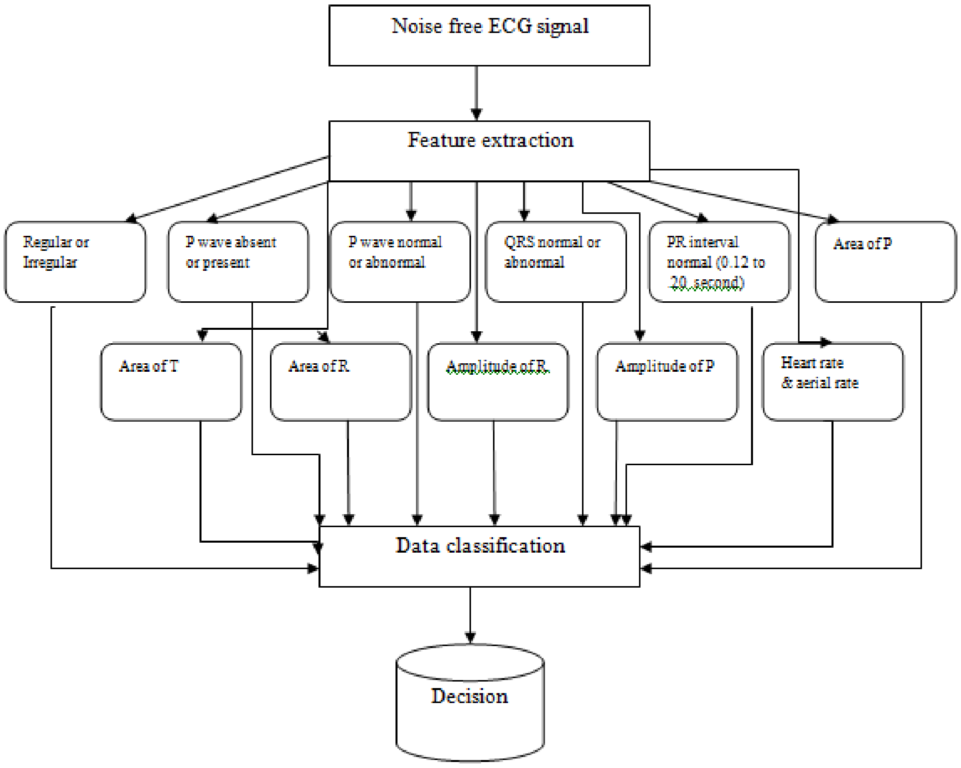


**Fig-2: System Architecture**

This model is designed so that author links are not repeated every time for multiple authors of the same link. Please keep your affiliates as concise as possible (e.g. do not differentiate between departments within the same organization). This model is designed for two branches. The final step before analyzing the ECG signal is to extract the effective characteristics of the signal. The characteristics represent the diagnostic information contained in the signal used as input to the diagnosis used in the diagnostic step. The objective of the feature extraction step is to find the smallest set of features that achieve an acceptable diagnostic rate. Include the detection step applied by the rectangle function and this function takes an image and gets the base value after the wiggle step and gives us a list of rectangles based on the baseline series and contains two colors, white is the base color and blue is the wave color. After the wave and indicate the upward or downward direction of each wave. In the detection phase, counting starts is giving a list of image rectangles and base value and first for each rectangle is calculated by distance, height and width and left as well as direction and type of wave (P, R, T). The wave type detection as follows starts by calculating the maximum peak height (R) *0.6. Vertex is R. All detects all R waves. Then calculates pre-R which is a space based and time series-based Q wave for the waveform and based on direction. In the same way, the point is to detect other waves on the same basis. This concept is illustrated in Figures 3 and 4, respectively. Before the diagnostic step, the measurement results are calculated by calculating the range for each parameter (QRS, PR, P, RR, PP) and calculating the evenness. Regularity (rhythm) of the ECG or irregularity. and based on this calculation heart rate (HR) with human help. Expert (doctor). Si Regular rhythm can be determined quickly by counting the number of large graphic boxes between two R waves. This number is divided by 300 to calculate bpm. HR=300 / Number of large histogram boxes between two R waves. A regular rhythm can be quickly determined using a 6-second ECG range to calculate heart rate. Recipe:

6 seconds (calculate R* 10 bpm) the area of the QRS complex, P wave and T wave is calculated using this formula

Area = Duration * Amplitude.

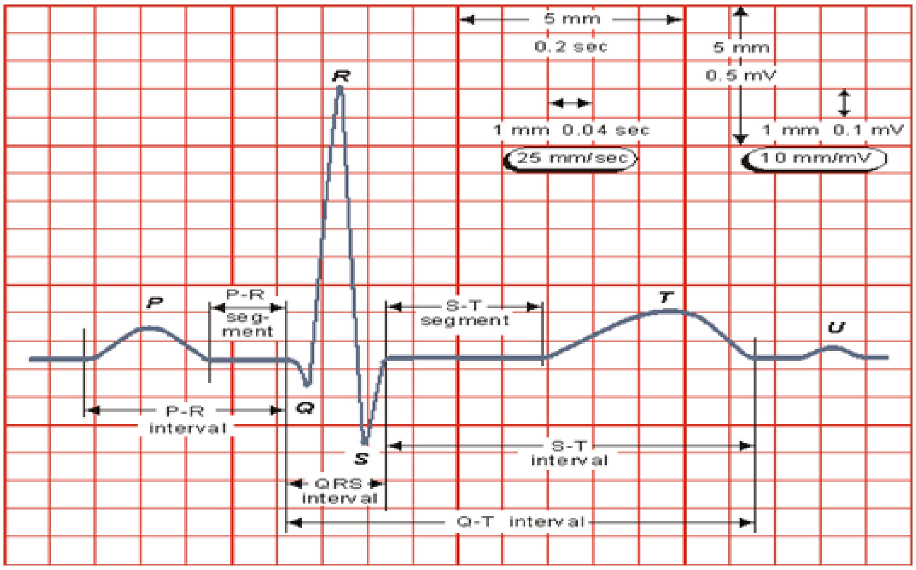


**Fig-3: ECG Range to Calculate Heart Rate**

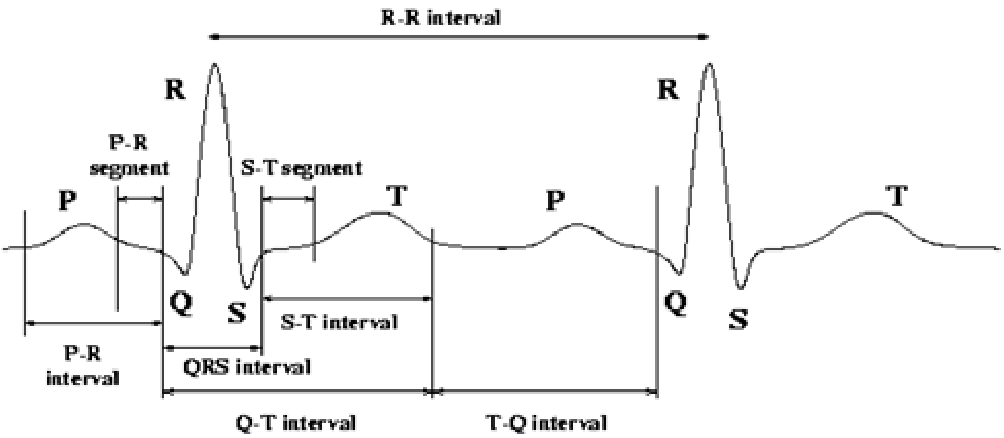


**Fig-4: Area of the QRS complex**

## 4. Result

This expert system can make eight types of decisions from the ECG signal, seven of which are disease and the other are normal based on 12 types of characteristics. This article has been decided into four disease categories based on eight characteristics. There is a dataset of 30 signals and the system can decide on 22 of them, giving the diagnostic system accuracy of 72%. When the heart rate of the signal is below 35, the cause of the disease is unknown. Wave motion is important, but properties don't show it. These characteristics are essential for making decisions about this disease. Cardiac signal analysis was performed using MATLAB software. The analysis was performed against the reference cases on the data available in the ALRTH Arrhythmia Database and it performed well. The experimental data results are given below-

**Table-1: Experimental Data**

| No | Beat rate | Arial rate | Rhythm | P wave | P wave | QRS | PR Interval | Decision |
|---|---|---|---|---|---|---|---|---|
| 1 | 73.7516 | 74.07091 | Regular | Present | Normal | Normal | Normal | Normal |
| 2 | 75.50336 | 76.12295 | Regular | Present | Abnormal | Normal | Prolonged | First Degree AV Block |
| 3 | 48.56661 | 47.55345 | Regular | Present | Normal | Normal | Prolonged | Sinus Bradycardia |
| 4 | 40.78645 | 40.83643 | Regular | Present | Normal | Normal | Short | Sinus Bradycardia |
| 5 | 37.2093 | 37.45611 | Regular | Present | Normal | Normal | Prolonged | First Degree AV Block |

## 5. Conclusion

In this article, the proposed system is mainly used to determine the human heart condition. In general, many people are affected by many diseases that can be detected by ECG signals. Thus, the diagnosis of the disease has been computerized to find the disease; this system will give us an approximate decision. Obviously, ECG signal is very important in medical science for diagnosis and many different diseases can be detected by this ECG. Therefore, there are several articles on ECG in which the detection of disease from ECG has been discussed. But in most articles, this system only provides four types of disease detection. This article introduces some characteristic data that help people decide what is normal and abnormal. And how to detect more than 4 diseases from ECG signal by computer and add extraction feature here has also been clarified.

## References

ABM Abdullah. ECG in medical practice. Pages 14-27, 3rd edition, 2010. Billauer, E. Peakdet

D. Haugland, J.G. Herber and J. H. Husoy. Optimized algorithm for compression of ECG data. In biomedical engineering - applications, bases and calculations. Pages 420-424, vol. 35, 1997.
Detecting peaks in MATLAB. September 2008. [quote:]

Electrocardiogram, http/en.wikipedia.org/wiki/Electrocardiography, last accessed August 25, 2013.

GD Clifford, F. Azuaje and P. McSharry. Advanced methods and tools for ECG data analysis. InArtech House Publishing House. Pages 102-107, Volume 3, 1995

I.K. Daskalov, I.I Christov. Electrocardiogram signal preprocessing for automatic detection of QRS limits. Medical Engineering and Physics. Pages 37-44, Volume 21, Issue 1, February 1999
Jiapupan, WJ Tompkins. Real-time QRS detection algorithm. IEEE Transactions in Biomedical Engineering. Pages 65-69, volume BME-32, no. 3, 1985.

M.K. Islam, A.N. M. M. Haque, G. Tangim, T. Ahmad và M. R, H. Khondokar. Study and analyze ECG signals using MATLAB and LABVIEW are powerful tools, International Journal of Electrical and Computer Engineering. Pages 404-408, vol. 4, number 3, 2012.

Muzhir Shaban Al-Ani and atiaf Ayal Rawi. Arule-based expert system for automatic ECG diagnosis. International Journal of Advances in Engineering and Technology, September 2013.

Muzhir Shaban Al-Ani and atiaf Ayal Rawi. ECG rhythm diagnostic method for printing based on expert system. International journal of engineering technology and advanced engineering. Volume 3, issue 4, April 2013.

October 16, 2011. [Online]. Available online:
R. Vishnubhotla. Preprocessing of ECG signals for use in ambulances. Northern Illinois University, IL 60115 USA, 2008.
www.bluer.co.il.