Automatic Digital ECG Signal Extraction and Normal QRS Recognition from Real Scene ECG Images

Shuang Wang , Shugang Zhang , Zhen Li , Lei Huang , Zhiqiang Wei

PII: S0169-2607(19)30139-7
DOI: https://doi.org/10.1016/j.cmpb.2019.105254
Reference: COMM 105254

To appear in: *Computer Methods and Programs in Biomedicine*

Received date: 29 January 2019
Revised date: 30 October 2019
Accepted date: 29 November 2019

Please cite this article as: Shuang Wang , Shugang Zhang , Zhen Li , Lei Huang , Zhiqiang Wei , Automatic Digital ECG Signal Extraction and Normal QRS Recognition from Real Scene ECG Images, *Computer Methods and Programs in Biomedicine* (2019), doi: https://doi.org/10.1016/j.cmpb.2019.105254

Highlights

- An Adaptive Filter Algorithm (AFA) is proposed in this paper for uneven illumination problem. The AFA method relieves the effect of uneven illumination that commonly exists in real-scene ECG image to a large extent.
- A channel dependent hierarchical scheme for automatic ECG binary image extraction from ECG image in real scene is introduced. The two layers aim separately at gridline removing and illumination relieving depending on different channels' properties, but they also complement each other and finally generate a good denoised ECG signal. 1D ECG signal extraction is based on ECG binary image and low difference between extracted value and real value shows excellent ability of the proposed approach.
- We also bring forward a new QRS recognition method using fused image features. This provides a way to utilize spatial relationships of special points to assist in disease diagnose. A candidate point set strategy is used for further computing optimization.

Computer Methods and Programs in Biomedicine

Automatic Digital ECG Signal Extraction and Normal QRS Recognition from Real Scene ECG Images

Shuang Wang, Shugang Zhang, Zhen Li, Lei Huang, Zhiqiang Wei*

Department of Computer Science and Technology, Ocean University of China, Qingdao 266100, China;

ws@stu.ouc.edu.cn; zhangshugang@hotmail.com; lizhen0130@gmail.com; huangl.ouc@gmail.com; weizhiqiangouc@163.com

* Correspondence: weizhiqiangouc@163.com

Abstract:

Background and objective: Electrocardiogram (ECG) is one of the most important tools for assessing cardiac function and detecting potential heart problems. However, most of the current ECG report records remain on the paper, which makes it difficult to preserve and analyze the data. Moreover, paper records could result in the loss significant data, which brings inconvenience to the subsequent clinical diagnosis or artificial intelligence-assisted heart health diagnosis. Taking digital pictures is an intuitive way of preserving these files and can be done simply using smartphones or any other devices with cameras. However, these *real scene* ECG images often have some image noise that hinders signal extraction. How to eliminate image noise and extract ECG binary image automatically from the noisy and low-quality real scene images of ECG reports is the first problem to be solved in this paper. Next, QRS recognition is implemented on the extracted binary

images to determine key points of ECG signals. 1D digital ECG signal is also extracted for accessing the exact values of the extracted points. In light of these tasks, an automatic digital ECG signal extraction and normal QRS recognition from real scene ECG images is proposed in this paper.

Methods: The normal QRS recognition approach for real scene ECG images in this paper consists of two steps: ECG binary image extraction from ECG images using a new two-layer hierarchical method, and the subsequent QRS recognition based on a novel feature-fusing method. ECG binary image extraction is implemented using sub-channel filters followed by an adaptive filtering algorithm. According to the ratio between pixel and real value of ECG binary image, 1D digital ECG signal is obtained. The normal QRS recognition includes three main steps: establishment of candidate point sets, feature fusion extraction, and QRS recognition. Two datasets are introduced for evaluation including a real scene ECG images dataset and the public Non-Invasive Fetal Electrocardiogram Database (FECG).

Results: Through the experiment on real scene ECG image, the F_1 score for Q, R, S detection is 0.841, 0.992, and 0.891, respectively. The evaluation on the public FECG dataset also proves the robustness of our algorithm, where F_1 score for R is 0.992 (0.996 for thoracic lead) and 0.988 for thoracic S wave.

Conclusions: The proposed method in this article is a promising tool for automatically extracting digital ECG signals and detecting QRS complex in real scene ECG images with normal QRS.

Keywords: Adaptive Filter Algorithm (AFA); ECG signal extraction; ECG image; QRS recognition

1. Introduction

According to the report of American Heart Association, 11.5% of American adults (27.6 million) were diagnosed with heart disease [1]. Electrocardiogram (ECG), a record of the electrical activity of the heart, is an important clinical tool for diagnosing cardiovascular diseases. An accurate and long-term ECG recording can not only help to evaluate the functional alterations of the heart or some other circulation related diseases, especially for the unhealthy population with cardiac problems or the pregnant population for fetal heart rate detection [2], [3], but also benefit medical research by providing precious clinical data. However, most ECG devices, including conventional 12-lead electrocardiograph and cardiotocograph that records fetal heartbeat still provide paper reports. Those digital and high-quality ECG signals are not always preserved in the machines. In fact, the ECG records in most cases are only available in printout clinical reports that are kept by individuals. This further hinders the popularization of Electronic Health Record (EHR) and poses problems for data management as these paper reports are not convenient for collecting, storing and analyzing. Moreover, deep learning methods have been utilized on ECG signal[4] , which requires a large amount of training data. For example, a recent study [5] by Andrew Y. Ng team, which collected 64,121 ECG records from 29,163 patients for arrhythmia classification. Therefore, a large number of printed ECG images need to be processed and extracted.

Aiming at the above problems, a reasonable solution is to construct a robust ECG signal extraction algorithm, which could accurately separate ECG signals from noisy real scene photos using image processing techniques. Digital signals can be easily stored in personal health electronic record and used for medical research after erasing personal information. Generally, two problems needed to be solved in a real scene ECG signal extraction task. The first one is grid removing. The grid in ECG images is an auxiliary tool to measure the signal. The grid is organized exactly as 1mm×1mm square units, with one horizontal unit indicating 0.04 seconds and the vertical unit indicating 0.1 mV. Removing these grids is a challenging task, especially in photos obtained in real scene with different illumination environments, because the low quality of these images results in that part of the grids are indistinguishable and cannot be removed accurately. Also, the upstroke of R wave could be very thin, which requires a well-crafted approach to capture. Secondly, the uneven illumination problems should be solved when pictures are taken in real scene. This problem makes the grid removing task more difficult as ECG traces and grids in different illuminated parts should be processed separately. To the best of our knowledge, the researches about ECG image processing in real scene are still relatively rare. With respect to grid elimination, most efforts[6]–[11] are devoted to scanned ECG images, in which illumination problem is not existing. These approaches mostly followed a procedure of color image - gray image - binary image, and then the binary image is transferred into one-dimensional (1D) digital signal. These techniques achieved good performance on scanned images, but they

are still not robust enough for real scene cases. Mitra et al. proposed a threshold based ECG extraction method for non-scanned ECG pictures [12], which partially solved the grid removing and insufficient light problems. However, the proposed method additionally caused the edge noises and the loss of ECG signal information.

Relying on the extracted ECG binary image, some basic properties of ECG can be determined. The QRS complex, a combination of three waves on a typical electrocardiogram, which corresponds to the depolarization of ventricles and has important clinical significance. The structure, duration and amplitude of QRS complex are important indicators and evidences for cardiovascular disease including conduction abnormalities, ventricular hypertrophy, and myocardial infarction, etc.

Currently, lots of approaches of QRS recognition have been proposed and achieved good performances. Wavelet transform [13], [14], Hibbert transform [15], [16] and QRS waveform templates [17] are all used for QRS recognition. There are also some other methods which utilized adaptive filter to recognize QRS [18], [19]. Moreover, deep learning methods [20] are also introduced for ECG biometric recognition. These methods were applied on 1D ECG signal, which focused on the properties of 1D signal. But the ECG morphological characteristics which play a vital role in the diagnosis of disease, are visualized by 2D image. Some special image features such as Local texture feature and morphological feature will be omitted if they are transformed into 1D signal. Due to this consideration, 1D ECG signals are projected onto a 2D space including P, QRS complex, and T waves for

user recognition based on convolutional networks [21]. There is rare direct method of recognizing QRS in ECG binary image, especially for real scene images. The water reservoir [22] feature for QRS recognition in ECG image labels the crest portion as R and the trough portion as S or Q, which regards QRS complex as a reservoir. It captures the general shape of the QRS but it is not robust for QRS complex when the big QS height difference is very large.

The main contributions of the algorithm proposed in this article are summarized as follows:

1. An Adaptive Filter Algorithm (AFA) is proposed in this paper to solve the problem of uneven illumination. The image is divided into small blocks and a filter is designed for elimination of illumination in each block. If the value distribution of each block changes, the filter adaptively changes to make the illumination as background all the time.

2. A channel dependent hierarchical scheme for automatic ECG binary image extraction from ECG image in real scene is introduced. The two layers separately aim at removing gridline and illumination depending on different channels' properties, but they also complement each other and finally generate a pure ECG binary image. 1D ECG signal extraction is based on ECG binary image and low difference values show excellent ability of the proposed approach.

3. A QRS recognition method based on fused image features is proposed. This provides a way to utilize spatial features of QRS complex for disease diagnosis. A candidate point set strategy is adopted for further computational optimization.

2. Methods

The automatic digital ECG signal extraction and normal QRS recognition from real scene ECG images is composed of two parts: ECG binary image extraction and normal QRS recognition using feature fusion. In the first part, a hierarchical algorithm for ECG binary image extraction in real scene is proposed. The first layer focuses on signal integrity and the second layer focuses on signal purity. The second part QRS recognition with feature fusion is implemented on ECG binary image. The feature fusion consisting of morphological feature and local texture feature is designed. Candidate points for QRS are screened depending on the corner commonality present in the image. QRS recognition using QRS feature fusion is carried out on this candidate set. Besides, 1D signal can also be extracted based on ECG binary image.

2.1 Hierarchical algorithm for ECG binary image extraction

A colorful image generally contains three channels: R(red), G(green), B(blue). The image under each channel can be described as matrices $R_{(m,n)}$, $G_{(m,n)}$ and $B_{(m,n)}$ respectively. m and n represents the number of rows and columns of the matrix. The pixel value of matrix for each channel is between the range of 0~255. Since different image channels have their own characteristics that the gradient and distribution of pixel values are quite different. Pixel values are clustered into three levels by 3-means clustering [23] in each channel. The three levels are displayed in Figure 1 (a). It can be observed that the illumination effect is more obvious in R

channel and grid lines mostly appear in G, B channels. Inspired by these observations, a hierarchical algorithm for ECG binary image extraction consisting of two layers is implemented on R,G,B. Overall framework is depicted in Figure 1(b) and an illustrating example is shown in Figure 1(a).

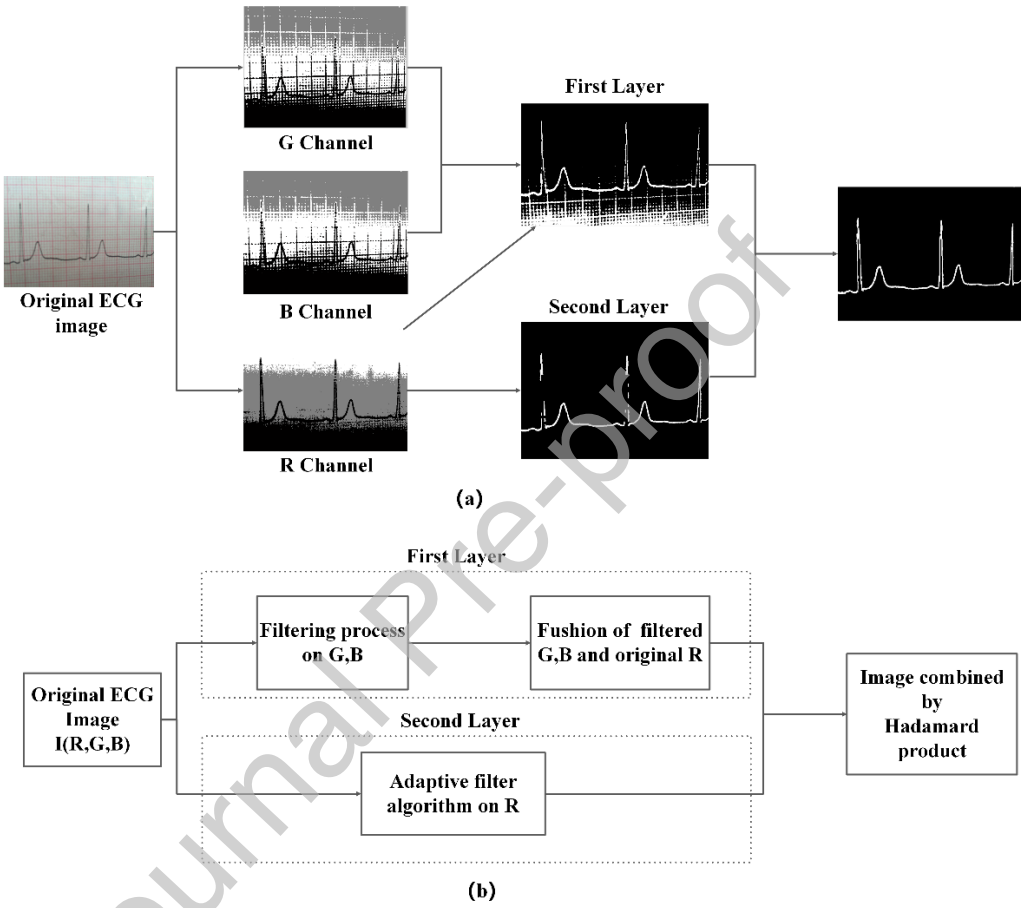

Figure 1. Framework for hierarchical algorithm for ECG binary image extraction: (a) An illustrating example of ECG binary image extraction from ECG image in real scene; (b) Framework of hierarchical algorithm.

2.1.1 First layer: Sub-channel filtering

The first layer is implemented on G, B channels for grid filtering. As grids themselves own a relatively high frequency, they can be well removed in the frequency domain of G, B channels. Two-dimensional Fourier transform and

Butterworth filter are utilized in frequency domain of G and B channels to remove grids, and then the two channels are transformed back into spatial domain by inverse Fourier transformation. Finally, these two processed channels are stacked on the original R channel (which remains unchanged) to get the output 3-channel image I'. In order to extract the ECG signal waveform, image binarization approach by Otsu et al. [17] is applied on I'. Above procedure is summarized as:

$$I' = \text{Binary}([R, G', B']) \quad (1)$$

I': Output of the first layer, i.e., the stacked channels.

R: R channel signal of original image.

G', B': G and B channel signal after processed by Butterworth filter on frequency domain.

The output of first layer is illustrated in the image of top line and the third column within Figure 1 (a). It can be observed that about two thirds of the grid lines are removed. Nevertheless, there are still some grid lines that have not been removed yet due to the uneven illumination effects. The second layer is further introduced to tackle it.

2.1.2 Second layer: An Adaptive Filter Algorithm (AFA)

Targeting on the R channel, the second layer conducts an adaptive filter algorithm (AFA) on it to eliminate the effects of illumination. Specifically, the matrix of R channel $R_{(m,n)}$ is firstly divided into numerous non-overlapping image blocks with the stride of one tenth of the height of the image. The next step is to

adaptively decide whether a certain block contains signal or not. This can be determined by calculating variance for each block, which indicates the degree of dispersion of the block. In a block that contains the ECG signal, the variance is prone to be larger (more dispersive) than background blocks. By defining a mean value of variances of all blocks as a threshold, all those blocks are thus classified into two classes. The one that has a larger variance than the threshold is actually the class containing the signal. This kind of blocks are further processed by Ostu algorithm [24]. For the other class, we set 0 to them as background. At last, these blocks are re-assembled into an entire binary image. Algorithm is summarized in Algorithm1 and illustrated in Figure 2.

Algorithm1 Adaptive filter algorithm

Input: Original R channel image matrix $R_{(m,n)}$

Output: Binary ECG image $I_{(m,n)}$ without illumination

1: Select an adaptive filter F with fixed size $f \times f$;

2: Padding $R_{(m,n)}$ with neighbor pixel value to make

$$(m+a)\%f \to 0 \;\&\&\; (n+b)\%f \to 0$$

3: Divide the image matrix $R_{(m,n)}$ into K image blocks, $K = \frac{(m+a)}{f} \times \frac{(n+b)}{f}$;

4: Set the stride→f, $B_k \to 0$;

5: Set the adaptive filter $F \to \bar{x}$ of corresponding image block $im_{(f,f)}$, $B_k \to \frac{\sum_{i,j=1}^{n}(F_{(i,j)}-im_{(i,j)})^2}{f \times f}$;

6: All image blocks obtain their specified B_k. Calculate the average value of all B_k as $\bar{B}$.

7: if $B_k \le \bar{B}$ then

8: Corresponding image block $im \to 0$;

9: if $B_k > \bar{B}$ then

Corresponding image block $im \to Ostu(im)$;

10: Image blocks are assembled to get a new binary image containing ECG signal and no illumination.

Comparing with global threshold method, the proposed adaptive filter algorithm that applied on local block units effectively takes local properties of image into consideration by measuring the dispersion degree using variance (which is proved in our experiments to be one of the inherent differences between no-signal blocks and signal blocks), thus successfully extract the ECG traces. The subsequent Ostu algorithm generates a binary image that offers idealized input image for the QRS recognizer in the next stage.

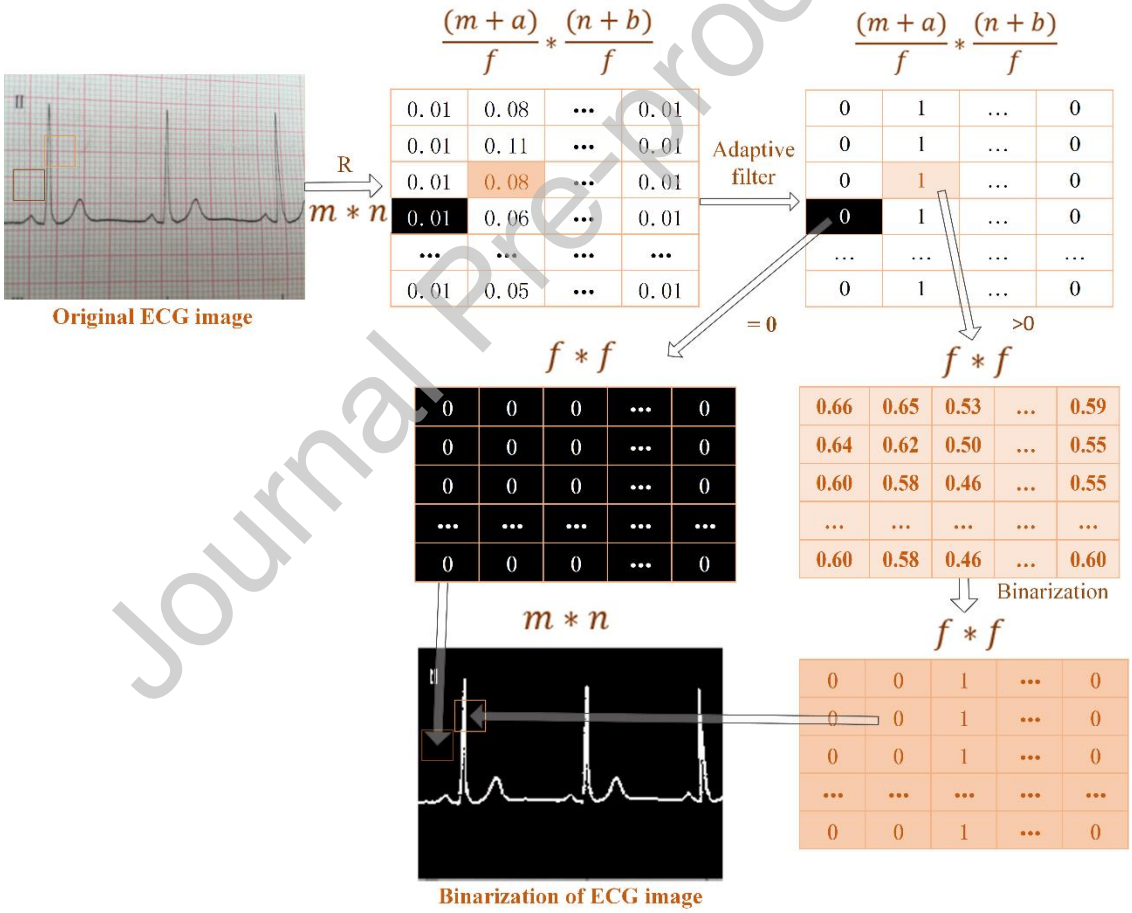

Figure 2. Adaptive filter algorithm for illumination noise elimination: The image of R channel is divided into several minor blocks with the same size of filter F. The adaptive filter distinguishes ECG waveform and background by calculating the difference of pixels in each small block. The blocks with background are set 0 and

blocks including ECG waveform are processed by Ostu algorithm to separate waveform from background. After the procedure of the algorithm, binary image with white ECG waveform and black background is acquired.

2.2 QRS recognition based on feature fusion

2.2.1 Candidate point set

After ECG binary image extraction, ECG signal is presented as a white signal line with black background. If feature extraction is performed on the whole signal, it would be computationally expensive. To minimize the time cost, series of candidate points that might be Q,R,S, were firstly elected as *candidate point set* using Harris corner detection [25] before feature extraction. This is inspired by the observation that Q, R, S points are mostly the corner points in 2D image, which can be well extracted using Harris corner detection.

2.2.2 Feature fusion for QRS

For QRS recognition, our algorithm was carried on the basis of fusion feature of local texture feature and morphological feature. The characteristics of different kinds of points are diverse. So two different methods for feature extraction are used: combination of local texture feature and shape feature for R, combination of local texture feature and spatial feature for Q, S.

- The Local texture feature

One of the characteristics that differ QRS points from other candidate points is that they are prone to have sharp gradients (peak-shape) comparing with others.

This unique local morphology could be well described by Local Binary Patter (LBP), a delicately designed local texture descriptor that taking into account information of neighboring pixels around each point [26]. Specifically, a 3×3 local grid is obtained for each candidate point. If the value of a neighbor point is bigger than the center point, it is assigned to 1, otherwise it is assigned to 0. The pixel values (0 or 1) of eight neighbors are weighted by 2^n (n is the position index in the 3×3 grid ranging from 0 to 7). The sum of all the weighted pixel values is the final LBP feature value for the center pixel. LBP values range from 0 to 255.

Although different candidate points have diverse LBP values, the values of same kind of ECG key points (Q, R, S) are similar. In training process, LBP values for each kind of key points are grouped in a value set (i.e., LBPQ, LBPR, LBPS). Take Q for insurance, if there are n points of Q with their LBP values stored in LBPQ via training stage, the predicted probability of candidate point p being Q is:

$$\mathrm{P}_{LBP}(p = Q) = \frac{\sum_{i=1}^{n} f_{LBP}(p, p_i)}{n} \tag{2}$$

where p_i is the labeled Q in training process, f_{LBP} is a simple function for comparing the LBP values of two points and can be calculated as

$$f_{LBP}(p_x, p_y) = \begin{cases} 1 & if\ LBP(p_x) = LBP(p_y) \\ 0 & else \end{cases} \tag{3}$$

If there are large number of points in LBPQ owning the same LBP value with that of candidate point p, the probability of p being Q will reasonably increase. Equations for R and S can be obtained in similar way.

- The morphological feature

The QRS complex has its own unique morphological properties. For instances, R is always the highest point that is easy to be determined, while Q and S can be subsequently located by calculating appropriate horizontal distance and scanning the rest points towards left (for Q) and right (for S) from the detected R peak. These properties are exactly the morphological features that our algorithm attempts to extract.

R is determined firstly as it is relatively easy to recognize using the amplitude features. A frequency histogram that indicates the value distribution of R amplitude, is established in training process. The feature fusion of R point contains LBP value and amplitude value. The fusion feature of R can be described as follows:

$$W(p,R) = \alpha * \mathrm{P}_{LBP}(p = R) + \beta * fre_{\,ampli(p \in R)} \quad (4)$$

where p is the candidate point, $\mathrm{P}_{LBP}(p = R)$ is the calculated LBP probability in last section, $fre_{\,ampli(p \in R)}$ can be determined from the frequency histogram, α and β represent the weights for two features and are assigned to 0.3 and 0.7 separately.

The horizontal distances of QR or RS are utilized as the global morphological feature for Q and S. It is noticeable that this can only be obtained after R is determined. Here we highlight how the distance feature is acquired supposing that R has been correctly detected, and more details of the process flow will be given in the next section. Specifically, the distance between each point and the nearest R peak is calculated and termed *Horizontal Distance from R (HDR).* Similar to the frequency histogram of R amplitude, two HDR frequency histograms are also

obtained for Q and S in the training process. The feature fusion for Q and S point contains LBP value and *HDR* value, i.e., the fused feature for Q can be described as follows:

$$W(p,Q) = \alpha * \mathrm{P}_{LBP}(p = Q) + \beta * fre_{HDR(p \in Q)} \quad (5)$$

Where p is the candidate point, $\mathrm{P}_{LBP}(p = Q)$ is the calculated LBP probability in the last section, $fre_{HDR(p \in Q)}$ can be determined from the frequency histogram, α and β represent the weights for two features and are assigned to 0.3 and 0.7 separately.

2.2.3 QRS recognition

In the real scene ECG image, there is more than one waveform in single lead. The algorithm first recognizes all R points in corresponding QRS complexes in ECG image. The point with the maximum y coordinate in the candidate set is defined as R_{peak} of the whole ECG image. $ampli(p)$ is the amplitude of point p which belongs to candidate point set. If it exists c_i satisfying with equ.6, this point is regarded as candidate of R point marked as R_c.

$$ampli(R_{peak}) - ampli(c_i) < \frac{1}{4}\, ampli(R_{peak}) \quad (c_i \in C) \quad (6)$$

If there are some points of R_c close in the x coordinate less than 1/6 length of R-R interval, the point with max $W(p,R)$ is selected as final R of each waveform.

In a consecutive ECG signal, Q or S point is located between two neighboring R points. Q and S could be determined by scanning towards both directions from the determined R. Thus two sequences (namely SQ and SS) containing possible Q and S

point collections are generated by collecting candidate points between two adjacent R, as shown in Fig. 3.

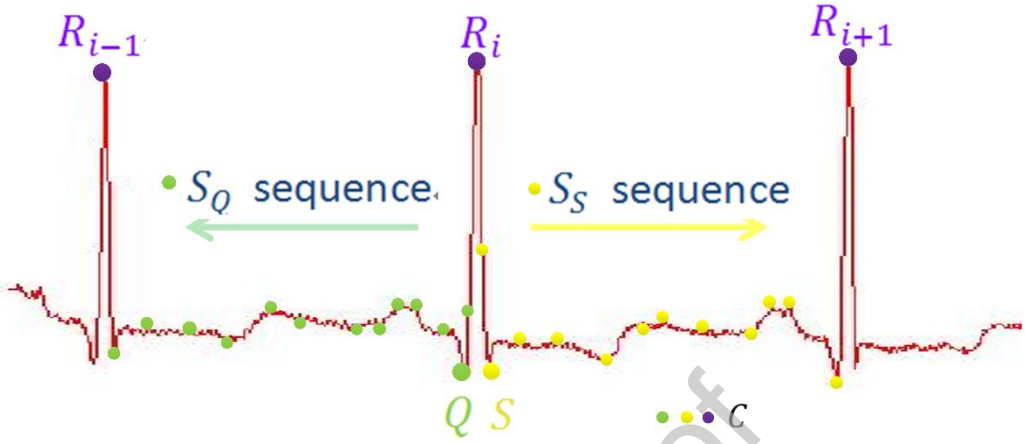

Figure 3. Calculation of QRS complex: for each R, Q point exists in S_Q sequence and S point exists in S_S sequence.

Finally, the combining probability using fused features are calculated as:

$$F(q,r,s) = W(q,Q) \times W(r,R) \times W(s,S) \quad (q \in S_Q, r \in R_c, s \in S_S) \tag{7}$$

If there are m points in S_Q, k points in R_c and n points in S_S, there are $m \times k \times n$ combinations of corresponding QRS complex. Calculating each $F(q,r,s)$ and the maximum of them is determined as the final QRS.

$$QRS = argmax_{q \in S_Q, r \in R_c, s \in S_S}\left(F(q,r,s)\right) \tag{8}$$

2.2.4 1D signal extraction

In addition to recognizing QRS directly from ECG binary image, 1D ECG signals can also be extracted. Since image is mathematically represented by a $m \times n$ pixel matrix. And each large grid which is composed of 25 small grids is mathematically represented by a $g_v \times g_s$ pixel matrix in initial image. Each large

grid represents 0.5 mV in height and 0.2 ms in length. According to the ratio between pixel and real value for the grid, the time position and amplitude of each point in the ECG signal can be calculated as follows:

$$\text{mvRate} = \frac{g_v}{0.5} \tag{9}$$

$$\text{msRate} = \frac{g_s}{0.2} \tag{10}$$

$$p_{mv} = \frac{p_{row}}{\text{mvRate}} \tag{11}$$

$$p_{ms} = \frac{p_{col}}{\text{msRate}} \tag{12}$$

where g_v indicates the pixel numbers of row, mvRate indicates the ratio of amplitude. g_s indicates the pixel numbers of col, msRate indicates the ratio of time. p_{row} is the row value for p point and p_{mv} is corresponding amplitude of point p. p_{col} is the col value for p point and p_{ms} is corresponding time position of point p. Generally, the width for ECG trace is more than one pixel and the medium row number of signal trace is used to represent p_{row}.

2.3 Datasets

To the best of our knowledge, there is no public dataset or any relevant criteria for real scene ECG images. To evaluate the performance of the proposed hierarchical algorithm, a real scene ECG image dataset was built. Images in this dataset were obtained in real scene from the book *Atlas of Complicated Electrocardiogram* [27]. To assess the performance of hierarchical algorithm for ECG binary image extraction on multiple image data sources, these images with 300dpi resolution cover four photographing angles and two kinds of light environment.

The dataset mainly covers lead Ⅱ. The amount of ECG obtained in real scene is 326, which contains four kinds of waveform numbers in a single image. 5-fold cross validation was applied to divide datasets into training data and testing data separately.

Another public dataset, *Non-Invasive Fetal Electrocardiogram Database (FECG)* [28] is also used for algorithm evaluation. The QS coordinates are manually labeled as the ground truth, and R coordinates are obtained from annotation documents. The fetal electrocardiogram (FECG) database contains a series of 55 multichannel fetal electrocardiogram (FECG) recordings. And each recording is consisted of six channels: two thoracic signals and four abdominal signals. Our experiments collected 3742 waveforms on thoracic signals and 4077 waveforms on abdominal signals.

Our algorithm mainly focuses on ECG signal with normal QRS with the shape of qRs and Rs. The ECG images containing these two kinds of QRS shapes were collected for datasets.

3. Results and discussion

3.1. Results for ECG binary image extraction

Representative results of the binary image extraction are shown in Figure 4, where the three columns are separately the original input images, the output images of *global thresholding* approach [11], and the output images of the proposed hierarchical algorithm in this article.

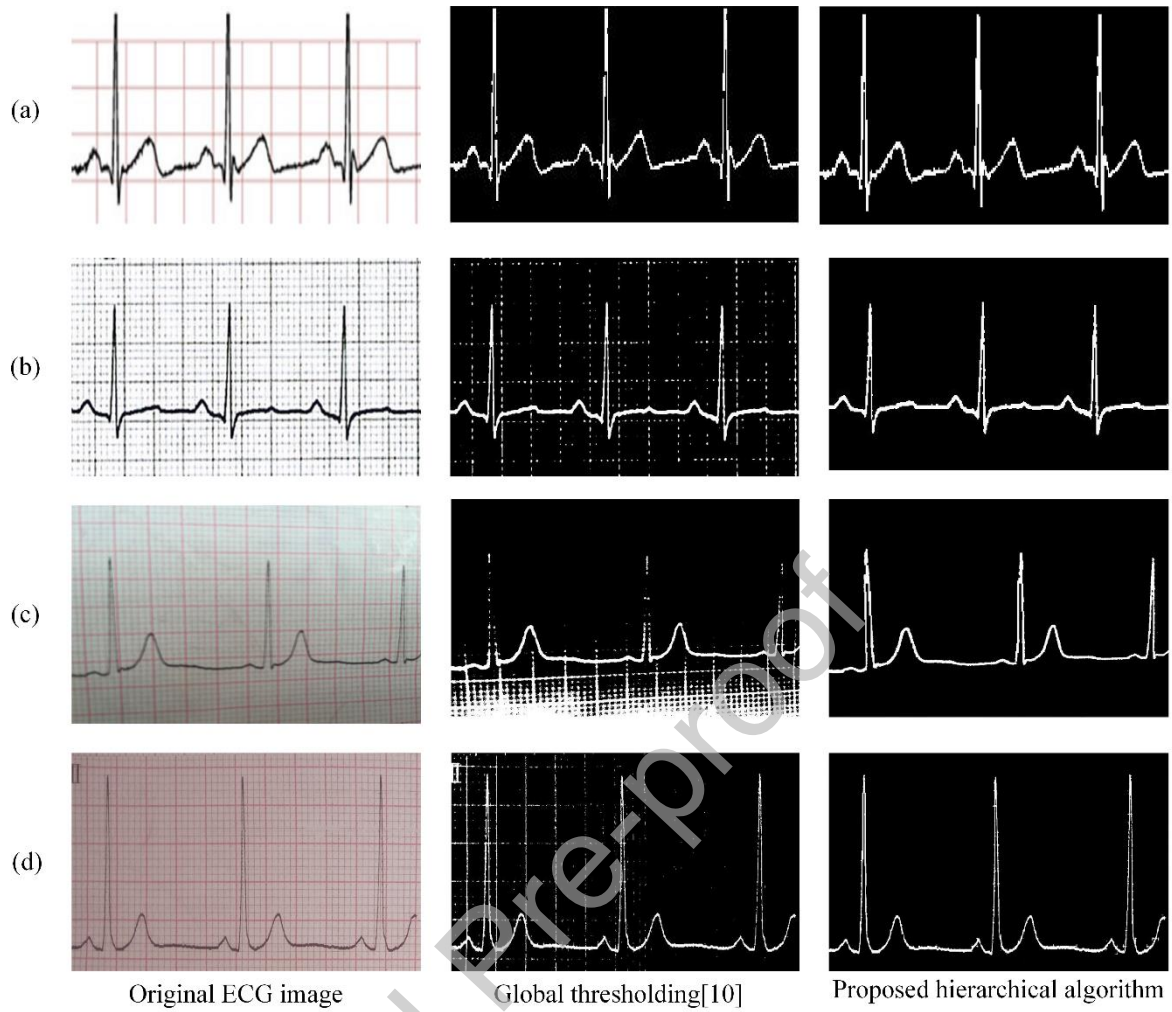

Figure 4. Results for binary image extraction. Four kinds of common real scene ECG images are shown in four rows (a-d), namely (a) scanned image with color grid; (b) scanned image with white and black grid; (c) ECG photo with uneven illumination from top to bottom; (d) ECG photo with uneven illumination from left to right;

From the scanned ECG image with red grids in Figure 4 (a), it can be observed that both two methods perform well. What is more, for those scanned ECG images with black-and-white grids in Figure 4 (b) where the grid pixel value is rather close to that of ECG signal trace, our method shows a better performance than the global thresholding approach. It is noteworthy that the three channels (R, G, B) in the

black-and-white ECG images are the same, so only AFA algorithm (second layer) was applied on those images. Finally, the last two rows present two uneven illuminated real scene ECG images either from bottom to top (Figure 4 (c)) or from left to right (Figure 4 (d)). The proposed method is also apparently superior to the compared algorithm in these two cases. In summary, the introduced method proves robust for common noises appearing in diverse kinds of real scene ECG images.

3.2 Results for QRS recognition using fusion feature

3.2.1 Candidate point set

Some examples for candidate points screened out from ECG signals are shown in Figure 5. The average number of points for each QRS trace was 2994 and the number of candidate points dramatically decreased to 86, meaning approximately 97% computation costs were saved.

Figure 5. Screened candidate point set

3.2.2 Accuracy of QRS recognition

In the recognition phase, the R peak was determined preceding Q and S recognition. For those images with multiple image size, images were normalized by R-R interval. Euclidean distance [29] of detected point and the corresponding ground truth point was calculated to determine whether the point was hit or not. Specifically, those paired points (i.e., the detected one and the ground truth one) that generated a Euclidean distance less than the linewidth of ECG trace were considered successfully recognized. In order to objectively evaluate the proposed method and keep in line with other results published in the literature, we used three statistical indexes, namely **Precision**, **Recall** and **F_1 score**, which are calculated as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{9}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{10}$$

$$F_1 = \frac{2\times\text{Recall}\times\text{Precision}}{\text{Recall+Precision}} \tag{11}$$

Precision refers to the percentage of correctly detected specified points (i.e., Q, R, S) out of all detected points, while **Recall** reflects the percentage of correctly detected specified points out of all actual specified points. **Precision** and **Recall** are commonly counterparts of each other considering the performance of a system, algorithm, etc. Thus we adopted **F_1 score** balancing both of them as the final criteria for our proposed approach. In equations, **TP** stands for the number of true positive samples (i.e., Q, R, or S that were correctly recognized), **FN** stands for false negative

samples (the number of Q, R, S that not spotted by the algorithm), and **FP** stands for false positive samples (non-QRS points but falsely recognized as Q, R, or S).

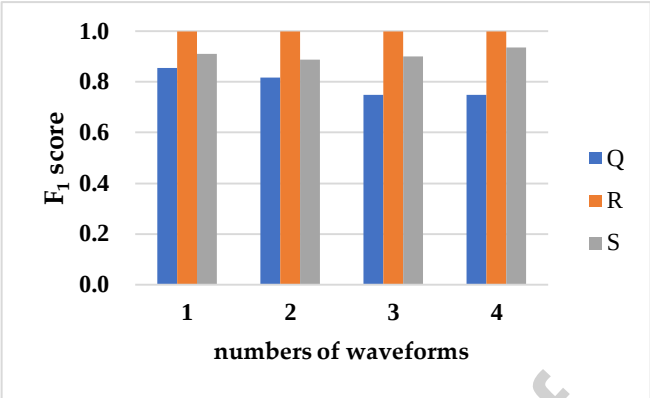

Figure 6. F_1 score of QRS recognition for different numbers of waveforms in real scene ECG image dataset

The result of QRS detection from ECG image in real scene is shown in Figure 6. For QRS detection from ECG image, if there are several waveforms in one image, the detection will be a bit more difficult due to waveform segmentation. Multi-wave recognition was considered in our recognition algorithm, so the proposed algorithm could handle recognition of multiple waveforms. To make a fair evaluation of our algorithm, we chose four different kinds of samples containing one to four waveforms in an ECG image. QRS detection was applied on each kind of sample and the accuracy is shown in Figure 6. It can be seen from the figure that the detection accuracy of QRS wave is not sensitive to QRS number, indicating that our algorithm has strong robustness to multi-wave detection. On average, our experiments achieved 0.992 F_1 score for R, 0.841 F_1 score for Q and 0.891 F_1 score for S with real scene ECG image dataset. F_1 score of R point is the highest among all the three waves, which is almost 1.0. As for F_1 score for Q and S,

the latter is higher than the former. This is because that the distribution of LBP value of Q is more disperse than that of S, which causes an impact on the probability calculation since that some LBP values have similar probability. Some of the results of recognition for QRS on real scene ECG image are shown in Figure 7

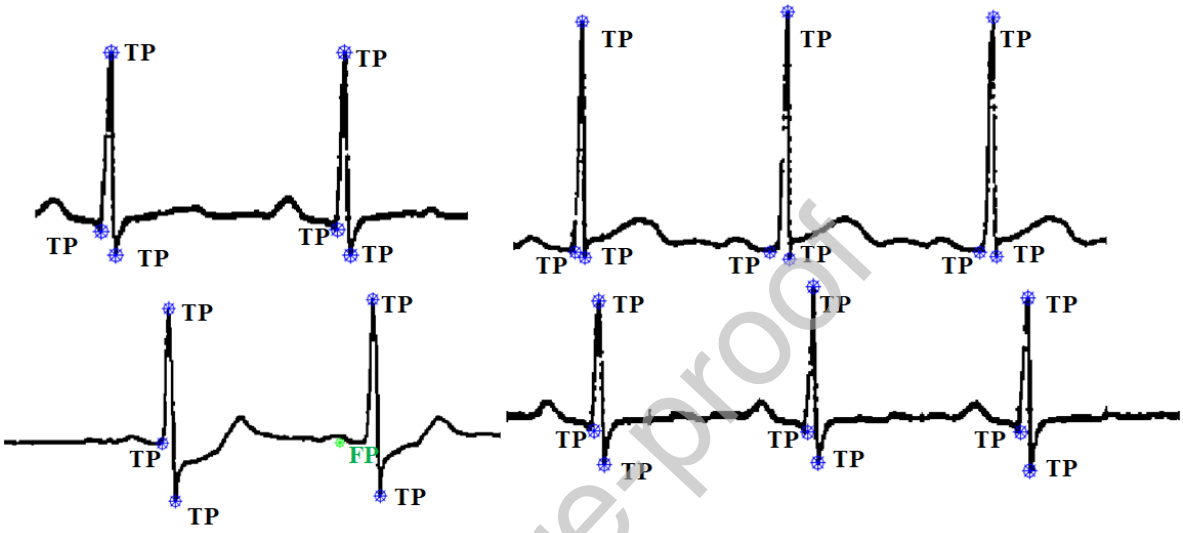

Figure 7. QRS recognition on Real scene ECG images

The results of QRS recognition on FECG database of two thoracic signals are shown in

TABLE I. This result mainly evaluates the proposed method of identifying special points with image features. A total of 3742 thoracic signal waveform segments were collected from FECG database. Noteworthy that the FECG data was transformed into 2D digital image, as shown in Figure 8. Since Q waves could not be accurately marked in FECG chest lead recording, we only tested R, S recognition using fusion features on these images. Without the influence of image noise, the

QRS recognition accuracy rate is significantly improved. As far as we know, there is rare studies on the identification of Q/S directly from ECG images, so these two are not compared with other methods.

TABLE I. QRS recognition results of thoracic signals in FECG database

	TP	FN	FP	Precision	Recall	F_1 score
R	3714	28	0	1	0.993	0.996
S	3657	39	46	0.988	0.989	0.988

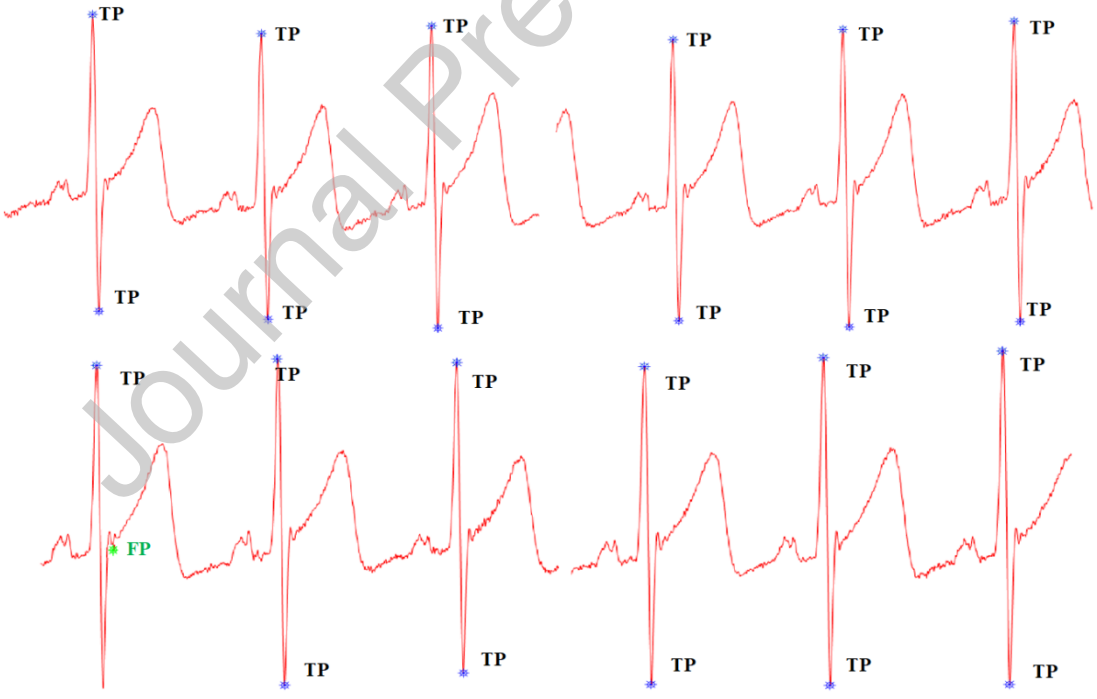

Figure 8. QRS recognition on FECG image (transformed from 1D signal)

As for R recognition, Hammad's algorithm [30] was compared with the proposed method in TABLE II. R detection was performed on all R peaks (including

3742 waveforms on thoracic signals and 4077 waveforms on abdominal signals) in FECG dataset. As shown in the TABLE II, the average F_1 score of Hammad's algorithm is 0.984. And the average F_1 score of the proposed algorithm is 0.992. Compared with Hammad's algorithm, the results of the proposed method show better performance. The average Recall of Hammad's algorithm is 1, and its Precision is 0.968. Its high Recall is related to its detection method that detects all peaks and divide them into P, R, and T separately. The average Recall of our algorithm is 0.984, and the Precision is 1. Compared with Hammad's algorithm, the Recall of both are nearly same, and the Precision of the proposed method is better.

TABLE II. Comparison of performance for R peak detection in FECG database

	Hammad's algorithm			**our algorithm**		
Record	**Re**	**Pr**	**F_1 score**	**Re**	**Pr**	**F_1 score**
102	1.000	1.000	**1.000**	0.987	1.000	0.993
115	1.000	0.970	0.985	0.980	1.000	**0.990**
127	1.000	1.000	**1.000**	0.981	1.000	0.991
154	1.000	0.966	0.983	0.983	1.000	**0.992**
192	1.000	0.966	0.983	1.000	1.000	**1.000**
244	1.000	1.000	**1.000**	0.988	1.000	0.994
252	1.000	0.969	0.984	0.985	1.000	**0.992**
274	1.000	1.000	**1.000**	0.981	1.000	0.991
300	1.000	0.968	0.984	0.978	1.000	**0.989**
308	1.000	0.935	0.966	0.968	1.000	**0.984**
323	1.000	0.968	0.984	0.988	1.000	**0.994**
368	1.000	0.937	0.968	0.990	1.000	**0.995**
384	1.000	0.967	0.983	0.990	1.000	**0.995**
392	1.000	0.970	**0.985**	0.964	1.000	0.982
410	1.000	0.939	0.969	0.986	1.000	**0.993**
416	1.000	0.944	0.971	0.988	1.000	**0.994**
436	1.000	1.000	**1.000**	0.991	1.000	0.996

444	1.000	0.914	0.955	0.970	1.000	**0.985**
445	1.000	0.966	0.983	0.986	1.000	**0.993**
515	1.000	1.000	**1.000**	0.972	1.000	0.986
571	1.000	0.968	0.984	0.990	1.000	**0.995**
585	1.000	1.000	**1.000**	0.990	1.000	0.995
595	1.000	0.917	0.957	0.980	1.000	**0.990**
597	1.000	0.968	0.984	0.980	1.000	**0.990**
621	1.000	0.969	0.984	0.994	1.000	**0.997**
Total	1.000	0.968	0.984	0.984	1.000	**0.992**

3.3 Results for 1D signal extraction

To assess the ability of the algorithm for extracting 1D signals from ECG binary image, another experiment was performed on MIT-BIH Arrhythmia Database [31], [32]. The records containing the shape of qRs and Rs in MIT-BIT were plotted per five seconds as ECG images with grids using Physionet [32] tools. These images were then printed out and photographed as real scene images. Finally, the proposed ECG binarization and 1D signal extraction were performed on this MIT-BIH real scene dataset to evaluate its performance regarding 1D signal extraction. Representative results for extracted 1D signal (blue) and real signal in MIT-BIH (red) are plotted in Figure 9 with highly overlapped traces, which stands for an effective and exact 1D extraction. Since the extracted 1D signal is primarily served to measure some clinical indexes, for instances, Q amplitude, QRS duration, etc., five QRS relevant indexes including amplitudes of Q, R, and S wave, as well as QRS duration and R-R interval were compared in this section.

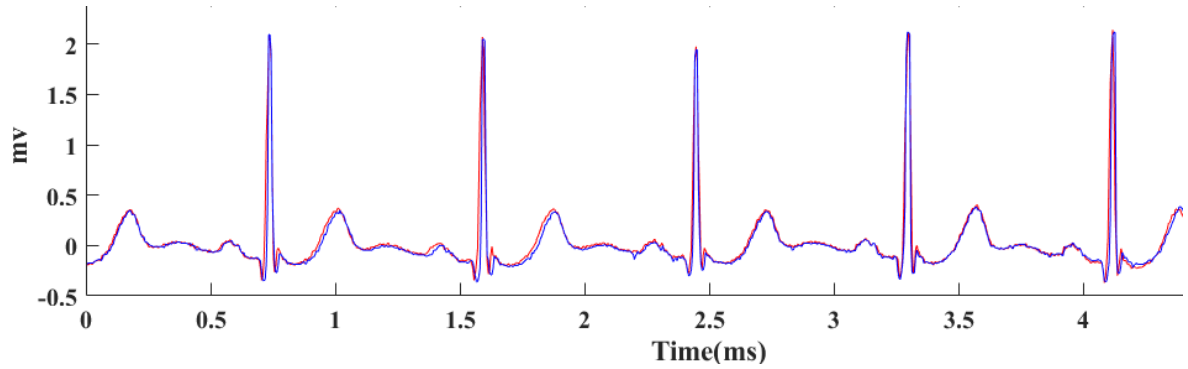

Figure 9 Comparison for extracted 1D signal (blue) and real signal in MIT-BIH (red)

Almost 1400 waveforms of MLII lead from MIT-BIH except for some abnormal QRS complexes were adopted for comparison. In order to demonstrate the performance of our algorithm for 1D signals extraction more intuitively, $Diff_i$ between the extracted value and the corresponding ground truth value is calculated to measure the difference through equ.12. Because MIT-BIH was obtained from different subjects and each subject owned unique ECG features, subject dependent $Mean$ and Std were utilized to measure the degree of deviation for real values and extracted values.

$$Diff_i = |x_i - \hat{x}_i| \tag{12}$$

$$Mean = \frac{\sum_{i=1}^{n} Diff_i}{n} \tag{13}$$

$$Std = \sqrt{\frac{\sum_{i=1}^{n}(Diff_i - Mean)^2}{n}} \tag{14}$$

where x_i is the extracted value and $\hat{x}_i$ is corresponding real value of i^{th} waveform. n is the number of waveforms in a subject's ECG record. $Diff_i$ indicates the difference between extracted value and real value.

The 1D signal extraction method was evaluated quantitatively by measuring the difference of five features (three vertical measurements and two horizontal measurements) including the amplitudes of Q, R, and S wave, as well as the QRS duration and R-R interval. The average amplitude differences of Q, R and S wave between extracted signal and original signal are respectively 0.018 mV, 0.026 mV, and 0.022 mV, which are all within a small range of 0.03 mV. And the average QRS duration and R-R interval differences are both only 0.007 ms. These quantitatively measured results suggest a close match of the extracted 1D signal and original record, which proves the effectiveness and robustness of the digital signal extraction method.

TABLE III. The difference between extracted 1D signal value and real value of MIT-BIH

	Q amplitude (mV)		**R amplitude (mV)**		**S amplitude (mV)**		**QRS duration (ms)**		**R-R interval (ms)**	
	Mean	**Std**	**Mean**	**Std**	**Mean**	**Std**	**Mean**	**Std**	**Mean**	**Std**
100	0.012	0.011	0.016	0.012	0.012	0.007	0.002	0.002	0.004	0.003
103	0.019	0.015	0.036	0.026	0.018	0.016	0.003	0.002	0.008	0.015
105	0.014	0.01	0.035	0.028	0.033	0.028	0.006	0.006	0.007	0.005
106	0.017	0.014	0.035	0.025	0.018	0.014	0.025	0.128	0.018	0.062
109	0.019	0.014	0.039	0.022	0.019	0.016	0.008	0.015	0.008	0.018
112	0.011	0.008	0.019	0.013	0.014	0.008	0.007	0.005	0.006	0.006
113	0.014	0.016	0.048	0.048	0.027	0.025	0.01	0.04	0.005	0.005
115	0.019	0.018	0.032	0.027	0.036	0.029	0.015	0.045	0.007	0.005
116	0.017	0.012	0.023	0.015	0.037	0.023	0.004	0.003	0.004	0.004
117	0.027	0.025	0.026	0.023	0.024	0.025	0.009	0.007	0.01	0.008
118	0.019	0.015	0.023	0.014	0.023	0.016	0.004	0.004	0.007	0.006
119	0.027	0.019	0.028	0.022	0.029	0.022	0.003	0.002	0.008	0.007
121	0.027	0.083	0.026	0.043	0.014	0.013	0.009	0.009	0.008	0.006
122	0.033	0.023	0.031	0.026	0.029	0.023	0.008	0.007	0.007	0.01
123	0.018	0.02	0.023	0.014	0.023	0.014	0.007	0.009	0.008	0.009
124	0.015	0.02	0.023	0.033	0.018	0.017	0.006	0.005	0.007	0.005

201	0.013	0.013	0.015	0.014	0.014	0.015	0.005	0.005	0.005	0.004
202	0.012	0.024	0.024	0.015	0.013	0.012	0.011	0.012	0.007	0.006
205	0.019	0.016	0.015	0.016	0.021	0.019	0.008	0.009	0.009	0.009
209	0.013	0.014	0.021	0.026	0.02	0.016	0.003	0.002	0.005	0.006
210	0.011	0.008	0.022	0.017	0.011	0.011	0.009	0.013	0.007	0.007
212	0.02	0.018	0.023	0.015	0.008	0.007	0.009	0.007	0.007	0.005
213	0.024	0.02	0.034	0.026	0.024	0.015	0.004	0.003	0.004	0.004
214	0.026	0.02	0.035	0.029	0.022	0.021	0.008	0.008	0.007	0.005
215	0.009	0.007	0.02	0.017	0.021	0.015	0.004	0.003	0.005	0.004
219	0.013	0.015	0.025	0.027	0.023	0.036	0.008	0.008	0.008	0.005
220	0.016	0.011	0.029	0.018	0.032	0.017	0.005	0.004	0.005	0.005
221	0.015	0.015	0.01	0.012	0.017	0.017	0.009	0.012	0.006	0.006
223	0.014	0.013	0.019	0.014	0.017	0.013	0.006	0.007	0.005	0.004
230	0.027	0.033	0.025	0.023	0.045	0.025	0.005	0.003	0.007	0.007
231	0.017	0.012	0.026	0.017	0.021	0.019	0.005	0.005	0.008	0.008
234	0.019	0.011	0.026	0.021	0.021	0.014	0.003	0.002	0.005	0.004
Ave	**0.018**	**0.018**	**0.026**	**0.022**	**0.022**	**0.018**	**0.007**	**0.012**	**0.007**	**0.008**

4. Conclusions

In this paper, a novel automatic digital ECG signal extraction and normal QRS recognition method applicable for real scene ECG images was proposed. The method is robust against grid or uneven illumination, and thus could precisely acquire the ECG binary image. A novel feature fusion method was then adopted to recognize normal QRS complex, which was proved to be effective on real scene ECG image datasets.

The proposed algorithm was evaluated on two datasets, including a real scene dataset collected by ourselves and a public dataset (FECG). And the feature-fusion based detection method achieved 0.992 F_1 score for R, 0.841 F_1 score for Q and 0.891 F_1 score for S in the first dataset. Moreover, the experiments on multi-wave images also achieved good performance, which suggested that our algorithm was robust on images containing multiple waveforms. We also evaluated the detection method

on the public dataset FECG and the method achieved 0.992 F_1 score for R (especially, 0.996 for thoracic R) and 0.988 F_1 score for thoracic S, in which F_1 score for R is superior to the state-of-the-art algorithm. As for 1D signal extraction, the amplitude differences are lower than 0.03 mV, besides the time differences of QRS duration and R-R interval are both 0.007 ms, all within a small error range.

The proposed method provids a feasible way for automatically extracting ECG signals from real scene ECG images and accurately detecting normal QRS complexes. In the future, we plan to further explore an effective way to detect the other two waves, i.e., P wave and T wave, to form a complete ECG waveform recognition framework. As for the abnormal ECG waveform, we will try to research on recognizing and classifying them in our future work. Also, disease dependent QRS recognition will be studied to support the cardiac disease classification. For example, the altered Q wave ("delta shape") is a representative morphological feature of woff-parkinson-white syndrome, which can be utilized for the disease dependent recognition and classification.

Funding: This research is supported by National Natural Science Foundation of China (No.61602430, No.61402428, No.61672475). It is also partially supported by The Aoshan Innovation Project in Science and Technology of Qingdao National Laboratory for Marine Science and Technology (No.2016ASKJ07).

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] A. H. A. S. Update, *Heart Disease and Stroke Statistics — 2018 Update A Report From the American Heart Association*. 2018.

[2] S. M. M. Martens, C. Rabotti, M. Mischi, and R. J. Sluijter, "A robust fetal ECG detection method for abdominal recordings," *Physiol. Meas.*, vol. 28, no. 4, p. 373, 2007.

[3] M. J. Rooijakkers, C. Rabotti, S. G. Oei, and M. Mischi, "Low-complexity R-peak detection for ambulatory fetal monitoring," *Physiol. Meas.*, vol. 33, no. 7, p. 1135, 2012.

[4] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," *Comput. Methods Programs Biomed.*, vol. 161, pp. 1–13, 2018.

[5] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," *arXiv Prepr. arXiv1707.01836*, 2017.

[6] H. K. Khleaf, K. H. Ghazali, and A. N. Abdalla, "Features Extraction Technique For Ecg Recording Paper," in *Presented at the International Conference on Artificial Intelligence in Computer Science and ICT (AICS 2013)*, 2013, pp. 243–248.

[7] S. Mitra, M. Mitra, and B. B. Chaudhuri, "Generation of digital time database from paper ECG records and Fourier transform-based analysis for disease identification," *Comput. Biol. Med.*, vol. 34, no. 7, pp. 551–560, 2004.

[8] M. C. Lewis, M. Maiya, and N. Sampathila, "A Novel Method for the Conversion of Scanned Electrocardiogram (ECG) Image to Digital Signal," in *International Conference on Intelligent Computing and Applications*, 2018, pp. 363–373.

[9] J. Wang, Y. Pang, Y. He, and J. Pan, "ECG Waveform Extraction from Paper Records," in *International Conference on Image and Graphics*, 2017, pp. 505–512.

[10] A. Rajani, "Digitization of Electrocardiography Data Sheet Through Image Processing Techniques.," *IUP J. Electr. Electron. Eng.*, vol. 9, no. 2, 2016.

[11] A. Sbrollini *et al.*, "Computer Methods and Programs in Biomedicine eCTG : an automatic procedure to extract digital cardiotocographic signals from digital images R," *Comput. Methods Programs Biomed.*, vol. 156, pp. 133–139, 2018.

[12] R. N. Mitra, S. Pramanik, S. Mitra, B. B. Chaudhuri, and F. Ieee, "Pattern Classification of Time Plane Features of ECG Wave from Cell-Phone Photography for Machine Aided Cardiac Disease Diagnosis," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 2014, pp. 4807–4810.

[13] J. P. V Madeiro, P. C. Cortez, F. I. Oliveira, and R. S. Siqueira, "A new approach to QRS segmentation based on wavelet bases and adaptive threshold technique.," *Med. Eng. Phys.*, vol. 29, no. 1, pp. 26–37, Jan. 2007.

[14] Z. Zidelmal, A. Amirou, M. Adnane, and A. Belouchrani, "QRS detection based on wavelet coefficients.," *Comput. Methods Programs Biomed.*, vol. 107, no. 3, pp. 490–6, Sep. 2012.

[15] D. Benitez, P. A. Gaydecki, A. Zaidi, and A. P. Fitzpatrick, "The use of the Hilbert transform in ECG signal analysis," *Comput. Biol. Med.*, vol. 31, pp. 399–406, 2001.

[16] N. M. Arzeno, Z. Deng, and C. Poon, "Analysis of First-Derivative Based QRS Detection Algorithms," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 2, pp. 478–484, 2008.

[17] C.-L. Chen and C.-T. Chuang, "A QRS detection and R point recognition method for wearable single-lead ECG devices," *Sensors*, vol. 17, no. 9, p. 1969, 2017.

[18] S. Jain, M. K. Ahirwal, A. Kumar, V. Bajaj, and G. K. Singh, "QRS detection using adaptive filters: A comparative study," *ISA Trans.*, vol. 66, pp. 362–375, 2017.

[19] S. Jain, A. Kumar, and V. Bajaj, "Technique for QRS complex detection using particle swarm optimisation," *IET Sci. Meas. Technol.*, vol. 10, no. 6, pp. 626–636, 2016.

[20] R. D. Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti, "Deep-ECG: Convolutional neural networks for ECG biometric recognition," *Pattern Recognit. Lett.*, 2018.

[21] M.-G. Kim, H. Ko, and S. B. Pan, "A study on user recognition using 2D ECG based on ensemble of deep convolutional neural networks," *J. Ambient Intell. Humaniz. Comput.*, pp. 1–9, 2019.

[22] S. Pramanik, R. N. Mitra, S. Mitra, and B. B. Chaudhuri, "A novel approach for delineation and feature extraction in QRS complex of ECG signal," in *Image Information Processing (ICIIP), 2011 International Conference on*, 2011, pp. 1–6.

[23] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 881–892, 2002.

[24] P. Smith, D. B. Reid, C. Environment, L. Palo, P. Alto, and P. L. Smith, "A Threshold Selection Method from Gray-Level Histograms," vol. C, no. 1, pp. 62–66, 1979.

[25] C. Harris and M. Stephens, "A combined corner and edge detector.," in *Alvey vision conference*, 1988, vol. 15, p. 50.

[26] Z. Li, Z. Wei, W. Jia, and M. Sun, "Daily Life Event Segmentation for Lifestyle Evaluation Based on Multi-Sensor Data Recorded by a Wearable Device," in *Conference: International Conference of the IEEE Engineering in Medicine & Biology Society IEEE Engineering in Medicine & Biology Society Conference*, 2013, pp. 2858–2861.

[27] Xia Liu, *ATLAS OF COMPLICATED ELECTROCARDIOGRAM*. Shanghai science and technology press, 2011.

[28] A. L. Goldberger *et al.*, "Physiobank, physiotoolkit, and physionet," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[29] L. Liberti, C. Lavor, N. Maculan, and A. Mucherino, "Euclidean distance geometry and applications," *Siam Rev.*, vol. 56, no. 1, pp. 3–69, 2014.

[30] M. Hammad, M. Ibrahim, and M. Hadhoud, "A novel biometric based on ECG signals and images for human authentication.," *Int. Arab J. Inf. Technol.*, vol. 13, no. 6A, pp. 959–964, 2016.

[31] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, 2001.

[32] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215--e220, 2000.