

Quick detection of QRS complexes and R-waves using a wavelet transform and K-means clustering

Yong Xia^{*}, Junze Han and Kuanquan Wang

Department of Computer Science and Technology, Harbin Institute of Technology, Harbin, China

Abstract. Based on the idea of telemedicine, 24-hour uninterrupted monitoring on electrocardiograms (ECG) has started to be implemented. To create an intelligent ECG monitoring system, an efficient and quick detection algorithm for the characteristic waveforms is needed. This paper aims to give a quick and effective method for detecting QRS-complexes and R-waves in ECGs. The real ECG signal from the MIT-BIH Arrhythmia Database is used for the performance evaluation. The method proposed combined a wavelet transform and the K-means clustering algorithm. A wavelet transform is adopted in the data analysis and preprocessing. Then, based on the slope information of the filtered data, a segmented K-means clustering method is adopted to detect the QRS region. Detection of the R-peak is based on comparing the local amplitudes in each QRS region, which is different from other approaches, and the time cost of R-wave detection is reduced. Of the tested 8 records (total 18201 beats) from the MIT-BIH Arrhythmia Database, an average R-peak detection sensitivity of 99.72 and a positive predictive value of 99.80% are gained; the average time consumed detecting a 30-min original signal is 5.78s, which is competitive with other methods.

Keywords: ECG, detection of QRS-complex and R-wave, wavelet transform, k-means clustering

1. Introduction

A typical electrocardiogram consists of five characteristic waves: the P wave, Q wave, R wave, S wave, and T wave (sometimes together with the U wave). This series of waveforms correspond to each phase of cardiac activities. The QRS complex and R-peaks are of pathological importance and their detection serves as an entry point to almost all automated ECG analysis algorithms. Based on the identified QRS complex and R-wave, the rest of the waves can be detected. In the past, many approaches to QRS and R-wave detection have been proposed, including algorithms from the field of signal processing, morphology, and machine learning. Afonso proposed detecting ECG beats by filter banks [1]. Kadambe presented a method to detect QRS with a wavelet transform [2]. These algorithms based on the simple signal processing method can obtain the detection results very quickly but with undesirable sensitivity (<99%). Strintzis used a multilayer neural network to extract ECG features [3]. Fanga proposed a novel algorithm based on the phase space trajectory of ECG [4]. Better results can

^{*} Address for correspondence: Yong Xia, Department of Computer Science and Technology, Harbin Institute of Technology, West Da Zhi St., Harbin, P.C. 150001, China. Tel.: +86 631 5687921; Fax: +86 631 5687927; E-mail: xiayong@hit.edu.cn.

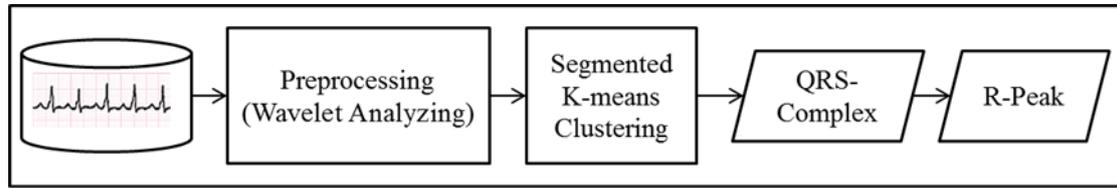


Fig. 1. Basic flow of detection.

be obtained based on machine learning and morphology theory [5-8], but it is very difficult to meet real-time application using these complicated algorithms. Therefore, new, more efficient algorithms need to be presented [9, 10].

This paper proposes a simple and quick detection algorithm for QRS complexes and R-waves. Unlike other approaches, which detect R-peaks first and then determine the QRS-complexes by the R-peaks, this method extracts the QRS region by K-means clustering first and then detects R-peaks by comparing the peaks just in the extracted QRS regions. A wavelet transform is often used to preprocess ECG signals [11-13], and we also used this method in this paper. Furthermore, different algorithms are compared based on the MIT-BIH Arrhythmia Database.

2. Methods

This paper proposes a novel scheme for detecting QRS-complexes and R-peaks that will be suitable for long time ECG monitoring. The basic flow is shown in Figure 1. First, the ECG is preprocessed by wavelet analyzing. Then, K-means clustering detects the QRS-complex. Finally, the R-peak is marked.

2.1. Preprocessing raw data using wavelet analyzing

The original signal contains some noise, such as industrial frequency noise and baseline drift. The classic analysis method is the Fourier transform. The Fourier transform can reveal the characteristics of the signal in the frequency domain. To some degree, a wavelet transform (WT) is a development of the Fourier transform. Any application using the Fourier transform can be formulated using wavelets to provide more accurately localized temporal and frequency information [1]. A smooth and quickly vanishing oscillating function was chosen as the mother wavelet. After scaling and translation, this mother wavelet function becomes the basis function in wavelet analysis. Because of the good localization of the mother wavelet, the wavelet is capable of detecting and characterizing specific phenomena in both time and frequency planes [1].

While the basis wavelet function is defined as Eq. (1) and $S(t)$ for the original signal, the wavelet decomposition function is defined as Eqs. (2) and (3). In the equation, cA_k stands for the approximation of the signal and $cD_1 \dots k$ represent detail information.

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad (1)$$

$$S(t) = \sum_b cA_0(k) \Psi_{a,b}(t) \quad (2)$$

$$S(t) = \sum_b cA_1(k) \Psi_{a-1,b}(t) + \sum_b cD_1(k) \omega_{a-1,b}(t) \quad (3)$$

This paper adopts the Sym8 wavelet to decompose the original signal at level 8. For example, using #103 from the MIT-BIH Arrhythmia Database, Figure 2 shows the result of wavelet analysis on the

electrocardiogram from 00:17:00 to 00:18:00. In the results, cA8 performs the overall approximation of the original, and cD1-cD3 contains the major noise information of the original signal. Thus the filter information contained in cA8, cD1, cD2, and cD3 eliminate the baseline drift and noise interference. The filtered result is shown in Figure 3.

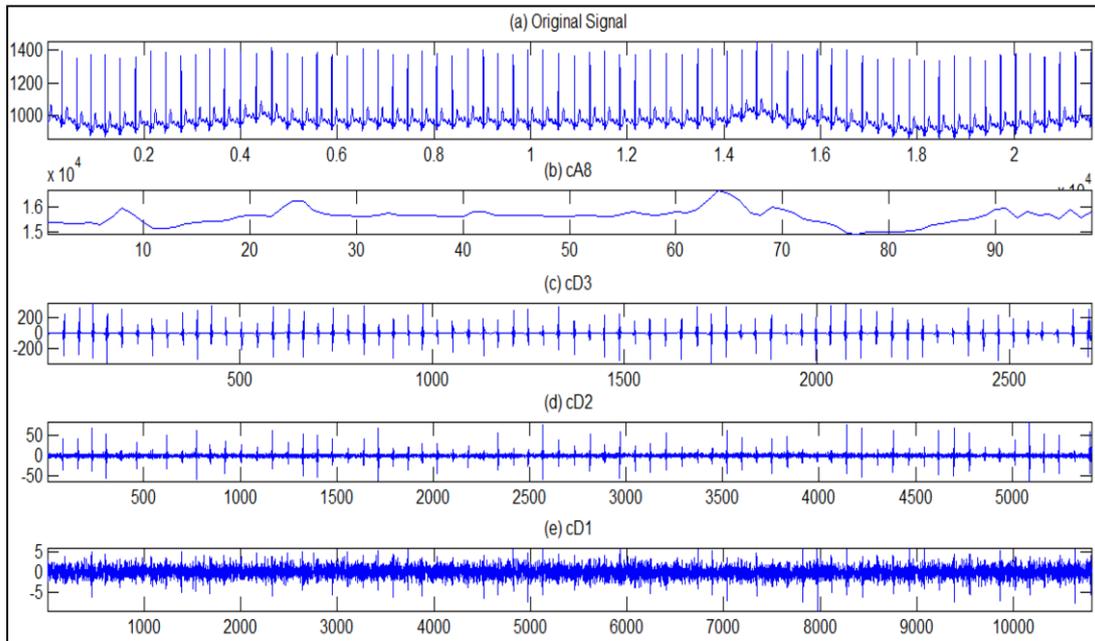


Fig. 2. Result of wavelet analysis. (a) the original signal; (b) approximation of signal at 8th level; (c, d, e) detail information at cD3-cD1 level.

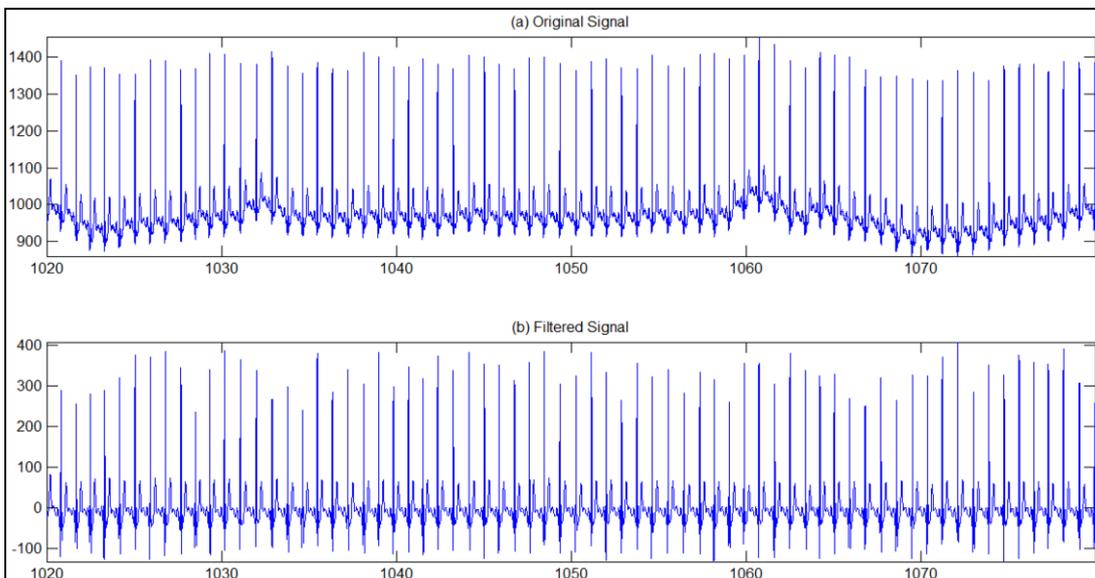


Fig. 3. Filtering result from wavelet transforms. (a) original signal; (b) filtered signal.

2.2. Detection of QRS-complexes using K-means clustering on segmented data

Points in a QRS-complex are characterized by large amplitude fluctuations. This characteristic can be reflected clearly by the slope value. So, in this section, this paper uses the slope information to extract the QRS-complex, and the procedures are as follows.

- a) Calculate absolute slope: from the input electrocardiogram data (n), calculate the slope between the adjacent two points.

$$\text{slope}(k) = \text{data}(k + 1) - \text{data}(k) \quad (4)$$

$$\text{abs}_{\text{slop}} = |\text{slope}| \quad (5)$$

- b) Cluster the absolute slope value of the divided data segment using the K-means method

The slope value is a direct reflection of the amplitude changes. The slope chart in Figure 4(a) reveals the distinguishing feature in the QRS region. During the QRS regions, very high positive slopes and very low negative slopes gather in the same area. Adopting the absolute value of the slope enhances the slope feature of the QRS region. Based on the absolute slope, we cluster the values into two classes using the K-means method.

When a number of samples are given and they must be grouped into K number of clusters, the K-means algorithm can be used. It is based on the minimization of the performance index, which is defined as the sum of the squared distances from all points in a cluster domain to the cluster center [5]. QRS regions in a relatively short time interval have more characteristics in common. The adopted K-means algorithm simply elects two cluster centers, one with a higher slope center and the other with a lower slope center. This paper adopts segmented K-means clustering; limiting the input clustering data to 1-minute long ensures the accuracy of K-means clustering.

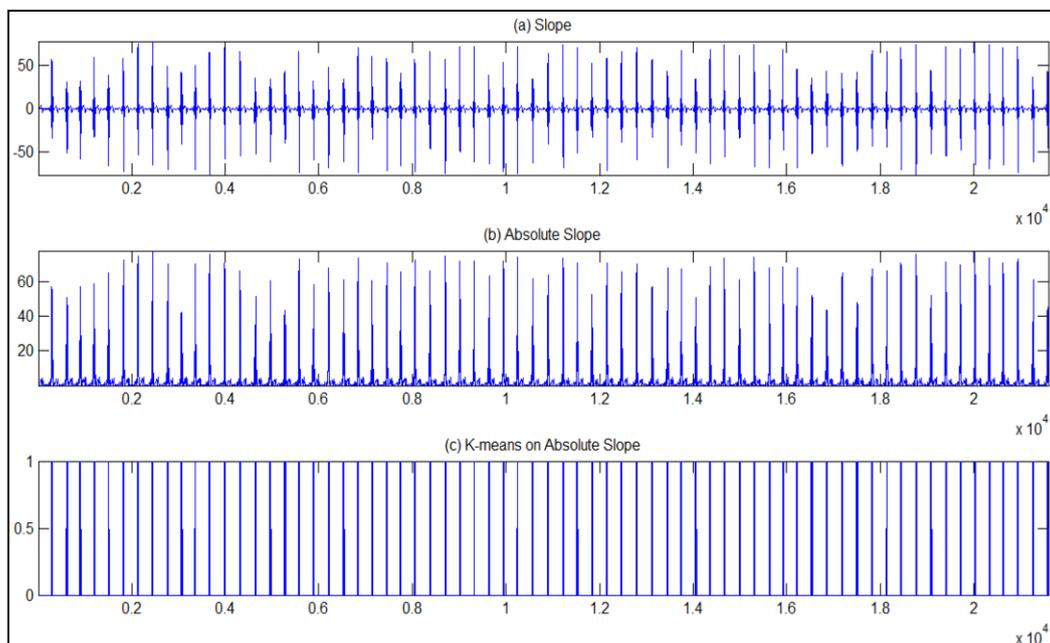


Fig. 4. Detection of QRS-complex based on absolute slope. (a) slope value; (b) absolute slope value; (c) result of K-means clustering on absolute slope.

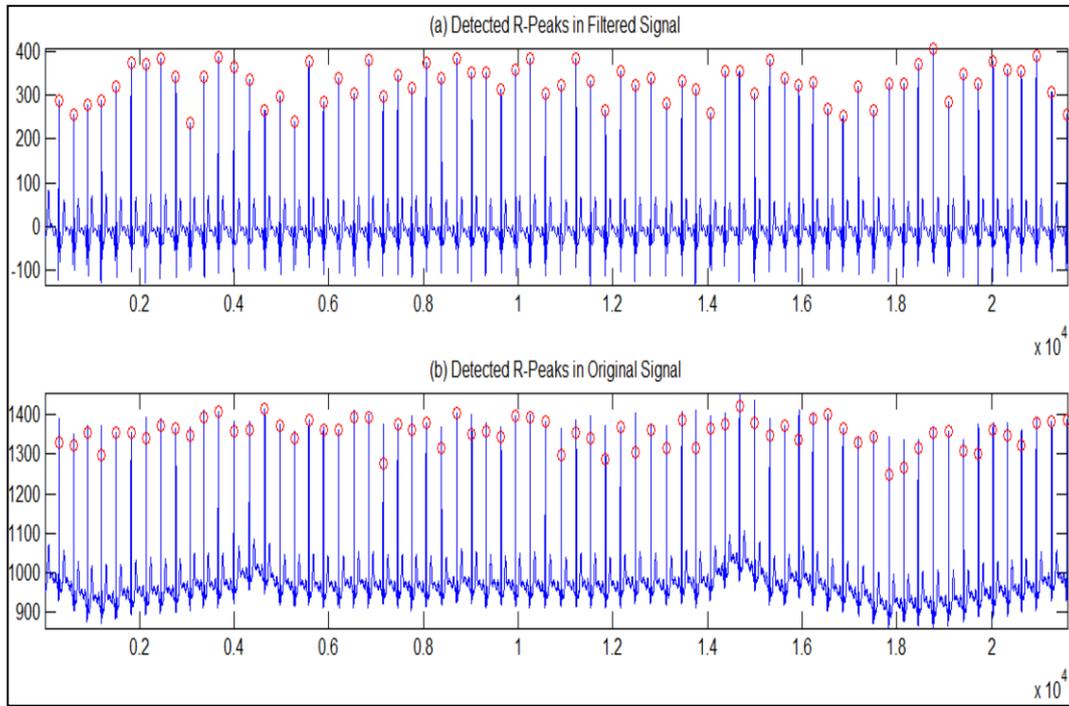


Fig. 5. Result of R-peaks detection. (a) detected R-peaks in filtered signal; (b) the same location in original signal.

2.3. Detection of R-peaks

When comparing the amplitudes during the extracted QRS region, we find the peak with the highest amplitude, which is only the R-peak in this QRS region. The result is shown in Figure 5.

- a) Find all the peaks in QRS region:
If

$$\text{slope}(k) > 0 \ \& \ \text{slope}(k + 1) < 0 \quad (6)$$

Then the point data (k) is classified as a peak.

- b) Compare the amplitudes of the peaks: in each QRS region, the highest peak is the R-peak of this QRS-complex.

3. Results

The MIT-BIH Arrhythmia Database contains 48 records. Each record consists of signals from two channels (MLII, V1, V2, V4, or V5). The above method was evaluated with Lead-MLII signals from 8 records in the MIT-BIH Arrhythmia Database. We selected 30-min signals from each record. To evaluate the detection result, we use two essential parameters to describe the performance of the R-peaks detection: detection sensitivity S_e

$$S_e = \frac{TP}{TP+FN} \quad (7)$$

and positive predictive value P^+

$$P^+ = \frac{TP}{TP+FP}. \quad (8)$$

where TP, FN, and FP are true positive, false negative, and false positive, respectively.

Referring to the annotation provided by the MIT-BIH Arrhythmia Database, the statistical results are shown in Table 1. Average R-peak detection sensitivity reached 99.72%, and average positive predictive value reached 99.80%. In addition, the results of several other algorithms are listed in Table 2. According to the comparison, our algorithm presents a competitive performance.

In addition to a desirable sensitivity performance, the proposed method has a good speed performance. A computer with Intel Core 2 Duo CPU 3.1 GHz is used for evaluation. By testing the 8 records, the average time consumed for detection on a 30-min original signal is 5.78 s, which is much faster than other methods in Table 2 (all more than 10 s). This is due to the new detection order, where the R-wave is detected after QRS is extracted, which is different from the usual method. The good speed performance may help the application in uninterrupted ECG monitoring system.

Table 1
Statistical data of R-peak detection results

Record	Annotated R-peaks	Detected R-peaks	Time Consumed (s)	TP	FN	FP	Se (%)	P ⁺ (%)
#100	2265	2265	5.935	2264	1	1	99.96	99.96
#101	1860	1862	5.186	1857	3	5	99.84	99.73
#103	2078	2077	5.870	2076	2	1	99.90	99.95
#105	2041	2052	5.736	2036	5	16	99.76	99.22
#109	2526	2520	5.970	2518	18	12	99.29	99.52
#115	1946	1945	5.462	1945	1	0	99.95	100
#122	2468	2468	5.904	2468	0	0	100	100
#209	3017	2998	6.218	2997	20	1	99.34	99.97
Total	18201	18187	46.281	18161	50	36	99.72	99.80

Table 2
Result with several algorithms

Reference	Method	Number of beats	Se (%)
Proposed in this paper	Wavelet Transform and K-means Clustering (MLII)	18201	99.72
Gritzali [6]	Length and energy transformation	14292	99.60
S.S. Mehta [5]	K-means algorithm	17856	98.63
Chouhan and Mehta [7]	Adaptive quantized threshold (Single-lead algorithm)	17856	98.56
Mehta and Lingayat [8]	Support vector machine (12-lead algorithm)	17856	99.93

4. Conclusion

The method proposed in this paper is simple and easy to implement. Especially for the long-term ECG monitoring, it can provide fast detection with high accuracy. A wavelet transform is an efficient tool for analysing an ECG signal. With a WT, the baseline drift is fully eliminated, and noise from the industrial frequency band is suppressed well. The segmented K-means clustering method not only has better accuracy, but also satisfies the uninterrupted monitoring needs. Because the R-peaks are elected by comparing amplitudes in each QRS region respectively and dynamically, this method can adapt to different ECG situations. Due to the combination of these simple algorithms, this detection method is qualified for an intelligent ECG monitoring system.

Acknowledgment

This work was financially supported by National Natural Science Foundation of China (61173086) and the Shandong Province Science and Technology Development Planning (2014GSF118152) and Shandong Province Natural Science Foundation (ZR2015FM028).

References

- [1] V.X. Afonso, et al., ECG beat detection using filter banks, *IEEE Transactions on Biomedical Engineering* **46** (1999), 192-202.
- [2] S. Kadambe, et al., Wavelet transform-based QRS complex detector, *IEEE Transactions on Biomedical Engineering* **46** (1999), 838-848.
- [3] M.G. Strintzis, G. Stalidis, et al., Use of neural networks for electrocardiogram (ECG) feature extraction, recognition and classification, *Neural Network World* **3** (1992), 313-327.
- [4] S.C. Fanga and H.L. Chan, QRS detection-free electrocardiogram biometrics in the reconstructed phase space, *Pattern Recognition Letters* **34** (2013), 595-602.
- [5] S.S. Mehta, D.A. Shete, et al., K-means algorithm for the detection and delineation of QRS-complexes in Electrocardiogram, *IRBM* **31** (2010), 48-54.
- [6] F. Gritzali, Towards a generalized scheme for QRS Detection in ECG waveforms, *Signal Process* **15** (1988), 183-192.
- [7] V.S. Chouhan and S.S. Mehta, Detection of QRS-complexes in 12-lead ECG using adaptive quantized threshold, *International Journal of Computer Science and Network Security* **8** (2008), 155-163.
- [8] S.S. Mehta and N.N. Lingaya, Combined entropy based method for detection of QRS complexes in 12-lead electrocardiogram using SVM, *Computers in Biology and Medicine* **38** (2008), 138-145.
- [9] M. Elgendi, et al., Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems, *Plos One* **9** (2014), e84018.
- [10] M. Raeiatibanadkooki, et al., Real time processing and transferring ECG signal by a mobile phone, *Acta Informatica Medica* **22** (2014), 389-392.
- [11] M. Patil, et al., Evaluation of QRS complex based on DWT coefficients analysis using daubechies wavelets for detection of myocardial ischaemia, *Journal of Mechanics in Medicine & Biology* **210** (2012), 273-290.
- [12] Z. Zidelmal, et al., QRS detection based on wavelet coefficients, *Computer Methods & Programs in Biomedicine* **107** (2012), 490-496.
- [13] A. Rute, et al., Fetal QRS detection and heart rate estimation: A wavelet-based approach, *Physiological Measurement* **35** (2014), 1723-1735.