Quest Journals
Journal of Electronics and Communication Engineering Research
ISSN:2321-5941 Volume 2 ~ Issue 1 (2014) pp: 01-06
www.questjournals.org

Research Paper

A novel approach to detect Atrial Fibrillation efficiently and accurately from 48 hours of ECG data

Jing Niu, Victoria Gokana, Shirley Goh,
Gaelle Lissorgues, Chee Teck Phua*

Received 03 January, 2014; **A**ccepted 10 January, 2014 **P**ublished with open access at **www.questjournals.org**

ABSTRACT:- *An algorithm for automatic detection of Atrial Fibrillation (AF) using R peak to R peak (RR) interval from electrocardiograph (ECG) signals has been developed and evaluated. The algorithm consists of a preprocessing which is based on Pan-Tompkins method to extract reliable QRS complex thus RR interval, followed by a 3 steps approach to detect AF using RR interval in ECG. The first step computes the Root Mean Square of the Successive Differences of the RR intervals extracted from a 48 hours ECG recording to identify the presence of arrhythmia. The second step defines the precise starting and stopping time of an arrhythmia episode through applying autocorrelation on the squared ECG signal within an arrhythmia window detected by the first step. The last step discerns AF from other type of arrhythmias through computing the Shannon Entropy (ShE) for the entire range of arrhythmia episode defined in step 2. The evaluation of the proposed algorithm was carried out utilizing normal sinus rhythm, AF and the arrhythmia ECGs provided in MIT-BIH database. The result demonstrates a 100% reliability in the detection of AF with accuracy of 99.5% in the identification of the start and stop time of an AF episode. In addition, the relationship between ShE threshold value of AF and duration of AF episode is studied with a formula proposed based on the results. The total processing time for the 48 hours ECG data is approximately 8 minutes, demonstrating the high efficiency of this algorithm.*

Keywords:- *Atrial Fibrillation, Electrocardiogram, Tom-Pankins, Autocorrelation, Shanon Entropy*

I. INTRODUCTION

Atrial Fibrillation (AF) is the most common arrhythmia, particularly in the elderly and those with heart disease.[1,2] It significantly increases the risk of cognitive dysfunction, thrombo-embolism and heart failure leading to mortality and morbidity [3-8]. Approximately 6 million Europeans and approximately 3 million Americans are affected by AF [1, 9]. The risk of developing AF increases with age, at age of 55 the risk is approximately 23% [1]. As the populations aging, it is highly possible that AF will become a public health burden [10]. Therefore, timely and accurate diagnosis of AF is crucial.

In an AF episode, an uncoordinated electrical and physical activity in the atria of the heart occurs. This will lead to disrupted electrical pathways thus quivery of the atria and irregular heart-beats with significant beat-to-beat variability and complexity. Such abnormally in heart activities can be detected through analysis of electrocardiogram (ECG). The ECG of AF episodes have irregular heart beat interval (RR) and/or low P-wave amplitude in the QRS complex. Different algorithms had been explored to distinguish AF from other kind of arrhythmias using RR. Examples of such algorithms are the Root Mean Square Successive Differences (RMSSD), Sample Entropy and Fast Fourier transform [11, 12]. However, the patients may not be aware of the occurrence of AF as its occurrence is usually unpredictable [13,14]. The prevalence of asymptomatic AF diagnosed parenthetically during clinical visit is approximately 20% [15,16]. Hence, Holter-monitoring is commonly utilized to trace the ECG signal of a patient for 24 to 48 hours [17]. It is time consuming and inefficient for a trained specialist to examine through such huge amount of data to search for abnormal episode which might be only less than one minute in length. This makes the automatic detection of AF episodes with various duration and random location from the recorded ECG data essential.

In this paper, a computationally efficient algorithm that is able to accurately recognize AF episode from ECG data recorded for up to 48 hours is proposed. The algorithm analyzes the dynamics of RR interval obtained from an ECG signal. The diagnosis is made by examining the variability and randomness of the RR interval.

*Corresponding Author: Chee Teck Phua
School of Engineering, Nanyang Polytechnic, Singapore

The algorithm consists of three major analyzing stages. The first stage is to identify QRS complexes and RR intervals based on Pan-Tompkins method. The second stage is to detect the presents of arrhythmia in the ECG signal and identify the exact start and stop time of the episode. The last stage is the essential part of the process which is to single out AF from other arrhythmia episode. Shannon entropy (ShE) is typically computed for this purpose. For arrhythmia episode with ShE higher than the threshold value it will be recognized as an AF episode. But the threshold value varies with time duration of the AF episode. To ensure a reliable diagnosis of AF episode with different time durations, the relationship between ShE threshold value and duration of AF episode is studied. The results show a linear relationship. A formula for computing the ShE threshold of AF episode with various durations is proposed.

This finding of a linear relationship between ShE threshold and time duration of an AF episode is essential as the duration of an AF episode varies. For patients diagnosed with an early stage of AF, the duration of an AF episode is usually short and recognition is crucial. The algorithm is validated using ECG data provided by the MIT-BIH database with different recording, period up to 48 hours. Both ECG signals with various types of arrhythmia and normal ECG signals are concatenated for evaluation. The algorithm demonstrates 100% reliability in detecting AF episodes with accuracy of 99.5% in determining the start and stop time of each AF episode.

II. METHOD

The ECG data utilized to optimize and evaluate the algorithm is obtained from several MIT-BIH databases which include Normal Sinus Rhythm (NSR), Arrhythmia and AF databases. The 48 hour ECG signal is constructed primarily from NSR ECG with randomly embedded Arrhythmias and AF episodes of different durations.

2.1 Pan-Tompkins method

Pan-Tompkins method is chosen for its simplicity and accuracy to identify the QRS complex from the 48 hour ECG signal and thereafter, extract the RR intervals. Pan-Tompkins method is reported to have up to 99.3% accuracy in identifying QRS complex in ECG signal and can be utilized to avoid false detection and reduces interference that is present in ECG signals [18]. In particular, this method is known to be capable of attenuating baseline drift, power line interference, electromyographic noise and T wave interference [18].

The ECG signal is first normalized and cleaned from the DC components, followed by passing through band-pass filter to remove unwanted noises. The QRS complexes in ECG are recognized through differentiation, squaring and integration analysis of the amplitude, slope and width of the ECG signal. The last step is to determine R peak by the decision rule algorithm in which a threshold value is typically employed [18].

2.2 RMSSD computation

To identify AF episodes from 48 hours of recorded signal requires significant amount of computational time. Thus, Root Mean Square of the Successive Differences (RMSSD) is used to isolate ECG segments with abnormal RR interval, which can be a result of either AF or other arrhythmias. As an efficient time-domain method, computation of RMSSD greatly reduces the total computational time. RMSSD is a common method to estimate beat-to-beat variability [19]. For an normal episode, the beat-to-beat variation of ECG signal is lesser than abnormal episodes so RMSSD will be smaller. RMSSD is calculated based on equation 1[19]:

$$x_{RMSSD} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} = \sqrt{\frac{x_1^2+x_2^2+\cdots x_n^2}{n}} \tag{1}$$

In the formula, n is the number of RR intervals; x_i represents the *i*th RR interval. The threshold value to distinguish arrhythmia episode from NSR episode is calculated and used to identify abnormal segments. The RMSSD value of the 48 hours long ECG single is computed using a sliding window method. The length of the sliding window is 5 minutes. The RMSSD value of the data fall into the first sliding window which starts from time 0 to 5 minutes is computed first. After that the RMSSD of signal falls within the second sliding window will be calculated. The second sliding window include the next 5 minutes data by increasing the starting time of the window from time 0 to 40 seconds, thus RMSSD of signal from time 40 seconds to 5 minute 40 seconds will be computed. The 40 seconds spacing between each window is optimised based on the study of percentage of relative errors and computational time cost. After computing of RMSSD the abnormal ECG segment is identified.

2.3 The Autocorrelation of the squared signal

Autocorrelation analyzes the cross-correlation of a random process between a given time series and a lagged version of itself. Thus, the degree of similarity between the two parties can be determined. Every abnormal segments identified from the previous step is divided into 5 sections each with a length of 1 minute

*Corresponding Author: Chee Teck Phua
School of Engineering, Nanyang Polytechnic, Singapore

and the autocorrelation of them is computed. For an arrhythmic episode the energy of the signal will be higher compare with the NSR episode. Thus, auto-correlating the squared signal allows highlighting the region of the signal presenting a high rate of energy and permits determination of the start time and stop time of the arrhythmic episode. The equation used for autocorrelation is shown in equation 2:

$$\overline{R_{xx}}(m) = E[x_{n+m}^2 \cdot (x_n^*)^2] = E[x_n^2 \cdot (x_{n-m}^*)^2] \quad (2)$$

where E[.] $=\sum_{i=1}^{\infty} x_i \; p_i$, x_i is ECG signal of length n, p_i is probability of occurrence of x_i and m is the lagging interval. To determine the exact starting and stopping time of the arrhythmic episode, a 10th degree polynomial is computed using equation below, after that the slope is computed as shown in equation 3:

$$y = \sum_{i=1}^{11} p_i x^{10-i+1} \quad (3)$$

2.4 The Shannon Entropy

She was first proposed in the 1940s and has been commonly utilized in the information science field [20]. It quantitatively measures the uncertainty of a random variable. The equation utilized in the process is shown in equation 4:

$$ShE(s_i) = \sum_i s_i^2 \log(s_i^2) \quad (4)$$

where, s_i is the coefficient extracted from the signal, in an orthonormal basis. For episodes with high uncertainty the results of ShE will be high. And the AF episode can be singled out due to the higher uncertainty of AF episode compare with other arrhythmic episodes. Thus, an accurate threshold ShE value is required for the process. But the threshold values vary with the length of ECG episodes. And the length of AF episode is random. To ensure that the result is precise the relationship between the length of AF episode and the threshold value is calculated and a liner relationship is realized.

III. RESULTS

3.1 Pan-Tompkins method

The Pan-Tompkins method determines the R peaks with an accuracy of up to 99.3%. After applying it to the ECG signals, the QRS complex is extracted and remapped to the original ECG data. The resultant waveform with denoted QRS complex is shown in figure below.

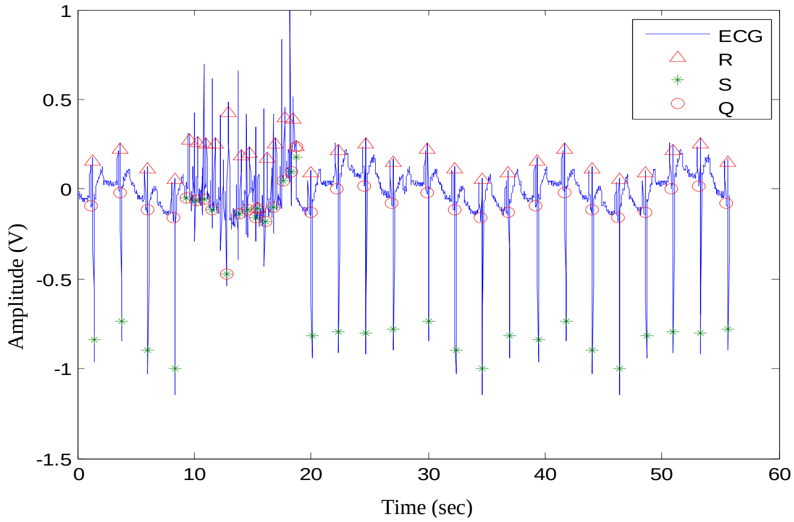

Figure 1. Detection of QRS complex using Pan-Tompkins method

3.2 RMSSD

To be able to evaluate the presence of an anomaly in the ECG signal, the threshold for RMSSD mean value needs to be determined. In order to determine this threshold, the RMSSD mean value of NSR, AF and other arrhythmia are computed as shown in Table 1. It is found that the RMSSD mean value of some arrhythmia episodes are very close to AF episode. This is reasonable since it has been reported that the range of RMSSD value for some other arrhythmia overlaps with AF.[19] Thus, subsequent process is necessary to distinguish AF from other arrhythmia episode. The RMSSD is only capable of isolating abnormal episode from NSR. The RMSSD mean value of NSR, arrhythmia and AF are normalized with the maximum value and the result is shown in Table 1. The threshold value is calculated by taking the average of the RMSSD mean value for an AF and NSR episode.

*Corresponding Author: Chee Teck Phua
School of Engineering, Nanyang Polytechnic, Singapore

	NSR	AF	Arrhythmia
RMSSD mean	0.209	0.379	**0.347**
Threshold calculated	**0.294**		

Table 1. RMSSD mean and threshold

Besides that the duration of spacing between sliding windows is also optimized based on the study of percentage of relative error and computational time. The percentage of relative error is computed based on equation 5:

$$P_{Error} = \frac{|x_{real} - x_{estimate}|}{x_{real}} . 100 \tag{5}$$

where, x_{real} is the real starting time of the abnormal segment and $x_{estimate}$ is the estimated starting time. Table 2 shows the relationship between the duration of spacing and percentage of relative error. Figure 2 illustrates the processing time needed for different spacing duration. For longer spacing duration the percentage of relative error is high, but the processing time is short. Based on the data obtained a spacing duration of 40 seconds is chosen which has comparatively low percentage of relative error (as shown in table 2) and short processing time (as shown in figure 2). After the RMSSD process, abnormal segments including AF and other arrhythmia sections are identified from the 48 hour signals from the MIT-BIH database.

Sliding window (seconds)	Percentage of relative error (%)
10	**0.1076**
20	**0.0646**
30	**0.1792**
40	**0.0190**
50	**0.0873**
120	**0.0227**
180	**25.0868**

Table 2. Percentage of relative error function of sliding window

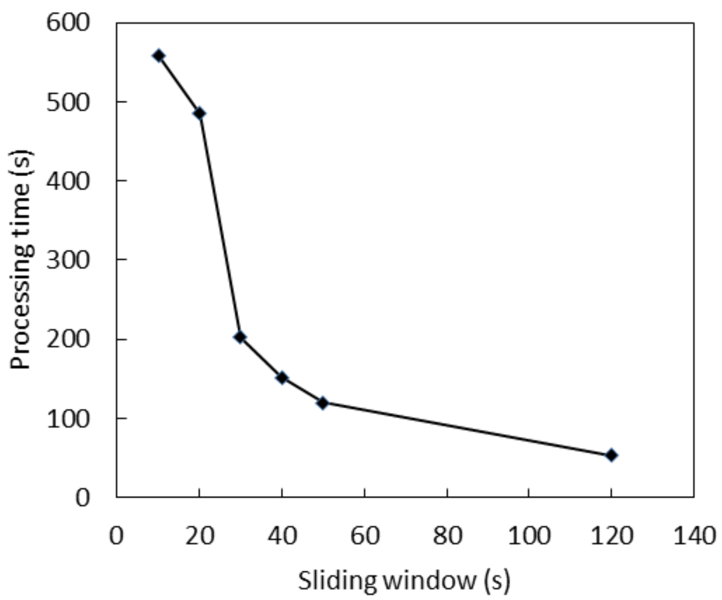

Figure 2. Duration of sliding window versus processing time

3.3 The autocorrelation of the squared signal

The autocorrelation process is applied on the above identified abnormal segments to determine the exact starting and stopping time of arrhythmia episode. The percentage of relative error is also computed for arrhythmia episode with various duration and random location. It is found that an accuracy of 99.5% can be achieved on determination of the starting and stopping time. Figure 3 shows the ECG signal after the autocorrelation and 10th degree polynomial process.

*Corresponding Author: Chee Teck Phua
School of Engineering, Nanyang Polytechnic, Singapore

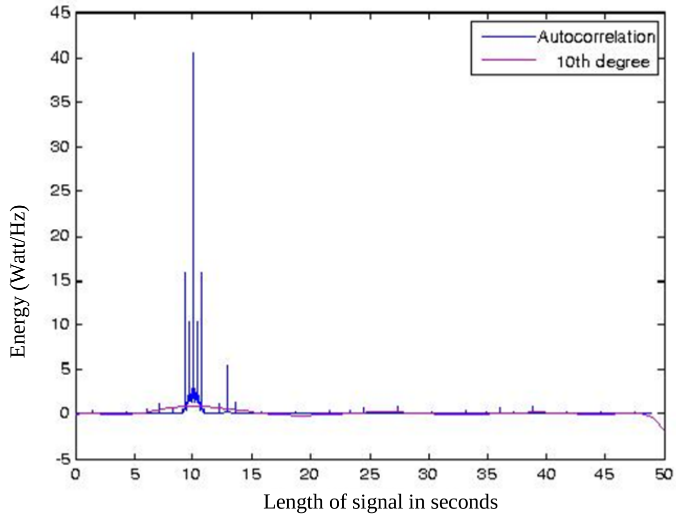

Figure 3. ECG after autocorrelation and 10th degree polynomial function process

3.4 The Shannon Entropy

To evaluate the relationship between ShE and the length of the episode. The relationship between the two parties is plotted as shown in Figure 4.

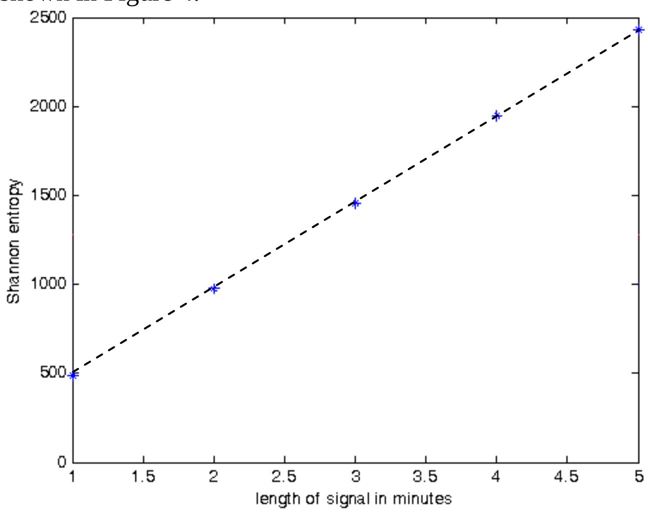

Figure 4. Evolution of the Shannon Entropy function based on signal length.

It is found that ShE evolves linearly with the length of the signal. A formula is proposed as shown in equation (6) based on the data obtained which can be utilized to calculate the ShE for various duration of episode. Through testing with the signal from MIT-BIH database, this step can discern 100% of the AF episode from other arrhythmias episodes:

$$Th_{ShE} = 500\ (\ Tstop\text{-}Tstart\ ) \qquad (6)$$

IV. CONCLUSION

In this paper, we show that AF episode with various duration can be accurately detected and differentiate from other kind of arrhythmia. The computational time needed is short, where a 48 hours ECG requires approximately 8 minutes. The algorithm applies the Pan-Tompkins method for ECG RR interval extraction with the combination application of the RMSSD, Autocorrelation of the squared signal and Shannon Entropy methods. In addition, an optimized RMSSD sliding window is proposed together with a formula for calculating the AF threshold of the ShE value. The latter is based on the study of the relationship between threshold values of ShE obtained as a result of different durations of an AF episode. This algorithm is capable of automatic determination of AF episodes within an ECG signal with an accuracy of up to 99.5% in the determination of the start and stop time of each AF episode.

*Corresponding Author: Chee Teck Phua
School of Engineering, Nanyang Polytechnic, Singapore

REFERENCES

[1]. J. Heeringa, D. A. van der Kuip, A. Hofman, J. A. Kors, G. van Herpen, B. H. Stricker, T. Stijnen, G. Y. Lip, and J. C. Witteman. "Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam study." European Heart Journal, vol. 27, pp.949–953, 2006.

[2]. 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.

[3]. Y. Miyasaka, M. E. Barnes, B. J. Gersh, S. S. Cha, K. R. Bailey, W. P. Abhayaratna, J. B. Seward, and S. T. Tsang, "Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence." Circulation, vol. 114, pp. 119-125, 2006

[4]. E. J. Benjamin, P. A. Wolf, R. B. D'Agostino, H. Silbershatz, W. B. Kannel, and D. Levy, "Impact of atrial fibrillation on the risk of death: the Framingham Heart Study." Circulation, vol. 98, pp.946–952, 1998.

[5]. S . Stewart, C. L. Hart, D. J. Hole, and J. J. McMurray, "A population-based study of the long-term risks associated with atrial fibrillation: 20-year follow-up of the Renfrew/Paisley study." The American Journal of Medicine, vol. 113, pp. 359–364, 2002.

[6]. W. B. Kannel, R. D. Abbott, D. D. Savage, and P. M. McNamara, "Epidemiologic features of chronic atrial fibrillation: the Framingham study." The New England Journal of Medicine, vol. 306, pp.1018–1022, 1982.

[7]. K. G. Peters and M. G. Kienzle, "Severe cardiomyopathy due to chronic rapidly conducted atrial fibrillation: complete recovery after restoration of sinus rhythm." The American Journal of Medicine, vol. 85, pp.242–244, 1988.

[8]. A. Ott, M. M. Breteler, M.C. de Bruyne, F. van Harskamp, D. E. Grobbee, and A. Hofman, "Atrial fibrillation and dementia in a population-based study. The Rotterdam study." Stroke; a Journal of Cerebral Circulation, vol. 28, pp.316-321, 1997.

[9]. W. B. Kannel and E. J. Benjamin, "Status of the epidemiology of atrial fibrillation." The Medical Clinics of North America, vol. 92, pp.17-40, 2008.

[10]. W. B. Kannel, et al "Prevalence, incidence, prognosis, and predisposing conditions for atrial fibrillation: population-based estimates." The American Jouranl of Cardiology, vol. 82, pp.2N–9 N, 1998.

[11]. D. Singh, K. Vinod, and S. C. Saxena, "Sampling frequency of the RR interval time series for spectral analysis of heart rate variability." Journal of Medical Engineering & Technology, vol. 28, pp.263-272, 2004.

[12]. I. Goldenberg, A. J. Moss, and W. Zareba, "QT Interval: How to Measure It and What Is "Normal"." Journal of Cardiovasc Electrophysiol, vol. 17, pp. 333-336, 2006.

[13]. C. W. Israel, G. Gronefeld, J. R. Ehrlich, Y.-G. Li, and S. H. Hohnloser, "Long-term risk of recurrent atrial fibrillation as documented by an implantable monitoring device: implications for optimal patient care." Journal of the American College of Cardiology, vol. 43, pp.47-52, 2004.

[14]. R. L. Page, W. E. Wilkinson, W. K. Clair, E. A. McCarthy, and E. L. Pritchett, "Asymptomatic arrhythmias in patients with symptomatic paroxysmal atrial fibrillation and paroxysmal supraventricular tachycardia." Circulation, vol. 89, pp.224– 227, 1994.

[15]. K. H. Humphries, C. R. Kerr, S. J. Connolly, G. Klein, J. A. Boone, M. Green, R. Sheldon, M. Talajic, P. Dorian, and D. Newman, "New-onset atrial fibrillation: sex differences in presentation, treatment, and outcome." Circulation vol. 103, pp.2365–2370, 2001.

[16]. C. Kerr, J. Boone, S. Connolly, M. Greene, G. Klein, R. Sheldon, and M. Talajic, "Follow-up of atrial fibrillation: the initial experience of the Canadian Registry of Atrial Fibrillation." European Heart Journal, vol. 17, pp.48–51, 1996.

[17]. F. Portet, "P wave detector with PP rhythm tracking: evaluationin different arrhythmia contexts." Physiological Measurement, vol. 29, pp.141–155, 2008.

[18]. J. Pan and Willis. J. Tompkins, "A Real-Time QRS Detection Algorithm." IEEE Transactions on biomedical engineering. vol. BME-32, pp.230-236, 1985.

[19]. B. Anuradha, K. Suresh Kumar and V.C. Veera reddy, "Classification of cardiac signals using time domain methods." vol. 3, ISSN 1819-6008, 2008.

[20]. C. E. Shannon, "A mathematical theory of communication," The Bell System Technical Journal, vol. 27, pp. 379–423, 623–656, 1948.

*Corresponding Author: Chee Teck Phua
School of Engineering, Nanyang Polytechnic, Singapore