IEICE TRANS. INF. & SYST., VOL.E101–D, NO.8 AUGUST 2018

PAPER

ECG Delineation with Randomly Selected Wavelet Feature and Random Forest Classifier*

Dapeng FU[†], ***Nonmember***, **Zhourui XIA**[††], ***Student Member***, **Pengfei GAO**[†††], **Haiqing WANG**[††††a)], **Jianping LIN**[†††††], ***and*** **Li SUN**[†††††], ***Nonmembers***

SUMMARY Objective: Detection of Electrocardiogram (ECG) characteristic points can provide critical diagnostic information about heart diseases. We proposed a novel feature extraction and machine learning scheme for automatic detection of ECG characteristic points. Methods: A new feature, termed as randomly selected wavelet transform (RSWT) feature, was devised to represent ECG characteristic points. A random forest classifier was adapted to infer the characteristic points position with high sensitivity and precision. Results: Compared with other state-of-the-art algorithms' testing results on QT database, our detection results of RSWT scheme showed comparable performance (similar sensitivity, precision, and detection error for each characteristic point). RSWT testing on MIT-BIH database also demonstrated promising cross-database performance. Conclusion: A novel RSWT feature and a new detection scheme was fabricated for ECG characteristic points. The RSWT demonstrated a robust and trustworthy feature for representing ECG morphologies. Significance: With the effectiveness of the proposed RSWT feature we presented a novel machine learning based scheme to automatically detect all types of ECG characteristic points at a time. Furthermore, it showed that our algorithm achieved better performance than other reported machine learning based methods.

key words: *ECG, random forest, wavelet transform*

1. Introduction

Automatic annotation of electrocardiograms (ECGs) has received increasing attention because of its vital role in the diagnoses of several cardiac diseases [1]. Most of the clinical information in ECGs can be inferred from the intervals and amplitudes of the ECG characteristic points (the peaks and limits of the individual QRS waves, P wave, and T wave).

For determining the position of ECG characteristic points several QRS complex detection and ECG delineation methods are used. Some methods focus on detection of QRS complexes [2]–[4], while others use empirical mode decomposition [3] and threshold-independent QRS detection algorithm [4]. However, most of the QRS detection methods cannot be straightforwardly applied to the detection of P and T waves. Delineation methods like low-pass differentiation (LPD) [5], [6], hidden Markov models [7], and spline representation [8], [9] can delineate P and T wave only with pre-defined model and manually adjusted data-sensitive parameters.

The wavelet transform is one of the most popular techniques in ECG characteristic point detection [10]. Inspired by this method, Martinez et al. [11] developed a single-lead ECG delineation system. Based on an improved QRS complex detection method proposed by Li et al. [10], their system estimated the P and T wave peaks, on-sets, and off-sets, which showed acceptable detection accuracy on public ECG databases. Researchers tried to apply different methods to improve the delineation accuracy. To improve the delineation performance on T wave, Chen et al. [12] modelled three categories of T wave and designed their decision rule accordingly. P wave detection and delineation are also achieved with phase free stationary wavelet transform proposed by Lenis et al. [13]. Besides, Dumont et al. [14] applied evolutionary algorithms to tune parameters of an ECG wavelet transform-based delineator, achieved similar detection accuracy like Martinez et al. [11].

Based on partially collapsed Gibbs sampler, Bayesian methods were proposed by Lin et al. [15], [16]. By exploiting the strong local dependency of ECG signals, the methods showed relatively high detection accuracy of ECG characteristic points on the QT database. However, their methods depend on their prior knowledge and pre-defined models.

Some recent studies adopted statistical machine learning techniques for the detection of ECG characteristic points. Saini et al. [17] proposed a K-Nearest Neighbor classification approach for ECG recognition, which was evaluated on multiple ECG databases [18]. However, the K-NN method has the drawback of high dimensionality, and the trained K-NN classifier model occupies lots of memory. In our previous work, a time-domain feature termed as randomly selected signal pair difference (RSSPD) was proposed for ECG characteristic point detection [19]. This detection scheme was preliminarily evaluated via pre-selected 30 records on QT database and achieved acceptable perfor-

Manuscript received December 21, 2017.
Manuscript revised April 5, 2018.
Manuscript publicized May 9, 2018.

[†]The author is with Beijing Zhong Guan Cun Hospital, Chinese Academy of Sciences Zhong Guan Cun Hospital, China.
[††]The author is with the Department of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, China.
[†††]The author is with the Department of Electronic Engineering, Tsinghua University, Beijing, China.
[††††]The author is with Research and Education Dept., Beijing Zhong Guan Cun Hospital, Chinese Academy of Sciences Zhong Guan Cun Hospital, China.
[†††††]The authors are with the Beijing XinHeYiDian Technology Co. Ltd., Beijing, China.

*This paper is partially supported by the Special Fund for Independent Innovation and Industry Development in the Core Area in Haidian District of Beijing (713-KJC-021).

a) E-mail: whq@zgchospital.com (Corresponding author)

DOI: 10.1587/transinf.2017EDP7410

mance.

In this study, we proposed a novel ECG recognition scheme, by leveraging wavelet transform and modern machine learning techniques to detect characteristic points in ECG waves. We devised a novel wavelet transform based feature, termed as randomly selected wavelet transform (RSWT) feature. The RSWT feature could effectively represent the morphology of ECG waveform (details of the feature importance are described in Sect. 3.3). The baseline removal and noise reduction process can be combined with the features selected by random forest classifier. With the feature pool of RSWT, random forest classifier can determine the class of each ECG sampling point. The detection result, given by the trained random forest classifier was further refined to produce the final detection output.

Our evaluation results on the QT database [20] and the MIT-BIH database [21] showed comparable sensitivity and precision to other state-of-the-art works. Especially our scheme with RSWT feature showed better performance than previous machine learning methods on QT database [18]. Since our scheme is based on machine learning, further advancement can be achieved with more annotated ECG data. As for the evaluation results on the MIT-BIH database, it showed that our method has good extendibility across different ECG measuring sources and achieved promising cross-database performance. Moreover, our method can process both single lead and multi-lead ECG data, therefore it can be applied to a variety of applications.

2. Methods

2.1 ECG Characteristic Points

Electrocardiogram (ECG) is a non-invasive way to measure the heartbeat rhythms. The heart's electrical activity can be captured by the voltage variance of ECG. The ECG signal from a normal heart has typical features such as the P wave, the QRS complex, and the T wave. Figure 1 is a segment of ECG sample from the record sel103 of QT database.

Our work can estimate eight types of ECG characteristic points: onset, peak, and offset of P wave; onset, offset of QRS complex, and its fiducial mark (typically at the R-wave peak, according to annotation definition in QT database); peak, offset of T wave, and other points (T_{others}), i.e. points with types other than the previously defined 8 types. Firstly, the classifier model labels each point in ECG waveform into one of the eight types of characteristic points or the T_{others} type (indicates current point is a non-target point). Secondly, this classification result will go through post-processing to derive the final detection output. According to the detected characteristic points, we could also segment ECG waveform into P, QRS, and T wave segments.

2.2 Algorithm Framework

The framework of our ECG recognition scheme is shown in Fig. 2. Firstly, we resampled the ECG time-domain signal to 250 Hz to avoid inconsistency of sampling frequency between training and testing signals. The frequency of 250 Hz was selected for both training and testing because the QT database has a sampling frequency of 250 Hz (any other suitable sampling frequency could also be chosen). Secondly, groups of point-pairs were randomly extracted from a set of windows of length L_w (details of L_w are described

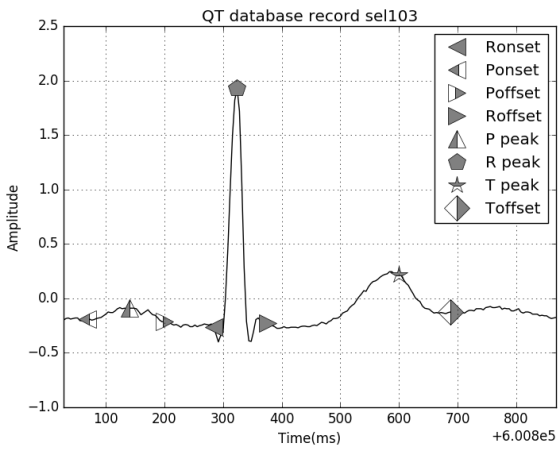

Fig. 1 Sample ECG segment of record sel103 from QT database, the sample point indexes are in range [150207, 150416], the annotations is the expert annotations in QT database.

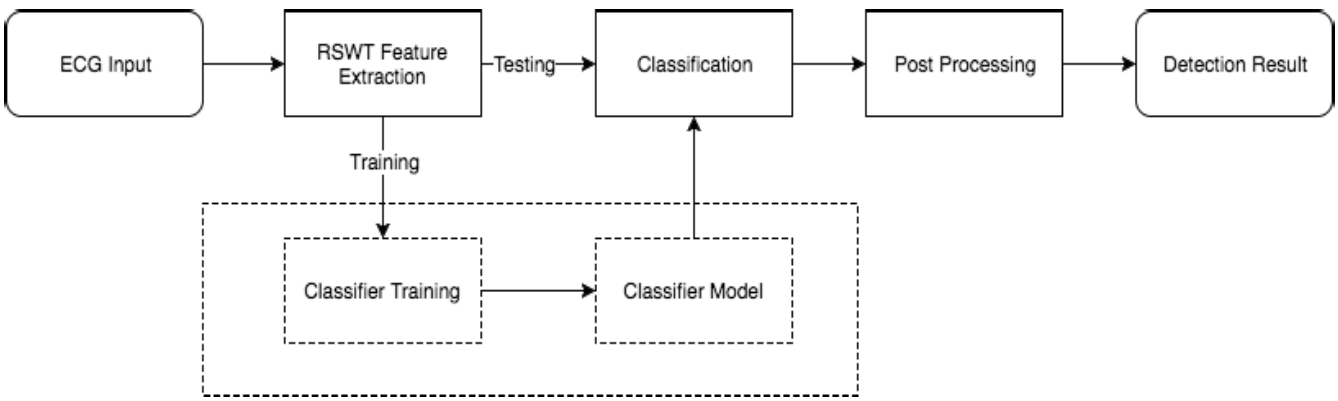

Fig. 2 The proposed framework for ECG recognition, the RSWT feature extraction algorithm is used for feature extraction in both training and testing stages.

in Sect. 2.3) at each level of wavelet coefficients. Then we computed the RSWT feature from all the levels of ECG wavelet coefficients according to the pair selection scheme (details of feature extraction are described in Sect. 2.3). Finally, in the training stage, a random forest classifier was trained to select the effective features, and the trained model was used for annotating ECG signals with RSWT feature.

In the detection stage, around the ground truth position, there were multiple positive responses with the classifier. Here, a post-processing method was adopted to group sample points with the same annotation and estimate the final position of the annotation, which is described in detail in Sect. 2.5.

2.3 RSWT Feature Extraction

The key module of our scheme was the extraction of RSWT feature. To illustrate the process of the RSWT feature extraction, Fig. 3 gives the pseudo code of the feature extraction algorithm.

The RSWT feature for each ECG sample position was extracted from a set of signal windows centered on the sample position guided by a pre-set pattern. This randomly selected pattern was fixed and unchanged throughout the entire training and testing process. Firstly, for each sample point position in the ECG time domain waveform, a virtual window with length L_w that centered on the given sample position was defined. The corresponding windows were also located on each level of WT coefficients in the corresponding position. Feature extraction procedures would only be conducted within these windows. Secondly, a pool of every combination of point-pairs within each window was generated, and pairs of time-axis positions were drawn from the coefficient pair pool, then we computed the amplitude difference and sign of the difference between the selected pairs as the feature representation of the current window, which is described in detail as follows.

Using wavelet, we can derive the stationary wavelet transform coefficients for the signals segment in the window. In this process, we chose the Daubechies2 (db2) wavelet because the Daubechies Wavelet is widely reported for the accuracy of details compared to other methods. Moreover, this wavelet shows similarity with QRS complexes and the energy spectrum is concentrated around low frequencies [22], [23]. The maximum level of the wavelet transform was set to 6, since experiment results showed that higher order mainly consisted of noise.

```
-for each sample point x_i in each ECG waveform:
    -compute the normalized signal sig_norm of ECG
        signal segment in window
        [x_i − L_w/2, x_i + L_w/2] as in (1)
    -compute WT coefficients wt_coef of sig_norm
      -for each level of coefficients in wt_coef:
        -repeat N_f/2 times:
          -select two different positions [x_1, x_2]
              in current WT coefficients
          -compute wt_coef_i(x_1) - wt_coef_i(x_2) and
              |wt_coef_i(x_1) − wt_coef_i(x_2)| as feature
              values
```

Fig. 3 Pseudo code of RSWT feature extraction.

In order to make the proposed feature more robust to ECG amplitude variations, the wavelet coefficients inside the fixed length window were normalized according to the maximum and minimum values of the time domain signal inside the window:

$$Norm(Coef_i) = \frac{Coef_i}{(S_{max} - S_{min})}, (i = 1, 2, \ldots, L_w) \quad (1)$$

Where $Coef_i$ is the i_{th} wavelet coefficient in the fixed length window, L_w is the length of the window (number of sample points in ECG signal), and S_{max} and S_{min} are the maximum and minimum values of the time domain signal in the window.

The RSWT feature contained the sign and absolute difference of WT coefficients as a feature representation of the selected point-pairs. The feature for a given pair (x_1, x_2) was represented by: $|coef(x_1) - coef(x_2)|$ and $sign(coef(x_1) - coef(x_2))$, where the function $coef(x)$ represents the x_{th} sample value of the current level of WT coefficients, and the sign function $sign(x)$ is given by:

$$sign(x) = \begin{cases} 1(x \geq 0) \\ -1(x < 0) \end{cases} \quad (2)$$

The WT coefficients in different levels contained information of ECG waves in different frequency bands. Therefore, computing point-pair features within each level of WT coefficients allowed the RSWT feature to obtain frequency information about ECG waveform. Note that ECG denoising in frequency domain was implicitly performed by the feature selection process in WT coefficients. This made the system more robust and allowed the direct application over the time-domain ECG signal without denoising.

The total number of different signal pairs was:

$$\begin{pmatrix} L_w \\ 2 \end{pmatrix} = \frac{{L_w}^2 - L_w}{2} \quad (3)$$

Where $\begin{pmatrix} L_w \\ 2 \end{pmatrix}$ denotes the number of 2-combinations in a set of L_w elements. It cannot be afforded when the window length L_w is too large. Therefore, we employed the random selection of pairs instead of storing all the possible pairs, which greatly reduced the memory cost.

The total feature number used in random forest was N_f, the value of N_f was set to 4000 in our experiments. Since a pair (x_1, x_2) could generate 2 features, therefore, the total number of $N_f/2$ pairs were randomly drawn from the wavelet coefficients.

There were two types of random pairs: For one pair, both of the ends were randomly selected, for the other pair, one end was fixed at the center position of the window and the other end was randomly selected. The number of pairs with one end fixed was $N_f/4$, half of the total number of pairs.

Random selection from the WT coefficient pairs could

decrease the number of pairs thus reducing the dimension of our feature. We reduced the dimension of the feature not only because it was computationally efficient, but also because of the ability to train the classifier more easily and efficiently on low dimensional features.

As for L_w, a higher value was set for it than the sampling frequency f_s, because the normal resting adult human heart rate ranges from 60 ∼ 100 bpm, and we wanted the fixed-size window to include more than one heartbeat cycle. We set $L_w = 3 \times f_s$ as the fixed-size window length. The detection accuracy was not sensitive to longer window sizes. The influence of L_w is further discussed in Sect. 3.4.

There are two reasons that make the proposed RSWT feature more effective. First, the point-pair difference feature is a good approximation of signal gradient and relative position between sample points. It captures the amplitude relationship between the target sample points and surrounding sample points. Second, the features in wavelet domain are easier for the classifier to decide which level of WT coefficients contains effective information for classification.

2.4 Random Forest Classification

The random forest classifier would extract effective features from RSWT feature pool. The RSWT feature was a local descriptor, which not only tried to describe the relative amplitudes between the window center with other points, but also described the relationship between neighboring points. It was calculated by the values of $|coef(x_1) - coef(x_2)|$ and $sign(coef(x_1) - coef(x_2))$, for each pair (x_1, x_2) generated in Sect. 2.3.

Since RSWT features were weak gradient-like pair local descriptors, random forest was suitable for the RSWT feature classification task [24]. The ECG characteristic point detection with random forest consisted of two processes: learning and testing. In the random forest learning process, a group of decision trees was randomly trained according to the expert annotations in the QT database. The training set contained expert annotations from 75 records of QT database, the other 30 records were used for testing.

We added negative samples (T_{others}) to the training set. The negative samples were taken from positions other than annotated positions in the ECG signal. In the negative sample selection process, in order to avoid selecting the samples that were characteristic points but without expert annotations, we manually marked the range of the expert annotations, and randomly selected negative samples within the annotated beats. The negative samples were also kept a minimum distance of 40ms (the distance between 10 points in the sampling frequency of 250 Hz) from the expert annotations to avoid high correlations between positive and negative samples.

In the random forest classifier training process, 30 decision trees were trained and the test results of the random forest was given by averaging the probabilistic prediction of the decision trees. If the total number of training samples was N, then each decision tree was trained by N samples randomly drawn with replacement from the training samples. In the decision tree training process, the best split position was determined by the best split among a random subset of the features. The size of the random subset was $\sqrt{N_f}$, where N_f is the number of features. The nodes were expanded until the maximum depth was reached or until the leaves contained less than 2 samples. The maximum depth of each decision tree D_{tree} was set to 35 to prevent overfitting.

In the testing process, the RSWT feature was first extracted from the test position in the ECG signal, then the test result was given by averaging the probabilistic prediction of decision trees from the trained model.

2.5 Result Post-Processing

The post processing included two stages: grouping and group filtering.

In the grouping process, we first formed groups by collecting nearby sample points (samples with a distance of 1) with the same prediction label, and skipped groups with only 1 sample. Then the representative sample position of a group was determined by the confidence-weighted mean of the sample positions.

In the group filtering process, the groups with same label were processed when their distance was less than a threshold calculated by the maximum human heart rate. The group with smaller size was skipped to reduce false alarms.

3. Results and Discussion

In this section, we described the experiments to evaluate the within-database and cross-database performance of our method. We will describe our overall performance on the QT database in 3.1. Furthermore, the better cross-database performance of our system is discussed in 3.2.

To evaluate the results, we adopted the same evaluation scheme as described by Martinez et al. [11]. We detected the characteristic points in both leads of the records in QT database and selected the best data to derive the final conclusion. We computed the detection accuracy, sensitivity (also referred as detection rate) $Se = TP/(TP + FN)$, and positive predicitivity $P^+ = TP/(TP + FP)$, where TP denotes the number of true positive detections (wave was present and was detected), FN represents the number of false negative (wave was present but was not detected), and FP indicates the number of false positive detections (wave was not present but was detected).

For an expert annotation, the true positive was evaluated using the standard beat-by-beat comparison procedure given by AAMI et al. [25]. The number of false negative was calculated by the number of expert annotations that couldn't find a true positive matching and the number of false positive was calculated by the number of detection results that didn't match the expert annotations, we also manually annotated the regions with incomplete expert annotations, so that when a test result was located in these regions, it was free from statistics calculation.

Table 1 Detection accuracy on the QT database

Method	Metric	P-on	P-peak	P-off	QRS-on	QRS-peak	QRS-off	T-peak	T-off	Average
RSWT	M±SD	0.9±12.3	-0.9±9.9	-4.9±10.0	-2.2±6.6	0.8±3.3	-1.8±8.1	6.0±11.6	2.7±16.4	2.5±9.8
(Our method)	Se	100.0%	99.9%	100.0%	100.0%	100.0%	100.0%	98.5%	98.1%	99.6%
	P^+	99.0%	98.5%	98.5%	100.0%	100.0%	100.0%	99.9%	99.9%	99.5%
RSSPD [19]	M±SD	0.4±22.0	N/A	2.1±12.9	0.2±10.2	N/A	0.5±14.4	N/A	1.4±17.2	0.9±15.3
(Our previous work)	Se	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	P^+	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
KNN([18])	M±SD	2.5±14.3	N/A	3.1±13.9	2.4±8.8	N/A	4.7±9.2	N/A	2.8±18.6	3.1±13.0
	Se	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	P^+	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
WT([11])	M±SD	2.0±14.8	3.6±13.2	3.5±18.0	4.6±7.7	N/A	0.8±8.7	0.2±13.9	-1.6±18.1	2.3±13.5
	Se	98.87%	98.87%	98.75%	99.97%	99.92%	99.97%	99.77%	99.77%	99.54%
	P^+	91.03%	91.03%	91.03%	N/A	99.88%	N/A	97.79%	97.79%	94.76%
Phase free SWT [13]	M±SD	-0.3±12.2	N/A	5.8±9.1	N/A	N/A	N/A	N/A	N/A	3.0±10.7
	Se	100%	N/A	100%	N/A	N/A	N/A	N/A	N/A	100%
	P^+	88.2%	N/A	88.2%	N/A	N/A	N/A	N/A	N/A	88.2%
Beat-to-beat BGS [16]	M±SD	3.4±14.2	1.1±5.3	2.7±9.8	N/A	N/A	N/A	-0.8±4.1	-3.1±14.0	2.2±9.5
	Se	99.93%	99.93%	99.93%	N/A	N/A	N/A	100%	100%	99.96%
	P^+	99.10%	99.10%	99.10%	N/A	N/A	N/A	99.30%	99.30	99.20%
Multi-beat PCGS [15]	M±SD	1.7±10.8	2.7±8.1	2.5±11.2	N/A	N/A	N/A	0.7±9.6	2.7±13.5	2.1±10.6
	Se	99.60%	99.60%	99.60%	N/A	N/A	N/A	100%	100%	99.76%
	P^+	98.04%	98.04%	98.04%	N/A	N/A	N/A	99.15%	99.15%	98.48%

†N/A indicates the statistics is not available.
††M±SD is the abbreviation of mean±standard deviation.
†††Se and P^+ are the abbreviation of sensitivity and positive predictivity.
††††We computed the average of the absolute values of the M and SD at each characteristic point for a more intuitive comparison.

3.1 Overall Detection Accuracy on QT Database

The QT database contains 105 records chosen from existing ECG database, including the MIT-BIH Arrhythmia Database, ST-T Database and several other ECG databases. At least 30 beats in each record, 3622 beats in all, were manually annotated in the database. The annotation includes the beginning, the peak, and the end of the P wave; the beginning, the peak, and the end of the QRS complex; the peak, and the end of the T wave.

In order to evaluate the performance of the proposed algorithm, we randomly selected 75 records from the QT database to train the random forest classifier. The remaining 30 records were tested to evaluate the algorithm. This process was repeated for 30 rounds. We did not detect the T-onset due to the lack of expert annotations in the QT database.

The performance of wave delineation was measured by the mean (denoted as M) and standard deviation (denoted as SD) of error values. It's noted that the SD will be more important when the M is comparable. For each expert label, the error was defined as the distance between the detection output and the corresponding expert label. The indicated time values (in ms) were based on a sampling frequency of 250 Hz.

The total average detection accuracy of 30 rounds are listed in Table 1. The test result showed that our method achieved comparable sensitivity, positive predictivity, and detection accuracy (mean and standard deviations) in ECG characteristic point detection. Note that only our RSWT delineation scheme managed to detect all types of points in a unified model simultaneously, thus simplifying the detection and delineation process.

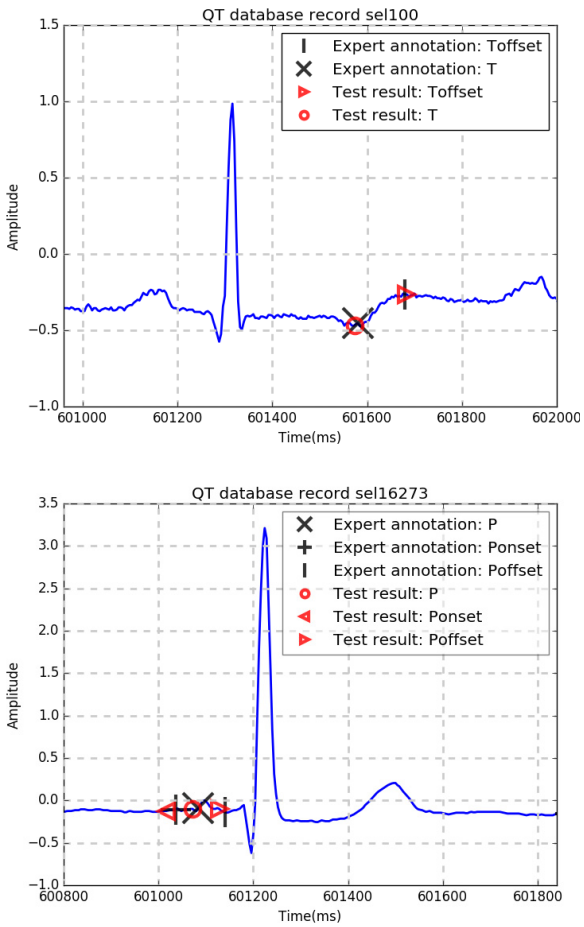

Fig. 4 Examples of test results on the QT database, the index range for sel100 and sel16273 is [150240, 150500] and [150200, 150460], respectively.

Among the machine learning based approaches (RSSPD [19] and KNN [18]), our algorithm achieved superior performance in SD at all types of characteristic point. It can also be observed that our results were comparable with state-of-the-art methods (BGS [16] and PCGS [15]), which were 2.5±9.8 compared to 2.2±9.5 and 2.1±10.6. However, BGS and PCGS methods depend on their prior knowledge and pre-defined models.

The delineation accuracy of each record in one of the rounds are presented in Table 3. As it is shown in the table, most of the records had high detection accuracy, however, some records in the testing set obtained larger mean error and standard deviation, for the lack of similar wave morphologies in the training set.

The ECG characteristic point detection result on the QT database is shown in Fig. 4. For the QT record sel100, even though the T wave is mono-phased, our algorithm correctly detected the position of T-peak and T-offset. For the QT record sel16273, we detected the P wave with relatively small amplitude.

The experiment results in this section on the QT database showed that the RSWT feature is an effective representation of ECG morphologies. This reduces the demand for a denoising procedure, since denoising often results in the loss of clinical information in the ECG waves.

3.2 QRS Detection Results on MIT-BIH Database

The MIT-BIH database includes specially selected Holter recordings with anomalous but clinically important phenomena. The annotations of QRS positions (R-peak) were used for the evaluation in this section.

In order to verify the enhanced cross-database performance, our classification model was trained on the QT database and then applied to the MIT-BIH database. Since the sampling frequency is 360 Hz for the MIT-BIH database and 250 Hz for the QT database, we down-sampled the ECG in the MIT-BIH database to 250 Hz. Table 2 shows the QRS detection sensitivity Se and positive predictivity

Table 2 QRS detection results on MIT-BIH database

	Sensitivity	P^+
This work	99.82%	99.37%
DPI ([4])	99.52%	99.70%
WT ([11])	99.80%	99.86%

Table 3 Detection accuracy with window size $3 \times f_s$ in QT database

Record Name	**Metric**	**P-on**	**P-peak**	**P-off**	**QRS-on**	**QRS-peak**	**QRS-off**	**T-peak**	**T-off**
sele0121	M±SD(ms)	8.3±7.8	10.3±5.6	4.0±6.8	0.6±7.3	1.6±2.4	-11.4±5.2	3.3±4.3	-30.7±12.7
sel307	M±SD(ms)	3.8±7.0	5.1±5.4	5.1±7.4	-2.3±5.0	2.0±1.5	-12.6±6.1	1.2±8.4	12.8±36.0
sel16773	M±SD(ms)	10.8±8.8	4.7±5.3	-7.3±10.3	-14.3±6.3	1.5±1.4	-1.9±4.3	-9.0±9.3	-6.3±10.1
sel16265	M±SD(ms)	6.3±8.3	-1.2±6.5	4.0±7.5	-16.5±6.1	-0.7±1.9	-15.0±9.2	7.0±8.9	-10.4±15.3
sel873	M±SD(ms)	15.6±11.8	-3.0±10.8	4.0±9.2	4.7±7.3	-2.0±1.8	-6.8±6.7	0.7±7.0	-3.3±11.3
sele0210	M±SD(ms)	-7.5±14.7	-6.3±7.3	-7.9±7.6	-4.8±6.1	0.6±1.2	-16.2±8.5	-1.2±6.7	0.8±9.7
sele0406	M±SD(ms)	21.3±18.0	12.0±8.0	8.8±8.8	-0.4±3.5	-0.6±2.5	-13.5±4.7	2.5±5.3	-3.8±10.1
sel808	M±SD(ms)	4.1±18.3	-0.7±7.0	-0.2±9.8	-1.4±11.7	-2.0±1.8	5.4±13.4	1.7±6.8	-0.6±18.8
sele0405	M±SD(ms)	32.2±12.7	20.7±11.5	21.4±10.1	7.2±8.9	-1.5±2.6	-5.5±8.5	3.9±8.4	0.9±17.2
sele0112	M±SD(ms)	0.3±11.5	2.4±6.9	4.3±5.8	-7.3±5.0	0.3±2.9	-11.8±7.4	-8.4±5.9	14.1±11.5
sel231	M±SD(ms)	4.2±10.4	4.2±8.5	-4.3±9.2	5.9±7.5	6.9±4.6	-2.8±19.1	67.0±63.4	40.6±44.9
sel47	M±SD(ms)	-28.6±7.8	-0.4±9.9	-4.0±7.6	-14.9±6.9	4.4±3.0	16.1±8.9	-20.4±11.1	20.5±14.9
sele0136	M±SD(ms)	6.1±11.4	3.3±7.4	-0.7±7.2	-2.7±4.4	0.3±2.5	-6.0±6.0	25.8±14.2	37.7±30.8
sel17453	M±SD(ms)	16.7±10.3	19.3±7.7	13.1±8.6	0.6±6.7	2.0±2.2	3.1±8.8	4.4±7.4	-0.5±13.2
sele0303	M±SD(ms)	0.9±8.7	10.3±8.3	4.4±6.3	-2.6±3.7	-2.1±2.8	-13.5±8.5	6.0±6.7	44.7±13.5
sel32	M±SD(ms)	15.8±9.9	15.5±13.3	3.6±6.3	0.9±3.4	1.3±2.5	4.4±6.7	8.2±18.8	21.7±27.7
sel306	M±SD(ms)	-8.2±8.3	12.5±8.5	-7.4±10.3	-2.0±5.1	0.4±3.4	-16.4±7.1	5.4±9.7	-3.9±10.0
sele0106	M±SD(ms)	2.4±7.9	7.6±9.7	0.2±11.4	-8.1±4.7	2.7±3.2	-2.8±4.2	42.6±10.7	36.1±15.4
sel16795	M±SD(ms)	20.2±11.5	16.3±7.8	10.9±8.2	1.9±7.0	1.6±3.7	-1.6±7.3	4.1±5.8	0.8±9.5
sel871	M±SD(ms)	-1.9±8.6	2.3±6.4	6.3±10.7	-14.5±7.1	2.1±3.4	-7.2±5.5	7.2±9.8	6.6±12.5
sele0129	M±SD(ms)	11.1±14.5	1.4±7.4	-6.4±8.6	-0.5±9.8	0.0±2.3	-2.6±5.0	1.0±5.5	22.3±12.2
sel16539	M±SD(ms)	17.0±10.2	-2.4±6.0	-24.1±19.2	0.9±2.9	1.4±2.3	-6.0±4.9	15.5±9.0	-0.6±9.6
sele0609	M±SD(ms)	8.5±9.6	8.1±9.3	3.8±8.3	-8.9±5.7	-0.6±2.0	-7.2±8.3	4.2±9.6	12.9±19.1
sele0612	M±SD(ms)	-72.1±15.3	-64.0±8.9	-49.2±13.0	-3.4±6.8	-3.6±2.0	0.6±7.7	15.6±14.2	12.0±13.2
sel302	M±SD(ms)	36.1±12.6	23.9±12.7	14.9±5.7	2.4±5.3	0.3±3.1	-2.5±6.0	-4.5±7.7	-4.3±11.1
sele0124	M±SD(ms)	-13.7±11.6	-9.2±8.1	-14.6±10.6	-0.5±7.2	2.0±3.9	-6.2±7.7	3.5±9.8	25.1±16.4
sele0122	M±SD(ms)	7.1±6.8	8.8±5.5	-1.8±6.4	-0.2±5.6	0.6±2.5	-7.8±6.4	-2.3±6.0	-41.3±9.0
sel16272	M±SD(ms)	53.0±16.3	45.2±10.2	28.2±7.3	6.6±6.1	0.2±2.3	-7.8±8.1	3.1±6.0	-6.6±10.6
sel114	M±SD(ms)	-5.3±13.4	-8.9±18.3	-15.7±14.5	-6.8±8.7	0.5±1.9	3.5±10.8	8.4±11.7	10.6±24.0
sele0704	M±SD(ms)	6.7±17.7	3.4±9.2	4.8±7.9	-1.6±5.3	-21.5±11.4	31.1±16.5	3.0±6.7	-9.6±9.0

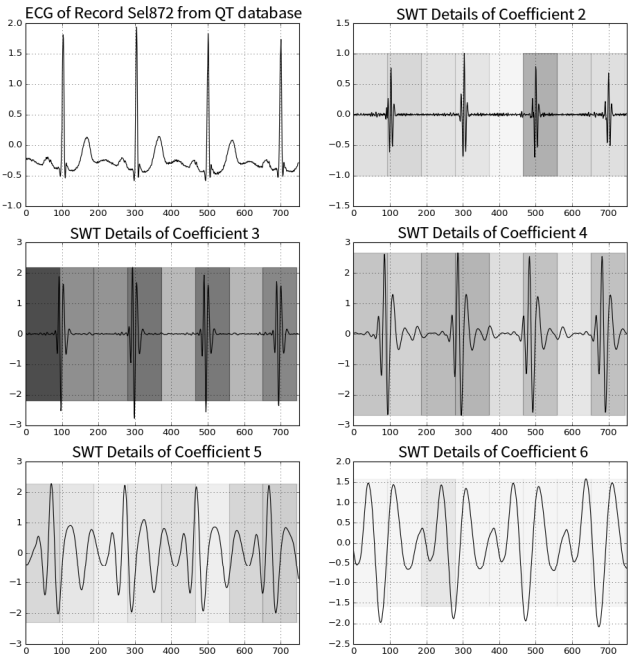

Fig. 5 Feature importance of the trained R-peak random forest classifier model. It showed that the R wave model had more important pairs in Coefficient 3.

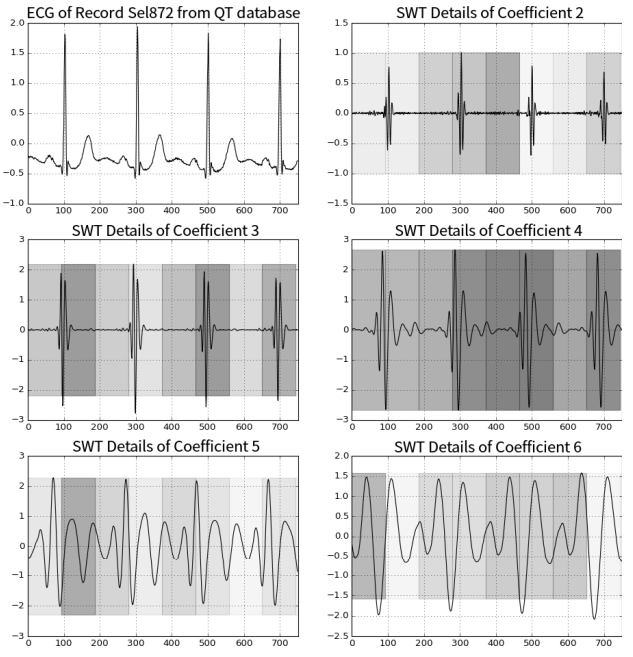

Fig. 6 Feature importance of the trained T-peak random forest classifier model. It showed that the P wave model had more important pairs in the front of the window center in Coefficient 3.

P^+ on the MIT-BIH database. The results showed that our scheme achieved higher sensitivity and comparable positive predictivity compared to other reported detection schemes. Note that both DPI [4] and WT [11] were tuned and evaluated solely on the MIT-BIH database. It demonstrated that our novel RSWT scheme has a robust cross-database performance.

3.3 RSWT Feature Importance Visualization

The feature importance in each level of Wavelet Transform (WT) coefficients for trained P-peak model, QRS fiducial mark (R-peak) model, and T-peak model are visualized in Fig. 5, Fig. 6, and Fig. 7 respectively. In order to compare the difference in the trained model of the different type of waves, we adopted the same set of pairs randomly generated in WT coefficients.

The feature importance was derived from the average of the number of samples that a feature can predict in the random forest. To visualize the importance of the selected pairs, we analysed the top 3% of the most important pairs, and evenly divided the time-axis into 8 sections. With increasing number of weighted pairs in a section the color became darker. The weight of a region Ω is given by:

$$\begin{aligned} weight(\Omega) = & \sum_{\substack{pair[0]_x \in \Omega \\ pair[0]_x \neq center_x}} Importance(pair) \\ + & \sum_{\substack{pair[1]_x \in \Omega \\ pair[1]_x \neq center_x}} Importance(pair) \end{aligned} \tag{4}$$

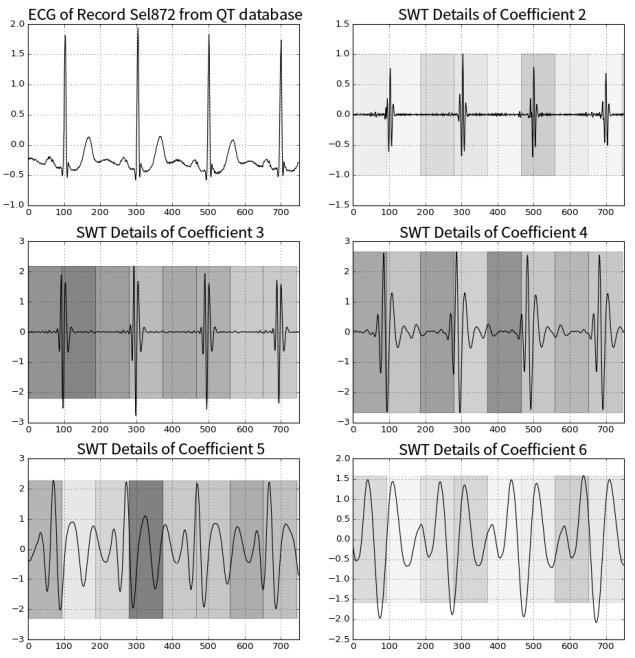

Fig. 7 Feature importance of the trained P-peak random forest classifier model. It showed that the T wave model had more important pairs in the back of the window center in Coefficient 3.

Where $pair[0]_x$ is the time index of the first point in the pair and $pair[1]_x$ is the time index of the second point in the pair; $center_x$ is the time index of the window center and $Importance(pair)$ is the feature importance of the pair. We didn't visualize the importance of the window center index because it was manually selected.

The most significant difference of the feature impor-

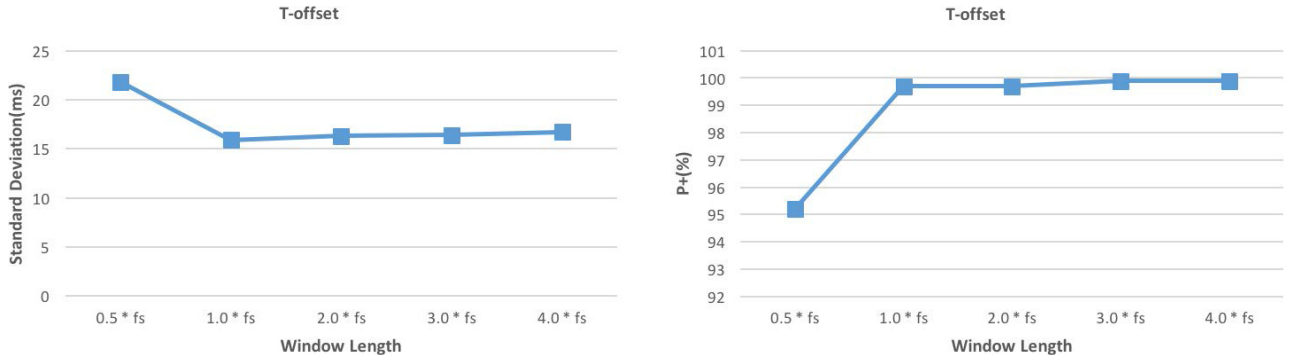

Fig. 8 Standard deviation(ms) and detection P^+ for different window lengths, L1~L4 represents window length of $1 \times f_s$ to $4 \times f_s$, respectively.

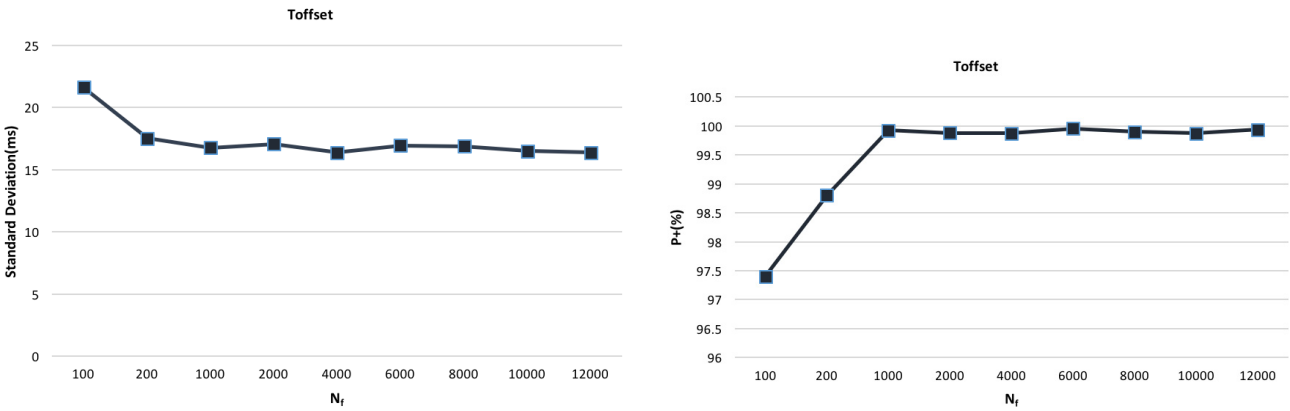

Fig. 9 Toffset detection standard deviation(ms) and detection P^+ for different N_f values, where N_f is the number of features drawn from the windows.

tance of these models lied in WT coefficient level 3 and 4. Comparing Fig. 5 with Fig. 6 and Fig. 7, the R wave model had more important pairs in WT level 3, since R wave usually has higher frequency than P and T wave. There were also differences among the important features of P and T wave. In wavelet transform level 3, the P wave model had more important pairs in the front of the window center, while the T wave model had more important pairs in the back of the center. This is because the P wave is usually in front of the QRS complex, having more pairs in the front can exploit this property. The similar explanation was applied to T wave also.

The difference between the important features of the P, QRS, and T wave models proved that the proposed scheme effectively captured the features of the ECG signal. The visualization result indicated that the first level of WT coefficient had less important features, since this level was often contaminated with high-frequency noises. This result also proved the effectiveness of the feature selection process.

The detection results on the QT database and the MIT-BIH database showed that the proposed method achieved superior performance to other studies, indicating that the proposed method is an effective method for ECG characteristic point detection.

3.4 Parameter Sensitivity Analysis

The delineation accuracy of our scheme is not sensitive to variations in parameter values. We evaluated the influence of window length L_w and number of features N_f on QT database with other parameters kept the same as in Sect. 3.1.

In order to evaluate the influence of window length L_w in RSWT feature extraction, we tested on different values of L_w for Toffset, as is shown in Fig. 8. According to the experiment results, the detection accuracy was not sensitive to window length, and we chose the window length of $3 \times f_s$ for a relatively better result.

The detection accuracy for Toffset at different values of N_f is also presented in Fig. 9. The detection accuracy was not sensitive to the value of N_f, and we chose $N_f = 4000$ in our work. The similar relation could be found for other characteristic points.

4. Conclusion

This work demonstrated a novel RSWT feature and a new detection scheme for ECG characteristic points. The RSWT feature exhibited a reliable feature to represent ECG morphologies. With the effectiveness of RSWT feature and the

power of machine learning, this novel scheme detected ECG characteristic points with comparable sensitivity, positive predictivity, and detection accuracy. This novel method can be further improved in speed with parallel computing since the feature extraction stage for ECG sample points are independent of each other, thus making this algorithm applicable for real-time applications.

References

[1] A. Gacek and W. Pedrycz, ECG signal processing, classification and interpretation: a comprehensive framework of computational intelligence, Springer Science & Business Media, 2011.
[2] B.U. Kohler, C. Hennig, and R. Orglmeister, "The principles of software QRS detection," IEEE Eng. Med. Biol. Mag., vol.21, no.1, pp.42–57, 2002.
[3] S. Pal and M. Mitra, "Empirical mode decomposition based ECG enhancement and QRS detection," Computers in biology and medicine, vol.42, no.1, pp.83–92, 2012.
[4] A.G. Ramakrishnan, A.P. Prathosh, and T.V. Ananthapadmanabha, "Threshold-independent QRS detection using the dynamic plosion index," IEEE Signal Process. Lett., vol.21, no.5, pp.554–558, 2014.
[5] P. Laguna, N.V. Thakor, P. Caminal, R. Jané, H.R. Yoon, A. Bayés de Luna, V. Marti, and J. Guindo, "New algorithm for QT interval analysis in 24-hour holter ECG: performance and applications," Medical and Biological Engineering and Computing, vol.28, no.1, pp.67–73, 1990.
[6] P. Laguna, R. Jané, and P. Caminal, "Automatic detection of wave boundaries in multilead ECG signals: validation with the cse database," Computers and biomedical research, vol.27, no.1, pp.45–60, 1994.
[7] R.V. Andreão, B. Dorizzi, and J. Boudy, "ECG signal analysis through hidden markov models," IEEE Trans. Biomed. Eng., vol.53, no.8, pp.1541–1549, 2006.
[8] F.G. Guilak and J. McNames, "A spline framework for ECG analysis," 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp.957–960, IEEE, 2011.
[9] F.G. Guilak and J. McNames, "A bayesian-optimized spline representation of the electrocardiogram," Physiological measurement, vol.34, no.11, p.1467, 2013.
[10] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," IEEE Trans. Biomed. Eng., vol.42, no.1, pp.21–28, 1995.
[11] J.P. Martinez, R. Almeida, S. Olmos, A.P. Rocha, and P. Laguna, "A wavelet-based ECG delineator: evaluation on standard databases," IEEE Trans. Biomed. Eng., vol.51, no.4, pp.570–581, 2004.
[12] P.C. Chen, S. Lee, and C.D. Kuo, "Delineation of T-wave in ECG by wavelet transform using multiscale differential operator," IEEE Trans. Biomed. Eng., vol.53, no.7, pp.1429–1433, 2006.
[13] G. Lenis, N. Pilia, T. Oesterlein, A. Luik, C. Schmitt, and O. Dössel, "P wave detection and delineation in the ECG based on the phase free stationary wavelet transform and using intracardiac atrial electrograms as reference," Biomedical Engineering/Biomedizinische Technik, vol.61, no.1, pp.37–56, 2016.
[14] J. Dumont, A.I. Hernandez, and G. Carrault, "Improving ECG beats delineation with an evolutionary optimization process," IEEE Trans. Biomed. Eng., vol.57, no.3, pp.607–615, 2010.
[15] C. Lin, C. Mailhes, and J.Y. Tourneret, "P- and T-wave delineation in ECG signals using a bayesian approach and a partially collapsed gibbs sampler," IEEE Trans. Biomed. Eng., vol.57, no.12, pp.2840–2849, 2010.
[16] C. Lin, G. Kail, A. Giremus, C. Mailhes, J.Y. Tourneret, and F. Hlawatsch, "Sequential beat-to-beat P and T wave delineation and waveform estimation in ECG signals: Block gibbs sampler and marginalized particle filter," Signal Process., vol.104, pp.174–187, 2014.
[17] I. Saini, D. Singh, and A. Khosla, "Delineation of ECG wave components using k-nearest neighbor (KNN) algorithm: ECG wave delineation using KNN," 2013 Tenth International Conference on Information Technology: New Generations (ITNG), pp.712–717, IEEE, 2013.
[18] I. Saini, D. Singh, and A. Khosla, "K-nearest neighbour-based algorithm for P- and T-waves detection and delineation," J. Medical Engineering & Technology, vol.38, no.3, pp.115–124, 2014.
[19] P. Gao, J. Zhao, G. Wang, and H. Guo, "Real time ECG characteristic point detection with random selected signal pair difference (rsspd) feature and random forest classifier," 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, EMBC, Aug. 2016.
[20] P. Laguna, R.G. Mark, A. Goldberg, and G.B. Moody, "A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG," Computers in Cardiology 1997, pp.673–676, IEEE, 1997.
[21] G.B. Moody and R.G. Mark, "The mit-bih arrhythmia database on cd-rom and software for use with it," Computers in Cardiology 1990, Proceedings., pp.185–188, IEEE, 1990.
[22] R. Haddadi, E. Abdelmounim, M.E. Hanine, and A. Belaguid, "Discrete wavelet transform based algorithm for recognition of QRS complexes," 2014 International Conference on Multimedia Computing and Systems (ICMCS), pp.375–379, April 2014.
[23] S.Z. Mahmoodabadi, A. Ahmadian, and M.D. Abolhasani, "ECG feature extraction using daubechies wavelets," Proc. Fifth IASTED International Conference, vol.2, no.2, pp.343–348, 2005.
[24] F. Moosmann, B. Triggs, and F. Jurie, "Fast discriminative visual codebooks using randomized clustering forests," Twentieth Annual Conference on Neural Information Processing Systems (NIPS'06), pp.985–992, 2006.
[25] Association for the Advancement of Medical Instrumentation, "Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms," ANSI/AAMI/ISO EC57:1998/(R)2008, pp.1–36, 2008.

Dapeng Fu is 45 years old. Associate chief physician, Chinese Academy of Sciences Zhongguancun Hospital, Director of the Beijing Medical Association, Member of the Institute of health management of the Beijing Medical Association.

Zhourui Xia is a senior undergraduate and will receive his B.S. degree in electronic science and technology from Beijing University of Posts and Telecommunications, Beijing, China, in 2018. His current research interests include medical signal processing and machine learning.

Pengfei Gao received the B.S. degree in electronic engineering from Beijing Jiaotong University, Beijing, China, in 2014. He received his M.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2017. His current research interests include information retrieval and natural language processing.

Haiqing Wang received the M.M. degree in General Practice and B.M. degree in medicine from Capital Medical University, Beijing, China. She is an Associate chief physician and director of Research and Education Department of Beijing Zhong Guan Cun Hospital (CHINESE ACADEMY OF SCIENCES ZHONG GUAN CUN HOSPITAL). Her current research interests include health education and health promotion, long-time ECG detection platform construction and clinical application. Dr. Wang is a member of Chinese Medical Association.

Jianping Lin received the B.S. degree in Electronic and Information Engineering from the University of Science and Technology of China, Hefei, China, in 2011, and the M.S. degree in Electronic and Communication Engineering from Tsinghua University, Beijing, China, in 2014. His current research interests include medical artificial intelligence and deep learning.

Li Sun received the B.S. degree in biomedical engineering from Harbin Engineering University, Harbin, Heilongjiang, China, in 2010. He received his M.S. degree in biomedical engineering form Shandong University, Jinan, China, in 2013. His current research interests include Bluetooth low energy wireless networking, weak signal detecting and digital signal processing.