Electrocardiogram Feature Recognition Algorithm with Windowing and Adaptive Thresholding

S I Purnama1, H Kusuma1, and T A Sardjono1

seviaind@gmail.com, hendraks@ee.its.ac.id, sardjono@bme.its.ac.id

1 Department of Electrical Engineering Faculty of Electrical Engineering Institut Teknologi Sepuluh Nopember Surabaya, Indonesia

Abstract. Heartbeat abnormalities of human body can be diagnosed by observing electrocardiogram (ECG) signal. Traditional methods of analyzing ECG signals to determine a person's abnormality are based on the expertise of cardiologists, where sometimes multiple interpretations or misinterpretations of the disorder occur. The development of pattern recognition methods nowadays have rapidly advanced so that make it possible to be applied to ECG signal. Certain feature of ECG required for pattern recognition are P, Q, R, S and T signals. In this paper, we propose a pattern recognition method for ECG features by using adaptive threshold to find P, Q, R, S, and T position. First, we find R signal defined by local peaks, P and T signal which are defined by maximum value from a specific window, S signal defined by local valleys of the ECG signal and the rest Q signal which is defined by minimum value between P and R signal. Then based on those information, we use 48 ECG signals that contain abnormality and 18 normal ECG signals from physionet database. Experimental results show that the accuracy level of our method to recognize P, Q, R, S and T signals are 96.52%, 95.88%, 96.56%, 98.35%, and 95.88% respectively for both normal and abnormal ECG signal.

1. Introduction
An electrocardiogram (ECG) is a measurement device for activity, transportation and restoration of the heart rate using electrodes that attached to the skin [1]. Sinus rhythm of the heart characterized by a sequence of P, Q, R, S, and T signal in ECG, and their combination complexes called QRS interval [2]. The P-wave of ECG shows the atrial depolarization [3]. Obtained ECG had the most prominent segment in every ECG cycle called QRS interval that characterized by a sharp waveform with a high amplitude [4] and T-wave of ECG shows the ventricular repolarization [3]. The QRS complex shown rapid depolarization of the ventricles with duration less than 0.09s. QRS complex has maximum amplitude in this interval called R-peak and the time period between two R-peak is RR-interval that used to calculate heart rate and its variability [4]. Normal heart rate has 60-100 beats per minute and it can be different for elderly and children. When human has an abnormalities in their body, their ECG signal pattern will change. There are several difficulties in processing signal ECG such as various distance between peaks, uncertain peak form, as well as the low frequency components presence due to patient breathing [5]. Many research about getting pattern from ECG signal such as extracting unwanted noise using a discrete wavelet transformation resulting in a reduction in noise in the frequency of the pass band entered, undecimated wavelet transform to eliminate noise from ECG signals, denoising signals for ECG using extended kalman filters which resulted in RR approach.
approaching RR original, R-peak detection on the wearable ECG using Matched Filtering and Hilbert Transform, relationship of anxiety with T wave abnormalities in subjects without heart disease [6]. Furthermore, the pattern for normal and abnormal ECG signal is completely different [5] [6]. In other hand, developing a method that can be applied for general signal especially for abnormal ECG signal is a challenge. This research proposed an ECG signal pattern recognition that can be applied in both normal and abnormal ECG signal to detect P, Q, R, S and T signal. After we get those signals we are able to recognize which one is normal or abnormal heartbeat.

2. The Proposed Method

The proposed method consists of preprocessing and heartbeat detection steps. Basically in the preprocessing step, the ECG signals are filtered by LPF and HPF, then passed to a normalization procedure. The next step is to detect P, Q, R, S, and T signals. This preprocessing step is needed because it is difficult to find those P, Q, R, S, and T signal in the raw ECG signals. The results is the detected P, Q, R, S, and T signals in both normal and abnormal ECG signal respectively. By knowing the position and value of those signals, we could recognize the pattern of normal and abnormal heartbeat. We use 48 abnormal ECG signal from physionet database obtained from physionet.org. The name of database is “Normal Sinus Rhythm Database (nsrdb) and MIT-BIH Arrhythmia Database (mitdb)”.

![Diagram Block Of The System](Figure 1)

2.1. Preprocessing

In preprocessing, both normal and abnormal signal will get the same process. The process is filtering by using high pass filter to remove low frequency noise. After we get the pure signal we normalize all of the signals by moving the negative portion of ECG signal to be positive value. ECG signal must be filtered since the possibility of noise from body movement. To remove noise below 5Hz baseline, we used 5th order zero phase IIR with cut off frequencies of 0.5Hz. This value relies on hypothesis that human body heart rate range varies between 40 (0.66Hz) and 200 (3.33Hz) [3]. The original ECG is shown in Figure 2(a), this signal is a raw ECG signal from the database before preprocessing applied. The filtered signal is shown in Figure 2(b). The original signal and filtered signal in one graph can be seen in figure 2(c) respectively.
And for last process for preprocessing step is the normalization step. Basically the Normalization step is needed for lifting up the negative signal. Thus all normal and abnormal signals do not have negative values. If $\mathbf{N}_s$ is normalized signal, $\mathbf{y}$ is a signal that we want to normalize, then equation 1 shows the formula to obtain $\mathbf{N}_s$ [7]:

$$\mathbf{N}_s = \sum_{n=0}^{N}(y_n - \min(y))$$

According to the equation above, we need to find the minimum value from the signal $\mathbf{y}$ and subtract the $\min(y)$, which is the value of the successive signal with the minimum value. $\mathbf{N}_s$ is a signal that can be process to detect the position and value of P, Q, R, S, and T signal in normal and abnormal signal. The aim of this preprocessing process is simplify raw signal from MIT-BIH Arrhythmia Database (mitdb) before we develop a system that can be recognize normal and abnormal ECG signal. The normalize ECG signal can be seen in figure 3.
2.2. Adaptive thresholding and windowing for finding P, Q, R, S, and T signal

First, we need to find the maximum value from the signal and multiply it with 50% and 55%. In this step we will get 2 values which are 50% and 55% from the maximum value. Secondly, we have to find how many signals that have a value above the 50% and 55% from the maximum. If the number of the signal above 50% is more than the number of the signal above 55% and the number of the signal above 55% is more than 20 and the percentage is below 86%, then we add the percentage 5%. So the percentage multiply value become 55% and 60%. In this method we use the fixed value such as 20 and 86%. The Fixed value of 20 is needed to avoid maximum value from the process. The percentage fixed-value of 86% is used to determine the threshold of the maximum value from the signal. Based on this method, when we implement to the normal ECG signal, all of R signals were found, and for arrhythmia ECG signals most of them were found.

Figure 3. Normalized Signal

Figure 4. Flowchart to Find R
To find P and T signals, we use window between two R signal in all ECG signal respectively. The P and T signals are defined by the maximum signal between two R signals. Therefore, we define P signal is a signal with value 40 – 70% of the maximum value from the signal. First we must find midpoint between two R signals. Subtract the region between two R signals into 2 region, first region is below midpoint from the first R, the second region is above midpoint, i.e. from the midpoint to second R. The signal in the first region which is 40 – 70% of the maximum value is T signal. For the next 40 – 70% of the maximum value that found in second region which is between midpoint and second R peak is P signal for second R.

Figure 5. The Flowchart to Find P and T Signal

Figure 6. Location of R, P and T
When R peak detected, then RR interval can be counted with this equation:

\[ RR\ int = \frac{\text{number of peak second peak location} - \text{first peak location}}{2} \]  

(2)

After the R peak detected, we then calculate the P and T signals by windowing them. The flowchart for windowing P and T shown in Figure 5. With the same way as flowchart in Figure 7, The Q and S can be obtained by finding its minimum value. The Q signal is a minimum value of P and R. The S signal is a minimum value between R peak and T. The result of finding this feature shown in figure 8.

Figure 7. The Flowchart to find Q and S

Figure 8. Feature detection of ECG signal
Figure 8 above shows the accuracy result of this algorithm for every signal. The P signal indicates by red +, Q signal indicates by green *, R signal indicates by red *, S signal indicated by blue o and T signal indicated by blue +. This signal in the picture is the feature of ECG Signal from a record.

3. Data Analysis
To know the effectiveness of our proposed algorithm, we tested it on 48 signals from MIT-BIH arrhythmia database, we calculate the accuracy level by the percentage of the correct data obtained. The accuracy rate of this recognition can be seen in table 1 below,

| Patient | Accuracy (in percentage) |
|---------|--------------------------|
| label   | P  | Q  | R  | S  | T  |
| 100     | 100.00 | 98.65 | 100.00 | 100.00 | 98.65 |
| 101     | 100.00 | 98.33 | 100.00 | 100.00 | 98.33 |
| 102     | 98.63 | 97.30 | 98.65 | 91.89 | 97.30 |
| 103     | 100.00 | 98.57 | 100.00 | 100.00 | 98.57 |
| 104     | 96.15 | 94.94 | 96.20 | 93.51 | 94.94 |
| 105     | 83.95 | 82.93 | 84.15 | 100.00 | 82.93 |
| 106     | 100.00 | 98.39 | 100.00 | 100.00 | 98.39 |
| 107     | 100.00 | 98.57 | 100.00 | 90.14 | 98.57 |
| 108     | 62.07 | 61.02 | 62.71 | 100.00 | 61.02 |
| 109     | 74.12 | 73.26 | 74.42 | 97.30 | 73.26 |
| 110     | 72.46 | 71.43 | 72.86 | 98.57 | 71.43 |
| 111     | 94.05 | 92.94 | 94.12 | 100.00 | 92.94 |
| 112     | 98.36 | 97.77 | 98.39 | 100.00 | 97.77 |
| 113     | 100.00 | 98.18 | 100.00 | 100.00 | 98.18 |
| 114     | 100.00 | 98.51 | 100.00 | 100.00 | 98.51 |
| 115     | 100.00 | 98.11 | 100.00 | 100.00 | 98.11 |
| 116     | 96.10 | 94.87 | 96.15 | 100.00 | 94.87 |
| 117     | 96.36 | 94.64 | 96.43 | 100.00 | 94.64 |
| 118     | 90.14 | 88.89 | 90.28 | 98.61 | 88.89 |
| 119     | 92.31 | 90.91 | 92.42 | 87.79 | 90.91 |
| 120     | 94.83 | 93.22 | 94.92 | 100.00 | 93.22 |
| 121     | 100.00 | 98.49 | 100.00 | 100.00 | 98.49 |
| 122     | 100.00 | 98.15 | 100.00 | 100.00 | 98.15 |
| 123     | 100.00 | 98.11 | 100.00 | 100.00 | 98.11 |
| 124     | 97.47 | 96.25 | 97.50 | 100.00 | 96.25 |
| 201     | 98.86 | 97.75 | 98.88 | 100.00 | 97.75 |
| 202     | 91.38 | 89.83 | 91.53 | 98.33 | 89.83 |
| 203     | 78.89 | 78.02 | 79.12 | 90.00 | 78.02 |
| 205     | 98.88 | 97.78 | 98.89 | 100.00 | 97.78 |
| 207     | 80.60 | 79.41 | 80.88 | 78.79 | 79.41 |
| 208     | 60.67 | 60.00 | 61.11 | 100.00 | 60.00 |
| 209     | 100.00 | 98.95 | 100.00 | 100.00 | 98.95 |
| 210     | 92.94 | 91.86 | 93.02 | 92.94 | 91.86 |
| 212     | 100.00 | 98.85 | 100.00 | 100.00 | 98.85 |
| 213     | 97.20 | 96.30 | 97.22 | 100.00 | 96.30 |
| 214     | 69.33 | 68.42 | 69.74 | 94.74 | 68.42 |
| 215     | 96.40 | 95.54 | 96.43 | 100.00 | 95.54 |
| 217     | 100.00 | 98.63 | 100.00 | 100.00 | 98.63 |
| 219     | 97.30 | 96.00 | 97.33 | 97.33 | 96.00 |
| 221     | 83.12 | 82.05 | 83.33 | 100.00 | 82.05 |
| 222     | 95.89 | 94.59 | 95.95 | 100.00 | 94.59 |
| 223     | 95.35 | 94.25 | 95.40 | 96.67 | 94.25 |
| 228     | 100.00 | 98.53 | 100.00 | 94.12 | 98.53 |
| 230     | 96.25 | 95.06 | 96.30 | 100.00 | 95.06 |
| 231     | 98.33 | 96.72 | 98.36 | 96.72 | |
| 232     | 98.28 | 96.61 | 98.31 | 98.31 | 96.61 |
| 233     | 89.11 | 88.24 | 89.22 | 78.43 | 88.24 |
| 234     | 100.00 | 98.89 | 100.00 | 100.00 | 98.89 |

The proposed method has detection accuracy of 93.13%, 93.04%, 95.88%, 95.52%, and 98.35% for R, P, T, Q, and S peaks respectively. If we apply this method for normal ECG database then we get 100% detection accuracy. Therefore by averaging all the detection accuracy results we have the detection accuracy of 96.52%, 95.88 %, 96.56 %, 98.35% and 95.88% for R, P, T, Q, and S respectively, when applied to all database used in this paper.

4. Conclusion
The proposed method in this paper can recognize P, Q, R, S, T signals for every ECG signals. The Features recognition method is tested on 48 database MIT-BIH Arrhythmia signals, obtained from the
physionet.org. The accuracy rate to find R peak, P, T, Q, and S are 96.52%, 95.88 %, 96.56 %, 98.35% and 95.88% respectively. Overall, the average accuracy is 93.27%.

5. References

[1] S. T. Tao Lin, "An ECG Signal Processing System Based on MATLAB and MIT-BIH," *Recent Advances in CSIE 2011, LNEE 125*, p. 787–792., 2012.

[2] D. Castells-Rufas and J. Carrabina, "Simple real-time QRS detector with the MaMeMi filter," *Biomedical Signal Processing and Control*, vol. 21, pp. 137-145, 2015.

[3] O. M. Wani, "Signal Processing of Stress Test ECG using MATLAB," *International Journal of Engineering Research & Technology (IJERT)*, vol. 6, no. 8, pp. 175-183, August 2017.

[4] T. Chanwimalueang, W. v. Rosenberg and D. P. Mandic, "Enabling R-Peak Detection in Wearable ECG: combining matched filtering and hilbert transform.," *2015 IEEE International Conference on Digital Signal Processing (DSP)*, pp. 134-138, 2015.

[5] M. Ponnusamy and S. M, "Detecting and classifying ECG abnormalities using a multi model methods," *Biomedical Research India 2017 special issue*, no. Special Issue: S81-S89, pp. s81-s89, 2017.

[6] B. K. Rehman, A. Kumar and P. Sharma, "A Novel Approach for R-Peak Detection in The Electrocardiogram (ECG) Signal," *ARPN Journal of Engineering and Applied Sciences*, vol. 11, pp. 13500-13503, December 2016.

[7] S. Hamdi, A. B. Abdallah and M. H. Bedoui, "A Robust QRS Complex Detection Using Regular Grammar and Deterministic Automata," *Biomedical Signal Processing and Control*, vol. 40, pp. 263-274, 2018.