

Research on Algorithm for Detecting Atrial Fibrillation Using RR Intervals

Yue Zhang

Laboratory of Embedded Systems and Technology
Graduate School at Shenzhen, Tsinghua University
Shenzhen, Guangdong Province, China
zhangyue@mail.tsinghua.edu.cn

Lieshuai Zhang

Laboratory of Embedded Systems and Technology
Graduate School at Shenzhen, Tsinghua University
Shenzhen, Guangdong Province, China
zhangl15@mails.tsinghua.edu.cn

Abstract—Atrial fibrillation is a kind of arrhythmias that is common but fatal. It causes many kinds of mortality such as stroke, transient ischemia attack and heart failure. In addition, as the age increases, the incidence rate of AF increases substantially. So it is very important to detect AF as early as possible. In this paper, an algorithm for detecting AF is proposed. The algorithm uses the feature of RR intervals and mainly includes two processes. Firstly, we pre-process the raw RR intervals and implement symbolic dynamic transformation and multiscale Shannon entropy to measure the disorder of the RR intervals to get a preliminary detection result. Then the delta RR interval distribution difference curve is used to modify the boundary between AF and normal beats to get a more stable result. Experiments show that in similar algorithms, our method achieves the best performance (sensitivity of 98.81% and specificity of 96.53% respectively). This method is suitable for ECG comprehensive analysis (detecting AF after collecting long ECG signals) in clinical use, which can help improve doctors' work efficiency. At the end of the article, we also discussed where the algorithm can be improved.

Keywords—AF detection; RR intervals; boundary modification

I. INTRODUCTION

With the development of economy and society, people are enjoying a continuous improvement of living standard. Meanwhile, people begin to pay more and more attention to health problems. In all kinds of diseases, the fatality rate of cardiovascular disease (CVD) ranks first. According to the research and statistics [1], in 2013 more than 2200 American people died of CVD each day, making an average of 3 death every 2 minutes.

Of all the cardiovascular diseases, the most common sustained arrhythmia is atrial fibrillation (AF). It is estimated that in the last 20 years patients with AF have raised by 66%. Now there are more than 2.2 million Americans having persistent or paroxysmal AF. Atrial fibrillation is a typical kind of malignant arrhythmias. Many kinds of mortality such as stroke, transient ischemia attack and heart failure are all associated with AF [2]. Based on the statistics, the occurrence of AF increases as high as about 9% with age.

Because of the danger of AF, it is extremely important to diagnose the AF as early as possible. Up to now, Electrocardiograph (ECG) is still the most effective means to detect arrhythmias. Fig. 1 shows a normal cardiac cycle of

ECG. When the electrical activity in the atria propagates in disorder, AF occurs. There are usually two features of AF: irregular RR intervals and P-waves devolving into f-waves. On the basis of the two features, many algorithms have been developed for detecting atrial fibrillation. But usually the f-waves are in very low amplitude and the detection of the absence of P waves is easily affected by signal noise. So algorithms for detecting AF based on P waves [3-5] have not achieved very good results in comparison with methods based on RR intervals [6-8].

This paper proposes a method to detect atrial fibrillation using RR intervals. This method mainly includes two processes. Firstly, we use median filtering, symbolic dynamic transformation and multiscale Shannon entropy (MSSE) to measure the disorder of the RR intervals to get a preliminary detection result. Then we use the delta RR interval distribution difference curve (dRRDC) [9] to modify the boundary between AF and normal beats to get a more stable result.

The structure of the article is as follows: Part 2 describes our method of detecting atrial fibrillation, including the detailed steps. Part 3 displays the performance of the method and puts forward some analysis. In the final, part 4 summarizes the article.

II. METHOD

Usually, the first step of ECG analysis is the detection of R peaks to segment heartbeats. In this study, we mainly focus on the process of AF detection and our test signals are from standard ECG database, so we simply use the position of R peaks annotated in the database here.

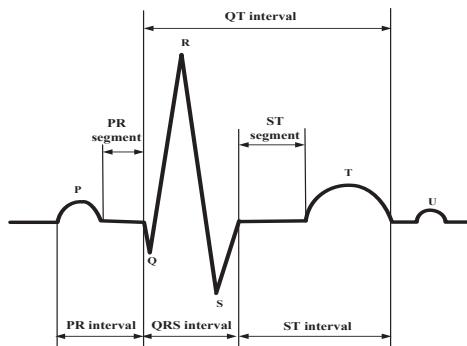


Fig. 1. A normal cardiac cycle of ECG.

This paper is supported by the Natural Science Foundation of China (No.61571268) and Shenzhen STP (JCYJ20150331151358150).

Our detection method includes two main processes. Next, we introduce the processes in detail.

A. Preliminary Detection

As mentioned above, when AF occurs, a very important feature is irregular RR intervals. Relative to similar algorithms [8, 10], we use unequally divided symbolic dynamic transformation and multiscale Shannon entropy here to measure the disorder of RR intervals. We use the following steps to get a preliminary detection result.

Step 1: signal pre-process: We use a 2 points median filter to process the original RR interval sequence. Then we add the raw RR interval sequence and the filtered RR interval sequence one by one to get the new sequence RR_n . Using the filter, we can suppress unwanted outliers on one hand. On the other hand, we find in the experiment that for different individuals, the degrees of disorder in RR intervals are not same. Through this transformation, we can suppress the individual difference. The sequence $\text{delta}RR_n$ is the difference between the successive RR_n .

Step 2: Symbolic Dynamic Transformation (SDT): This step we use (1) to transform the RR interval sequence to symbols. The purpose of this step is to map RR intervals to several bins, which can help facilitate statistical calculation. The bigger of the value of symbols, the more abnormal of the RR intervals. It should be noted that the length of each bin is not same. We determine each dividing threshold based on the statistics information of RR intervals. Experiments show that this segmentation method can improve the performance of the algorithm.

$$sRR_n = \begin{cases} 0 & \text{if } \text{abs}(\text{delta}RR_n / RR_n) < 1/24 \\ 1 & \text{elseif } \text{abs}(\text{delta}RR_n / RR_n) < 1/16 \\ 2 & \text{elseif } \text{abs}(\text{delta}RR_n / RR_n) < 1/12 \\ 3 & \text{elseif } \text{abs}(\text{delta}RR_n / RR_n) < 1/8 \\ 4 & \text{elseif } \text{abs}(\text{delta}RR_n / RR_n) < 1/4 \\ 5 & \text{else} \end{cases} \quad (1)$$

After getting the symbol sequence, we use the word sequence transformation to examine entropic properties. The transformation is defined as (2).

$$\begin{aligned} wRR_n = & (sRR_{n-3})^4 + (sRR_{n-2})^3 \\ & + (sRR_{n-1})^2 + (sRR_n) \end{aligned} \quad (2)$$

The implementing of the word sequence transformation can help improve the property of Shannon entropy during AF. During normal sinus, the symbols are most zeros, which can not be affected by the word sequence transformation. While during AF, It is obviously that this transformation helps improve the number of different elements, which would raise the value of Shannon Entropy.

Step 3: Shannon Entropy in Sliding Window: Shannon entropy (SE) is an index used to represent the complexity of variables. It is defined as (3).

$$SE_n^1 = -\sum_{i=1}^l p_i \log_2 p_i \quad (3)$$

Here the length of the sliding window is defined as 96 beats, which means the SE_n is decided by the word sequence from wRR_{n-95} to wRR_n . From wRR_{n-95} to wRR_n , we count the

number of different elements as l and number of each element as a_i , so the probability p_i in (3) is calculated as (4).

$$p_i = a_i / 96 \quad (4)$$

The normalized shannon is defined as follows [10].

$$SE_n^1 = -\frac{l}{96 \log_2 96} \sum_{i=1}^l p_i \log_2 p_i \quad (5)$$

Step 4: Multiscale Shannon Entropy: Costa [11] implemented the multiscale entropy method in the analysis of biological signals and proved its effectiveness. The multiscale transformation is calculated as (6).

$$y_j^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, 1 \leq j \leq N/\tau \quad (6)$$

In the equation, x_i represents the original discrete time series and τ is the scale factor. When $\tau=1$, the series y is exactly the original x .

After getting the new series from the original RR intervals, we can use step 2-3 to get new Shannon entropy. In this study, we choose the length of sliding windows as 96 heartbeats, and the scale factor as 2,3and 4, which can all exactly divide 96 beats. Here, for the convenience of calculation, in different scales, we all use 96 as the denominator. In this case, the Shannon entropy is not standard, but it can still express the distribution characteristic.

Step 5: Detection Results: After getting the Shannon entropy in different scales, we sum them as the final determination indicator. The superscript number represents different scale factor.

$$SE_n = SE_n^1 + SE_n^2 + SE_n^3 + SE_n^4 \quad (7)$$

A threshold is set here and when SE_n exceeds the threshold, the beat is detected as AF. To get the best threshold, we set various threshold values in the experiment and at last, we find the threshold of 0.75 provides the best performance. Fig.2 shows the receiver operating characteristic (ROC) curve in different threshold values.

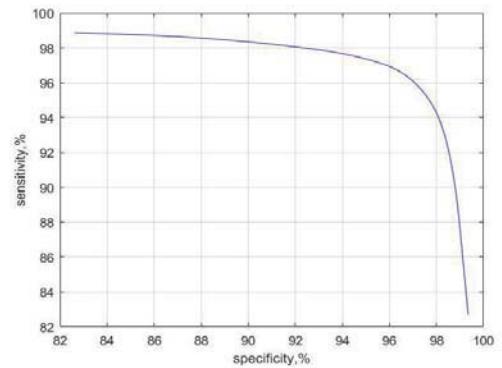


Fig. 2. ROC curve in different threshold values.

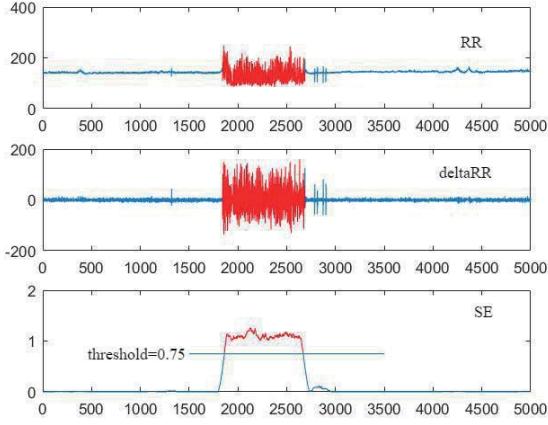


Fig. 3. An example of the data processing.

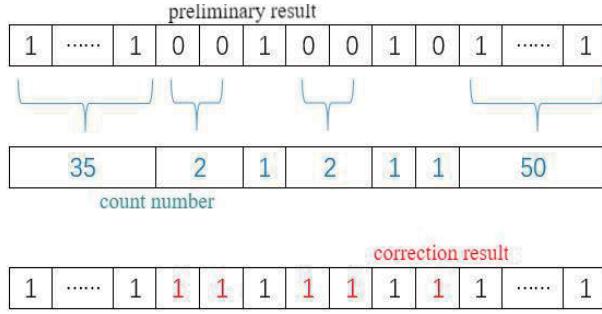


Fig. 4. An example of the correction of isolated points.

Fig. 3 shows an example of the data processing. Pictures from top to bottom are respectively raw RR intervals RR_n , δRR_n and the SE_n in sliding windows. The red part in the middle is the segment of AF in the signal. Obviously, after the transformation, it is easy to classify AF and non-AF in the last picture.

B. Boundary Modification Using dRDDC

AF is a typical kind of rhythm arrhythmia, which means that when AF occurs, it will last a period of several beats at least. In the preliminary detection, although we have used information of successive 96 heartbeats to give the detection result of one beat, we have ignored the boundary information between AF and normal. Chao [9] proposed an algorithm to detect the transition between AF and normal beats. But the algorithm needs to choose effective peaks from a lot of invalid peaks in dRDDC through complex processes and then implements Kolmogorov-Smirnov (K-S) test to detect AF.

In this study, due to the preliminary detection result, we can easily use dRDDC to modify the boundary between AF and sinus rhythm. This process mainly involves two steps.

Step 1: Correction of Isolated Points: After the first process, we have gotten a preliminary result of AF detection. But in the result, there are very short episodes of AF or sinus rhythm. Especially at some actual boundaries, AF and normal beat would appear alternately. So, in this step, we set a window of 25 heartbeats long empirically. We count the number of successive AF or normal beats in the preliminary result. For the durations less than 25 heartbeats, we correct the result label as its previous durations more than 25 heartbeats. Fig.4 shows an example of this step. In the figure,

number 1 and 0 represent AF and normal beat respectively. The symbol “.....” represents several numbers same as the previous number. This step is mainly to make the preliminary result having an initial boundary, preparing for the subsequent boundary modification.

Step 2: Boundary Modification: The density histogram of RR intervals is defined as (8).

$$H(i) = \sum_{j=1}^N sgn(floor(\frac{(\delta RR_j - \delta RR_{min}) \times M}{\delta RR_{max} - \delta RR_{min}})) = i-1 \quad (8)$$

In the equation, N represents the length of series, M represents the width of the histogram. δRR_{max} and δRR_{min} represent the range of δRR interval. The symbol “*floor*” represents rounded down. The symbol “*sgn*” is a symbolic function defined as (9).

$$sgn(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \quad (9)$$

Here, $N=48$, $M=21$ and the range is -1200ms to 1200ms. Fig.5 shows the comparison of AF and sinus histogram.

For one beat, we calculate the density histogram of 48 beats before and after the beat, defined as a_j and b_j . The difference $Diff$ is defined [9] as (10). In the durations of AF or normal sinus, the histograms are relatively stable. While at the transition between AF and normal beats, the value of $Diff$ becomes much bigger.

$$Diff = \sum_{j=0}^{20} (b_j - a_j)^2 \quad (10)$$

After getting the dRDDC, we implement a simple method to detect peaks. We set a threshold of 200 and a window of 50 beats length, which means that we only detect the peaks higher than 200 and once we detect a peak, we would not detect another peak in 50 heartbeats. This detection is very simple, so we can not assure all the detected peaks are valid boundaries, but it's enough in this study.

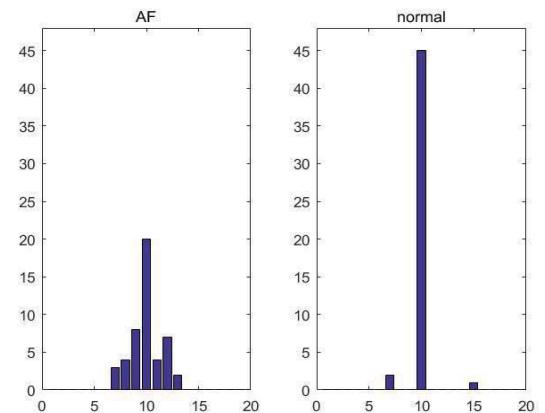


Fig. 5. Comparison of AF and normal sinus histogram.

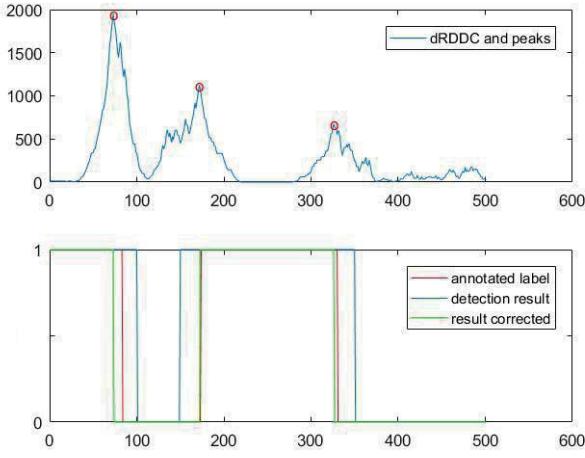


Fig. 6. Performance of the boundary modification.

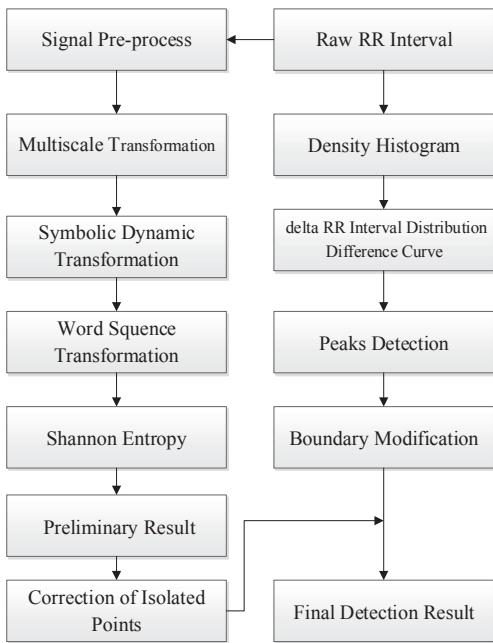


Fig. 7. Architecture of the overall process.

At last, we use the peaks to modify the boundaries. To be specific, we first find the primary boundaries in the detection result and then correct the boundaries using the peaks nearest the boundaries. Fig.6 shows the performance of this step. We can see that after the modification by dRDDC, the boundary becomes much closer to the real boundary (annotated label). The architecture of the overall process of this method is shown in Fig. 7.

III. RESULTS ANALYSIS

The proposed algorithm is tested using the AF database first developed in [12]. The database contains 25 long-term ECG recordings. Each recording has signals of two leads sampled in 250Hz, lasting 10 hours in duration. Because the raw data of the first two recordings are not available now, we use the rest 23 recordings. In the database, record 07162 and 07859 are persistent AF. Others are paroxysmal AF. We use the normal beats (annotated “N”) and AF beats (annotated “AFIB”) as test data. After statistics, the number of all beats

is 985358, in which 430135 beats are AF and the rest are normal beats.

In order to evaluate the performance of the proposed algorithm, three test indexes are chosen here. They are accurate, sensitivity and specificity rate, which can be calculated from true positives (TP), true negatives (TN), false positives (FP), and false negative (FN) as follows:

$$Se = \frac{TP}{TP+FN} \times 100\%. \quad (11)$$

$$Sp = \frac{TN}{TN+FP} \times 100\%. \quad (12)$$

$$Ac = \frac{TP+TN}{TP+FN+TN+FP} \times 100\%. \quad (13)$$

In terms of meaning, Se represents the percentage of correct detection of AF. Similarly Sp represents the percentage of correct detection of non-AF.

The results of this algorithm and some comparative algorithms are shown in TABLE I.

We can see from the table that in the preliminary detection, the sensitivity is 96.95% and the specificity is 97.12%. After modifying the boundaries using sRDDC, the sensitivity improves to 98.81% and the specificity declines slightly to 96.53%. The overall accuracy improves from 97.05% to 97.53%. We also list some results of similar algorithms for detecting AF. From the comparison, we can know that our method has the highest sensitivity and the specificity is relatively high as well. Meanwhile, our length of the sliding window is 96 heartbeats, while X.Zhou [10] uses 128 beats length. Smaller length of the sliding window means less detection delay.

In clinical use, we think that under the premise of similar accuracy, the sensitivity rate is more important. In that case, the algorithm can avoid missed detection to the greatest extent. In comprehensive analysis, if the sensitivity of detection algorithm is high enough, doctors would only need to review the ECG fragments which are detected as AF to give an authoritative diagnosis.

Through further analysis of the detection result, we find that in the records existing many episodes of short paroxysmal AF, like record 04936 and 08219, our method performs not so

TABLE I. COMPARISON OF AF DETECTION ALGORITHMS

Algorithm	Key Methods	Se (%)	Sp (%)
This paper	SDT+MSSE	96.95	97.12
	+sRDDC	98.81	96.53
S.dash [8]	SE+RMSSD	94.4	95.1
C.Huang[9]	sRDDC+K-S	96.1	98.1
X.Zhou[10]	Interger filter+SE	96.82	98.06
K.Tateno[13]	K-S+histogram	94.4	97.2
J.Ródenas [14]	Wavelet SE	96.47	94.19

well. We think this is mainly because in the sliding window detection, the information of short episodes may be ignored. But if we reduce the length of the sliding windows, the overall accuracy would be reduced too.

IV. CONCLUSION

In this paper, we have proposed an effective AF detecting algorithm using the feature of RR intervals. We first use symbolic dynamic transformation and multiscale Shannon entropy to get a preliminary result and then implement the delta RR interval distribution difference curve to modify the boundary between AF and normal sinus rhythm. Our method achieves high performance in the validation database, which we think is very suitable for ECG comprehensive analysis in clinical use.

As stated above, in the future, we need further refine our method in the detection of short AF episodes.

V. ACKNOWLEDGMENT

This paper is supported by the Natural Science Foundation of China (No.61571268) and Shenzhen STP (JCYJ20150331151358150).

REFERENCES

- [1] W. G. Members, D. Mozaffarian, E. J. Benjamin, A. S. Go, D. K. Arnett, M. J. Blaha, et al., "Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association," *Circulation*, vol. 133, p. e38, 2016.
- [2] S. Babaeizadeh, R. E. Gregg, E. D. Helfenbein, J. M. Lindauer, and S. H. Zhou, "Improvements in atrial fibrillation detection for real-time monitoring," *Journal of Electrocardiology*, vol. 42, p. 522, 2009.
- [3] M. Budeus, M. Hennersdorf, C. Perings, and B. E. Strauer, "Detection of atrial late potentials with P wave signal averaged electrocardiogram among patients with paroxysmal atrial fibrillation," *Zeitschrift Für Kardiologie*, vol. 92, p. 362, 2003.
- [4] D. Michalkiewicz, M. Dziuk, G. Kamiński, R. Olszewski, M. Cholewa, A. Cwetsch, et al., "Detection of patients at risk for paroxysmal atrial fibrillation (PAF) by signal averaged P wave, standard ECG and echocardiography," *Polski Merkuriusz Lekarski Organ Polskiego Towarzystwa Lekarskiego*, vol. 20, p. 69, 2006.
- [5] S. Ladavich and B. Ghoraani, "Rate-independent detection of atrial fibrillation by statistical modeling of atrial activity," *Biomedical Signal Processing & Control*, vol. 18, pp. 274-281, 2015.
- [6] J. Lian, L. Wang, and D. Muessig, "A simple method to detect atrial fibrillation using RR intervals," *American Journal of Cardiology*, vol. 107, p. 1494, 2011.
- [7] O. Maitas, J. Lee, D. Robotis, P. Bourell, K. H. Chon, and D. D. Mcmanus, "Detection of Atrial Fibrillation Using an iPhone 4S," *Circulation*, vol. 126, p. A9298, 2012.
- [8] S. Dash, K. H. Chon, S. Lu, and E. A. Raeder, "Automatic real time detection of atrial fibrillation," *Annals of Biomedical Engineering*, vol. 37, pp. 1701-1709, 2009.
- [9] C. Huang, S. Ye, H. Chen, D. Li, F. He, and Y. Tu, "A novel method for detection of the transition between atrial fibrillation and sinus rhythm," *IEEE Trans Biomed Eng*, vol. 58, pp. 1113-1119, 2011.
- [10] X. Zhou, H. Ding, B. Ung, E. Pickwellmacpherson, and Y. Zhang, "Automatic online detection of atrial fibrillation based on symbolic dynamics and Shannon entropy," *BioMedical Engineering OnLine*, vol. 13, pp. 1-18, 2014.
- [11] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale entropy analysis of biological signals," *Physical Review E Statistical Nonlinear & Soft Matter Physics*, vol. 71, p. 021906, 2005.
- [12] G. B. Moody and R. R. Mark, "NEW METHOD FOR DETECTING ATRIAL FIBRILLATION USING R-R INTERVALS," *Computers in Cardiology*, pp. 227-230, 1983.
- [13] K. Tateno and L. Glass, "Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and deltaRR intervals," *Medical & Biological Engineering & Computing*, vol. 39, pp. 664-671, 2001.
- [14] J. Ródenas, M. García, R. Alcaraz, and J. Rieta, "Wavelet Entropy Automatically Detects Episodes of Atrial Fibrillation from Single-Lead Electrocardiograms," *Entropy*, vol. 17, pp. 6179-6199, 2015.