How to cite this paper:

Mohamad Sabri Bin Sinal & Eiji Kamioka. (2017). Adaptive threshold based approach to perfectly detect heart cycle in ECG data in Zulikha, J. & N. H. Zakaria (Eds.), Proceedings of the 6th International Conference on Computing & Informatics (pp 492-498). Sintok: School of Computing.

ADAPTIVE THRESHOLD BASED APPROACH TO PERFECTLY DETECT HEART CYCLE IN ECG DATA

Mohamad Sabri Bin Sinal[1] and Eiji Kamioka[2]

[2]*Shibaura Institute of Technology, Japan,* {nb16109, kamioka}@shibaura-it.ac.jp

ABSTRACT. Heart cycle detection is considered as one of the most significant computational processes in automatic ECG data analysis. It is a crucial step to find any sign of cardiovascular abnormality especially in a long time series of ECG data for diagnosing early stage of cardiac diseases. However, both a high accurate heart cycle detection and a short computational time have not as yet been achieved for analysis. This is because the long time series of ECG data contains a lot of signal noises which make the heart cycle features undetectable with the conventional fixed threshold based approach and previous researches perform complex calculations. This paper proposes a new simple algorithm, which automatically, accurately and quickly detects all the heart cycles in a long time series of ECG data. It correctly identifies important ECG features based on an adaptive threshold approach. The evaluation results show that the proposed algorithm outperforms other existing methods in terms of detection accuracy and computational time.

Keywords: adaptive threshold algorithm, heart cycle detection, QRS morphology

1. INTRODUCTION

Cardiovascular disease is the primary cause of death in the world with more than 17 million victims within a year [1]. Basically, the cardiovascular abnormality can be diagnosed by using electrocardiogram known as ECG, which is a time series data showing myocardial contractility during the cardiac cycle. It is very important to analyze a long time series of ECG data recorded for more than several hours in order to diagnose the early stage of cardiovascular abnormality. However, it is very challenging to automatically, accurately and quickly detect all the heart cycles in the long time series data due to a huge number of noises. Although many studies have tried to overcome this issue focusing on accurate heart cycle detection, most of the approaches take a lot of time for the preprocessing, feature extraction and decision [2]. In addition, the accuracy of heart cycle detection is inappropriate for the long time series of data.

In this paper, a new algorithm, which automatically, accurately and quickly detects all the heart cycles in a long time series of ECG data, is proposed. It correctly identifies important ECG features based on an adaptive threshold approach. In addition, it does not use any complex mathematical calculations. As a result, the proposed algorithm outperforms other existing methods in terms of detection accuracy and computational time. However, the abnormalities of heart cycle detection will not be covered in this research as the focus will be on long duration heart cycle detection.

The rest of the paper is organized as follows: Section 2 will discuss five related. Section 3 introduces general features of ECG data to clarify the points focused by the proposed algorithm. Section 4 describes the proposed algorithm in detail. In section 5, the performance evaluation is discussed based on the experimental result. Finally, the paper is concluded with discussions on possible future works in Section 6.

2. RELATED WORKS

K.Muthuvel et al. proposed an algorithm to detect abnormal heart cycle signals in ECG data using Haar Wavelet Transform (HWT). In this algorithm, the preprocessing and classification were executed by using Feed Forward Neural Network. It achieved the overall detection accuracy of 73% with the symptom classification accuracy of 63%. There are two significant drawbacks in the proposed algorithm, namely, the large computational time and the low detection accuracy. Since the algorithm uses Haar Wavelet Transform and Feed Forward Neural Network, it takes a lot of time to obtain the computational results. In addition, the influence of using Haar Wavelet Transform on the sensitivity of signal detection was not clearly discussed. This might be the reason why the overall detection accuracy is not satisfactory, affecting the accuracy of classifying symptoms.

Indu saini et al. proposed a method to detect QRS-complex in ECG data using K-Nearest Neighbour (KNN) as a classifier. QRS complex is a combination of three of the graphical deflections seen in ECG data and it is a significant feature for detecting heart cycle. The detailed explanation of Q, R and S peaks will be described in section 3. In the proposed algorithm, the gradient of ECG signal was calculated and used as a feature for the KNN classifier. This method was evaluated using two types of database which are CSE database and MIT-BIH Arrhythmia database. The overall accuracy of the proposed method is 90% in detecting QRS complex. Even though the accuracy is high, the feature extraction process required large computational time.

Donati M et al. , proposed a practical ECG device with user-friendly and ergonomic feature which allows self-measurement of the high quality first lead ECG data for telemedicine. It is measured by holding or grip to the devices which allows an electric contact between the hands and the electrodes. Although this study proposed a new mechanism to record a quality ECG data continuously but analyzing any symptom from the data is out of the scope. Likewise Parak J et al. , proposed three heart rate frequency detection algorithms with custom digital filters to detect heart rate accurately. It is based on statistical and differential mathematical methods. The proposed mechanisms are well performed with computational time, this research did not focus on P peak, Q peak, S peak, and T peak. Therefore the reusability of the result for diagnosis will be limited. Christov I. I, proposed a real-time electrocardiogram QRS detection method based on three parameters as a threshold to detect QRS cycle. The accuracy of first algorithm is 99.69% while second algorithm is 99.74%.Although the accuracy of detection is high for both, the duration time for detection is 30 minutes only. Therefore a new mechanism is needed to determine diagnosis in a long duration of time.

3. TIME SERIES FEATURE IN ECG DATA

In a common diagnosis procedure of heart condition, a visual presentation of heart activities on pieces of paper is used. It refers to only a small part of detailed ECG pattern in a short time period or a general point of view in the long time period due to the impracticality of printing the detailed long time series data. However, some types of cardiovascular abnormality do not appear in such a short time period or in a general tendency without detailed features of ECG data. Hence, a new method, which automatically and accurately detects all the heart cycles in the long time series of ECG data, is required. Furthermore, the computational time must be as short as possible to quickly provide the result of diagnosis.

In this study, only normal heart cycle called Normal Sinus rhythm (see Section 3.1) is focused on for heart cycle detection. This is because the normal heart cycle has a specific feature with five peaks in the ECG waveform (see Section 3.2) and if such a feature cannot be detected appropriately, it shows the possibility of cardiovascular abnormality. However, if the proposed algorithm has a lot of miss-detections, it will cause many patients to be candidates of cardiovascular abnormality. Therefore, the normal heart cycle in the long time period must be detected accurately.

3.1 Normal Sinus rhythm

Normal Sinus rhythm represents a normally functioning conduction system in the body where there is no interference from any disease processes. Therefore, the electrical current from an electrocardiograph flows through the normal conduction pathway without any disturbance. If a patient has a cardiovascular abnormality, the tendency can be seen in a heart cycle or the number of normal heart cycles decreases.

3.2 Morphology of P, Q, R, S and T peaks

ECG signals of Normal Sinus rhythm has a specific feature of cardiac activities, which is composed of the P, Q, R, S, and T peaks, representing each activity inside the heart; the P peak represents the depolarization of the right and left atria, the QRS complex follows the P peak and depicts the activation of the right and left ventricles, and the T peak indicates repolarization of the ventricles. Therefore, it is indispensable to detect these 5 peaks in ECG data to diagnose the cardiovascular abnormality. Even though the ECG signals show a Normal Sinus rhythm, there exists individual variability in the waveform. Figure 1 depicts two different waveforms of Normal Sinus rhythm in ECG data. It is not appropriate to use absolute threshold values or to find the highest/lowest peak for detecting the above mentioned 5 peaks. For instance, from the left figure of Fig. 1, you may assume that the R and Q peaks can be detected by finding the peak with the value larger than 2.5mV and by finding the peak with the smallest value, respectively. However, both the R and Q peaks cannot be detected in the right figure of Fig. 1 with such conditions. Therefore, an adaptive threshold based approach considering the vicinal information of a peak is needed for a better heart cycle detection.

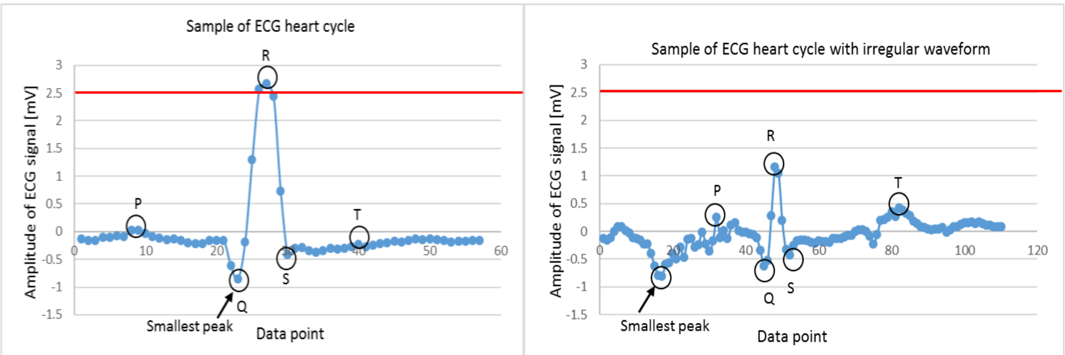

Figure 1. Two Different ECG Heart Cycle Behaviors with Normal Sinus Rhythm.

4. ADAPTIVE THRESHOLD BASED ALGORITHM

In this section, a new adaptive threshold based algorithm to detect heart cycle is proposed. It adapts to any kind of signal pattern and detects the P, Q, R, S and T peaks accurately. The relative signal height from one peak to the other adjacent peaks can be a significant trigger to detect each peak. By utilizing this idea, an accurate heart cycle detection can be achieved even in a long time series of ECG data. In general, the R peak has the largest value among a heart cycle, and thus it is not very difficult to detect the R peak. Therefore, this proposed algorithm will detect the R peak first. The algorithm searches all the local maximum values as peaks and investigates how steep their peaks are by comparing the value with the values of vicinal data points. If the peak is steep enough, it will be identified as the R peak. Then, the Q and S peaks will be detected by searching the minimum values backward and onward from the R peak in a specified time period. Finally, the P peak and T peaks will be detected by searching the maximum values backward from the Q peak and onward from the S peak in a specified time period in the same way as detecting the Q and S peaks.

This algorithm is composed of four different stages as shown in Fig. 2 in order to achieve the accurate peaks detection. The detailed procedure of detecting the P, Q, R, S and T peaks is as follows:

Stage 1: Search the local maximum value among successive 7 data points from the beginning of ECG data. Let V_i denote the detected local maximum value at i-th data point. Obtain the value difference between two data points such as V_i-V_{i-1}, V_i-V_{i-2}, V_i-V_{i-3}, V_i-V_{i+1}, V_{i}-V_{i+2}, V_i-V_{i+3}, V_{i-1}-V_{i-2}, V_{i-1}-V_{i-3}, V_{i-2}-V_{i-3}, V_{i+1}-V_{i+2}, V_{i+1}-V_{i+3}, V_{i+2}-V_{i+3}. Compare each value difference with each certain value determined in advance to decide if the detected local maximum value is the R peak or not. When the R peak is identified based on the above condition, go to Stage 3.

Stage 2: The considered successive 7 data points are shifted one data point onward and repeat the same process of Stage 1 until identifying the R peak.

Stage 3: When the R peak is identified, search the minimum value in the range of 8 data points backward from the R peak to detect the Q peak, and search the minimum value in the range of 8 data points onward from the R peak to detect the S peak.

Stage 4: Search the maximum value in the range of 40 data points backward from the Q peak to detect the P peak, and search the maximum value in the range of 40 data points onward from the S peak to detect the T peak.

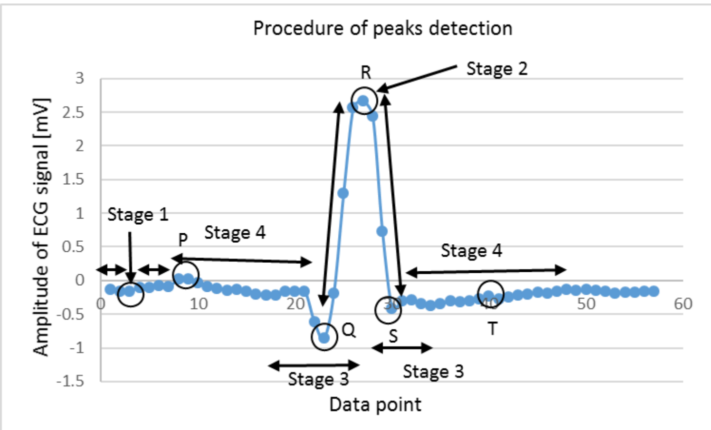

Figure 2. Proposed Adaptive Threshold Based Algorithm.

5. PERFORMANCE EVALUATION

In this section, the performance of the proposed algorithm to detect heart cycle is discussed. Sensitivity of the P, Q, R, S and T peaks detection is introduced as a metric of the performance evaluation. In general, the sensitivity is utilized to evaluate a reliability of statistical binary classification. In this experiment, the sensitivity represents the probability how correctly the proposed algorithm detects the actual peaks as peaks. The detail is described below:

Sensitivity = True Positive / (True Positive + False Negative)

Where,

True Positive = the number of actual peaks correctly detected as peaks

False Negative = the number of actual peaks not detected as peaks

The proposed algorithm was coded in Java programming language to confirm if it accurately detects all the five peaks in ECG data. MIT-BIH Normal Sinus rhythm database from PhysioNet was selected as the ECG testing data for this evaluation. PhysioNet is a free open source database that provides a lot of data related to biomedical research and development. All the ECG data used in this study comes from real patients with Normal Sinus condition. Firstly, 17 ECG patients' data with the sampling frequency of 129 Hz were analyzed to confirm if the data is appropriate enough for evaluation. Each data in one minute extracted from the 12 hours ECG data was checked in terms of the waveform feature for P, Q, R, S and T peaks. As a result, it was concluded that the quality of the ECG data is acceptable for the evaluation; including difficult conditions to detect each peak, such as sharp signal peak, tall P and T peaks, negatives QRS complex, small QRS complex and wider QRS in the data.

As the quality of the data has been validated, 17 long time series of ECG data, namely 12 hours ECG data of 17 patients stated above, were utilized for the evaluation. Note that it is almost impossible to manually verify if all the peaks are correctly detected or not since there are more than 40,000 heart cycles even in one patient data. In order to validate the proposed algorithm to the long time series of ECG data, 100 heart cycle samples in each patient data were randomly selected and the detection sensitivity was evaluated.

The detection sensitivity of heart cycle in 12 hours ECG data obtained from 17 patients was 100%. The total number of heart cycles detected in the data is different in each patient as shown in Fig. 3. The figure also showed that the minimum number of heart cycles produced in 12 hours is above 40,000. This number can be the first criterion to detect candidates of cardiovascular abnormality without any complicated manual investigations.

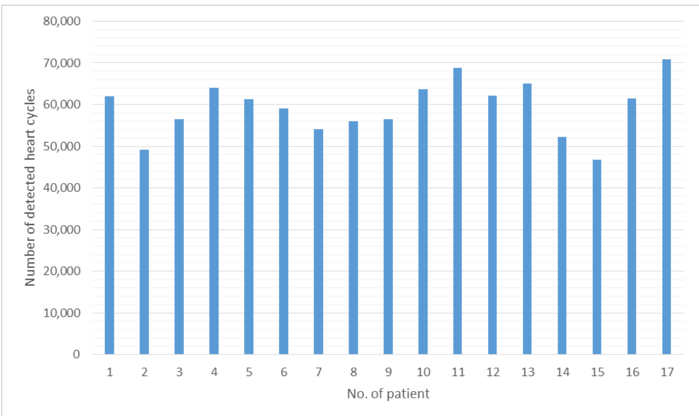

Figure 3. Total Number of Heart Cycles Detected in 12 Hours ECG Data.

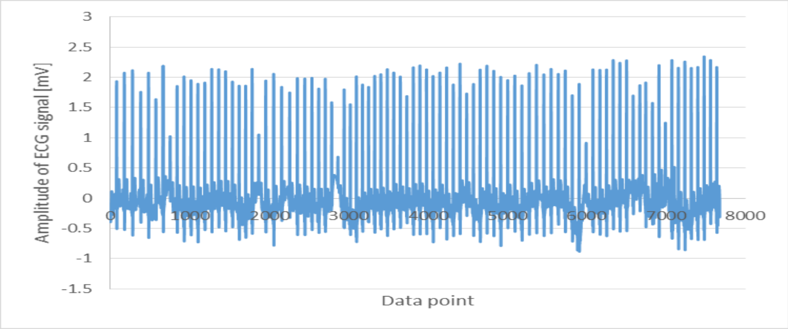

Figure 4. Unstable Heart Cycles in 1 Minute ECG Data.

From the result of detection sensitivity mentioned above, the effectiveness of the proposed algorithm has been verified. It means that the proposed adaptive threshold based approach works perfectly for the long time series of ECG data. The proposed mechanism has been overcome the existing approaches by Chandrakar C et al. [10] which achieved the accuracy of 99.57% for detecting heart cycle in 72 minutes only. This is the achievement that fixed threshold based approaches have not succeeded yet. ECG data includes unexpected and irregular waveforms that occur during the data recording such as muscle noise, baseline drift or even sudden change in QRS morphology. Therefore, applying threshold values to identifying the heart cycle induces detection errors or miss detections, particularly when the heart cycle has any unexpected behavior as shown in Fig. 4. This problem strongly influences the detection sensitivity for long time series of ECG data. As an alternative solution towards this issue, bandpass filters will be helpful to eliminate the baseline drift. However, the other problems cannot be solved. On the other hand, the mechanism to detect each peak with the proposed adaptive thresholds is based on a well-defined data distribution model in Normal Sinus rhythm, resulting in the tolerance for noise and irregular waveform change.

In order to complete this study, the computational time of the proposed adaptive threshold based approach was analyzed. The computational time of the proposed algorithm implemented in Netbeans 8 (on Intel® Core™ i7 2.50 GHz, 16 GB RAM, 64-bit OS) was less than 30 second for a 12 hours ECG data. This is because the proposed algorithm does not use any complex calculations such as Haar Wavelet Transform and Feed Forward Neural Network. This result outperforms one recently published algorithm [11] of a 12 hours EGC data duration. In that published algorithm, the computation time requires less than 40 milliseconds to capture the heart cycle segment of 30 seconds. The proposed method on the other hand, needs only 30 seconds for a 12 hours ECG data. Given this short computational time, the proposed algorithm is feasible and applicable in diagnosing long time series of ECG data.

The proposed algorithm has led to a significant finding in detecting heart cycle under different kinds of noises and artifacts without relying too much on the ECG filtering techniques. The effectiveness of the proposed method has been proven as the initial step for an automate ECG analysis.

6. CONCLUSION

In this paper, a new algorithm, which automatically, accurately and quickly detects all the heart cycles in a long time series of ECG data, were proposed. To show the effectiveness of the proposed algorithm in terms of peaks detection accuracy and computational time, the performance evaluation was discussed through the experimental results. The proposed method achieved 100% of peaks detection accuracy verified by using 100 heart cycle samples randomly selected from 12 hours ECG data for 17 Normal Sinus rhythm data. The computational time of the algorithm was less than 30 seconds for 12 hours ECG data which is definitely acceptable for practical use. As future work, this algorithm will be applied to the ECG data obtained from patients who have cardiovascular abnormality to confirm the number of heart cycles detected is much smaller than the Normal Sinus rhythm. Furthermore, a framework which diagnoses the type of cardiovascular abnormality will be discussed based on machine learning techniques.

REFERENCES

Ala Alwan, (2011). *Global status report on noncommunicable diseases 2010*. Geneva, Switzerland, World Health Organization, 2011.

Asgari, S., Mehrnia, A., & Moussavi, M. (2015). Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine. *Computers in Biology and Medicine*, 60, 132-142. doi:10.1016/j.compbiomed.2015.03.005

Christov, I. I. (2004). Real time electrocardiogram QRS detection using combined adaptive threshold. Real time electrocardiogram QRS detection using combined adaptive threshold. 3(1) (28). Chronic patients. 2016 IEEE Symposium on Computers and Communication (ISCC).

Chandrakar, C., & Sharma, M. (2015). A real time approach for QRS detection of ECG signal using network intrusion detection technique. *2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*. doi:10.1109/icecct.2015.7226131

Donati, M., Benini, A., Celli, A., Iacopetti, F., & Fanucci, L. (2016). A novel device for self-acquisition of ECG signal in telemedicine systems for

Kohler, B., Hennig, C., & Orglmeister, R. (2002). *The principles of software QRS detection*. IEEE Engineering in Medicine and Biology Magazine, 21(1), 42-57.

Muthuvel, K., Veni, S. H., Suresh, L. P., & Kannan, K. B. (2014). ECG signal feature extraction and classification using Haar Wavelet Transform and Neural Network. *2014 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2014*].

PhysioNet. (2016, November 5). Retrieved from http://physionet.org/

Parák, J., & Havlik, J. (2011). ECG Signal processing and heart rate frequency detection methods. ECG Signal processing and heart rate frequency detection methods.

Senapati, M. K., Senapati, M., & Maka, S. (2014). Cardiac Arrhythmia Classification of ECG Signal Using Morphology and Heart Beat Rate. *2014 Fourth International Conference on Advances in Computing and Communications.*

Uchida, S., Mori, A., Kurazume, R., Taniguchi, R., & Hasegawa, T. (n.d.). *Logical DP Matching for Detecting Similar Subsequence*. Computer Vision – ACCV 2007. Lecture Notes in Computer Science, 628-637.