Single-Feature Method for Fast Atrial Fibrillation Detection in ECG Signals

Lucie Marsanova[1], Andrea Nemcova[1], Radovan Smisek[1,2], Martin Vitek[1], Lukas Smital[1]
[1] Department of Biomedical Engineering, Brno University of Technology, Brno, Czech Republic
[2] Academy of Sciences, Brno, Czech Republic

Abstract

Atrial fibrillation (AF) is the most common arrhythmia in adults and is associated with a higher risk of heart failure or death. Here, we introduce simple and efficient method for automatic AF detection based on symbolic dynamics and Shannon entropy. This method comprises of three parts. Firstly, QRS complex detection is provided, than the raw RR sequence is transformed into a sequence of specific symbols and subsequently into a word sequence and finally, Shannon entropy of the word sequence is calculated. According to the value of Shannon entropy, it is decided, whether AF is present in the current cardiac beat. We achieved sensitivity Se=96.32% and specificity Sp=98.61% on MIT-BIH Atrial Fibrillation database, Se=91.30% and Sp=90.8% on MIT-BIH Arrhythmia database, Se=95.6% and Sp=80.27% for Long Term Atrial Fibrillation database and Se=93.04% and Sp=87.30% for CinC Challenge database 2020. The achieved results of our one-feature method are comparable with other authors of more complicated and computationally expensive methods. Our ECG experts found that public databases contain errors in annotations (in sense of AF). It means that results are affected by errors in annotations. Many errors were found in Long-Term AF database, several also in MIT-BIH AF database and MIT-BIH Arrhythmia database. Testing algorithms on poorly annotated databases cannot bring reliable results and algorithms useful in real medical practice. The examples of such annotations are reported in this study.

1. Introduction

Cardiovascular disorders are still the most common cause of death worldwide. Due to the ease of use, non-invasiveness and cheapness, electrocardiogram (ECG) is nowadays still the most available and widely used method for the cardiovascular system examination [1]. Atrial fibrillation (AF) is the most common arrhythmia in adults and is associated with a higher risk of heart failure or death.

AF is a supraventricular tachyarrhythmia which is represented by inconstant atrial activation and, therefore, dysregulation of atrial contractions. This cause uncompleted blood transfer from atria to ventricles and decrease the efficiency of heart functioning. This can result in serious complications such as ischemia, stroke, or early mortality [1]. Therefore, early detection of AF is crucial for effective treatment. Automatic detection of AF in ECG is still problematic, as was shown by the results of previous studies. Here, we introduce simple and efficient method for automatic AF detection based on symbolic dynamics and Shannon entropy.

2. Method

This method comprises of three parts. Firstly, QRS complex detection is provided by detector based on phasor transform, and sequence of RR intervals is computed. In the second part, the raw RR sequence is transformed into a sequence of specific symbols and subsequently into a word sequence. Finally, Shannon entropy of the word sequence is calculated. According to the value of Shannon entropy, it is decided, whether AF is present in the current cardiac beat.

2.1. QRS complex detection

Firstly, signal is filtered by bandpass filter with cut-off frequencies 12 and 19 Hz. In this frequency range lies the most of QRS complex energy. Using this filter suppresses P and T waves and also high frequency artifacts. On the other hand, QRS complexes are highlighted.

After that, phasor transform (PT) is applied. PT transforms each sample of the signal into a complex value preserving the signal information [2].

Converting each ECG sample into a phasor enhances changes in the ECG signal (the waves). The degree with which ECG waves are enhanced in phasor signal is determined by value R_V. The value of R_V is always within the interval 0-1. A constant value R_V is considered as a real part, whereas the imaginary component is the original value of the ECG sample:

$$y(n) = R_V + jx(n), \tag{1}$$

where $y(n)$ is the phasorial signal and $x(n)$ is the original sample of signal. The magnitude $M(n)$ is computed as

$$M(n) = \sqrt{R_V^2 + x(n)^2}, \tag{2}$$

and phase (phasor) $\varphi(n)$ is computed as

$$\varphi(n) = tan^{-1}\left(\frac{x(n)}{R_V}\right). \tag{3}$$

In phasor signal $\varphi(n)$, the QRS complexes have in all cases higher amplitude than the other ECG components. This applies also in a case of original ECG signals with smaller amplitude of QRS complex than T wave. Due to this fact, the QRS complexes in phasor signal are easier detectable [2].

In the case of QRS detection, phasor transform of ECG signal with R_V = 0.001 is performed. The example of phasor transform (grey curve) of ECG signal (black curve) prepared for QRS detection is in Fig. 1. Than we search for maxima in sliding window with 300 ms size. The next step is a check, whereas the found maxima are higher than used adaptive threshold. At the end it is checked, whether the interval between two subsequent QRS complexes (RR interval) is 1.75x higher than previous one. If it is accomplished, backward searching is performed. This detector was used and tested also in our previous studies [3], [4], [5].

The example of phasor transform of ECG signal is shown in Fig. 1. In phasor signal $\varphi(n)$, the QRS complexes have in all cases higher amplitude than the other ECG components. This applies also in a case of original ECG signals with smaller amplitude of QRS complex than T.

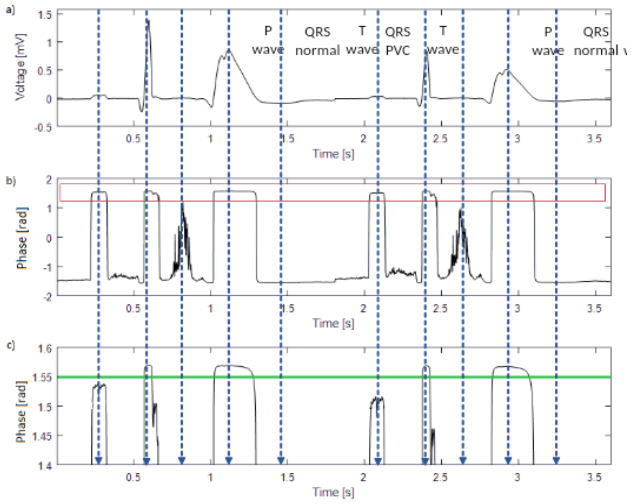

Figure 1 An example of (a) an original ECG signal; (b) phase signal φ(n); (c) detail of the top part of the phase signal (illustrated by red box in (b)), the green line represents the threshold for QRS detection. [3]

2.2. Symbolic dynamic and Shannon entropy

The purpose of using symbolic dynamics (symbols and words) is to describe the dynamic of heart rate. During atrial fibrillation, high variability of RR intervals is present and thus also high value of Shannon entropy [6].

Firstly, heart rate *hr(n)* is derived from RR intervals (*RR(n)*). *RR(n)* are computed from length two consecutive heart beats (*R(n)*). The sequence of raw RR intervals is transformed into the heart rate (*hr(n)*) according to the eg. (4) and then quantified into symbol sequence (*Sy(n)*) [6] according to the equation (4)

$$hr(n) = 60/RR(n), \tag{4}$$

$$Sy(n) = \{\ 63 \quad if\ hr(n) > 315\ \lfloor \frac{hr(n)}{5} \rfloor \qquad other\ , \tag{5}$$

The sequence of symbols *Sy(n)* represent instantaneous state of heart rate transformed into the 64 possible symbols. To facilitate the analysis of *Sy(n)* is used 3-symbol template for examination of entropic properties of sequence. This transformation is computed according to the eg. (6) and after that we obtain sequence of word (*wv(n)*) [7]. Each word takes into account 3 successive symbols (3 heart beats).

$$wv(n) = (sy(n-2).2^{12}) + (sy(n-1).2^{6}) + sy(n), \tag{6}$$

Finally, Shannon entropy (*SH*) is computed (7). SH is a statistical tool that quantifies a time series (in our case length of heart beats) in terms of the information size [8]. At first, we define the discrete probability space of a dynamic system as *A = (A|P)*. *A* represent set of characteristic elements $A = \{a_1, \cdots, a_k\}$, and $P = \{p_1, \cdots, p_k\}$ $(1 \leq k \leq N)$ is relevant probability. Each element a(i) has probability $p(i) = N(i)/N$ $(0<p(i)<1, \sum_{i=1}^{k} p(i) = 1$ where N is number of all element in *A* and *N(i)* is total number of element *a(i)* in *A*. Thus, the *SH* of *A* is defined as [26],

$$SH(A) = -\frac{k}{N.log_2} \sum_{i=1}^{k} p(i) log_2 p(i) \tag{7}$$

In our work, the dynamic system *A* consists of 95 consecutive word elements from *wv(n-47)* to *wv(n+47)*. The value of *SH* lies in interval <0,1>. It means that on computing of *SH* for *R(n)* is needed information of *RR(n)* of 95 consecutive heart beats. According to the value of SH is determined if atrial fibrillation is present or not in actual heart beats. In our work, we found that the best threshold the value for discrimination between AF and non AF is T=0.733.

In Figure 2 process of classification is illustrated. Subgraph b) shows values of Shannon entropy transformed from length of RR intervals and also threshold (blue line) for decision of presence of atrial fibrillation (over line) or other rhythm (below line). It is clear visible that value of Shannon entropy correlated with annotation of atrial fibrillation c) (value 1 represents AF, value 0 represents no AF).

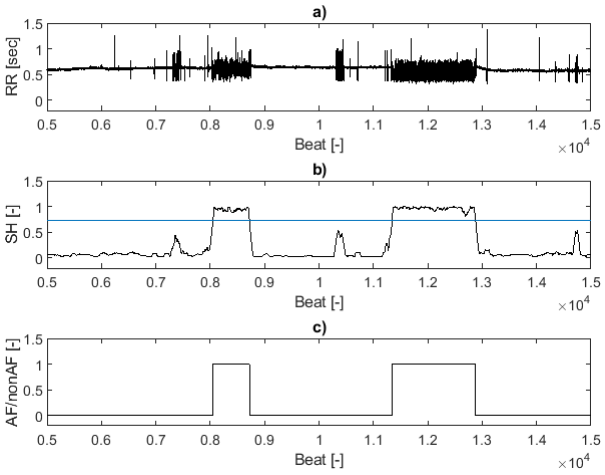

Figure 2 a) Length of RR intervals, b) Shannon entropy of RR intervals – blue line is treeshold for decision of presence AF, c) annnotation – value 1 represents AF, 0 represents nonAF

3. Results

We used four publicly available databases for testing of our algorithm. The results are summarized in Table 1. The exact values of true positive (TP), true negative (TN), false positive (FP), false negative (FN) are reported in Table 2. We achieved sensitivity Se=96.32% and specificity Sp=98.61% on MIT-BIH Atrial Fibrillation database, Se=91.30% and Sp=90.80% on MIT-BIH Arrhythmia database, Se=95.6% and Sp=80.27% for Long Term Atrial Fibrillation database and Se=93.04% and Sp=87.30% for CinC Challenge database 2020.

Table 1 The performance of AF detection algorithm signals from MIT-BIH arrhythmia database, MIT-BIH atrial fibrillation database, CinC chalange database (1sted) and Long term atrial fibrillation database (Se – sensitivity; PP – positive predictivity).

Database	SE [%]	PP [%]
MIT-BIH AF	96,32	98,61
MIT-BIH AR	98,42	90,78
CinC Ch. (1sted)	93,04	87,30
Long Term AF	95,60	80,27

Table 2 The values of true positive (TP), true negative (TN), false positive (FP), false negative (FN) heart beats of AF detection algorithm on signals from MIT-BIH arrhythmia database, MIT-BIH atrial fibrillation database, CinC chalange database (1sted) and Long term atrial fibrillation database.

Database	TP	TN	FP	FN
MIT-BIH AF	80964655	125029215	1764240	3093610
MIT-BIH AR	2891746	23887217	2424716	46321
CinC Ch. 1sted)	1145	4908	76	748
Long Term AF	437589872	284396825	69905003	20145419

The achieved results of our one-feature method are comparable with other authors of more complicated and computationally expensive methods [9], [10]. Fast a simplicity of our method is useful for example in application where the time is limitation and also for mobile devices, where the low computational complexity is important. Biosignals processing is nowadays an actual topic.

In Figure 3, examples of errors caused by our detection algorithm, are shown. The blue lines indicate annotation from database, green lines indicate our results, value 1 indicates AF, value 0 indicates nonAF. In a) is shown mistake caused by earlier termination AF of our detector – probably due to the fact, that SH is computed from 95 RR intervals and following RR intervals without AF decreased value of SH. These mistakes of inaccurate termination and beginning of AF is relatively frequent. In b) is shown that our detector pointed to the left part of the signal, where supraventricular arrhythmia is present, as atrial fibrillation, probably due to the fact of higher variability of RR intervals caused by this type of arrhythmia as with AF is also present.

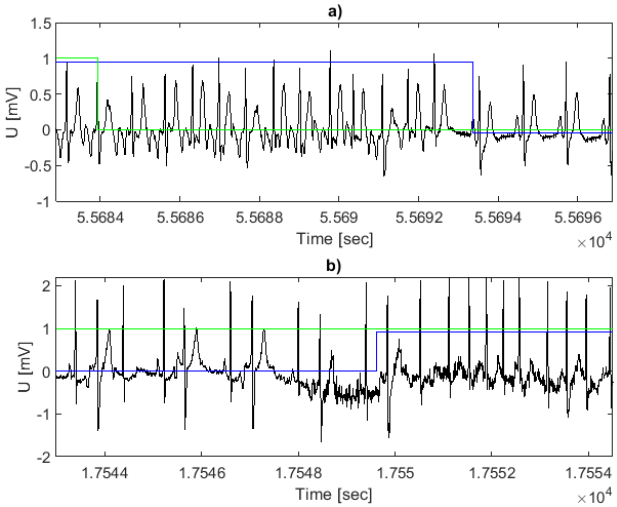

Figure 3 Examples of errors caused by our detection algorithm (blue line – annotation, green line - our result, 1-AF, 0-nonAF), a) signal 100 from Long term atrial fibrillation database, b) signal 05261 from MIT-BIH atrial fibrillation database.

4. Discutions

During our work, our ECG experts made **a** very important finding. Public databases contain many errors in annotations (in sense of AF) and also in other pathologies. It means that results of all authors, who used these databases, are affected by errors in annotations. Many errors were found in Long-Term AF, several also in MIT-BIH AF database and also in MIT-BIH Arrhythmia database. Testing algorithms on poorly annotated databases cannot bring reliable results and algorithms

useful in real medical practice. The examples of such annotations are reported in Figure 4. In the left part of subgraph a) is present AF, but in annotation from database (blue line) is not marked (signal no. 05261 from MIT-BIH atrial fibrillation database), correct annotation is marked by red line. In subgraph b) is present AF all the time, but according to the annotation is it not true. The part of the signal where the extrasystoles are, is not marked as AF.

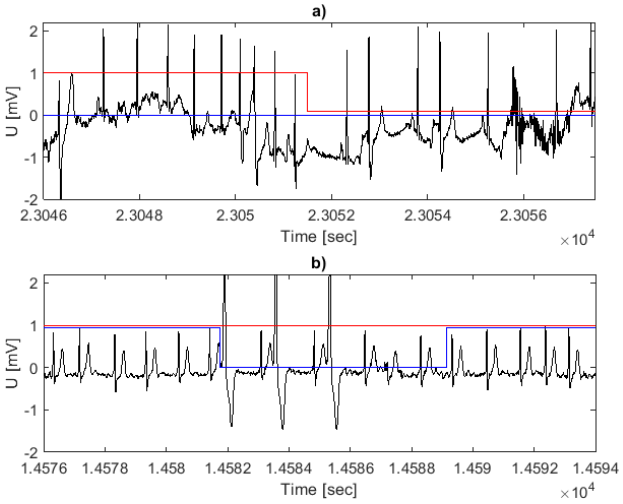

Figure 4 Examples of errors in annotation of AF (blue - annotation in database, red - correct annotation, , 1-AF, 0-nonAF); a) signal 05261 from MIT-BIH atrial fibrillation database, b) signal 100 from Long term atrial fibrillation database.

5. Conclusion

In this work, highly efficient single feature algorithm for atrial fibrillation is proposed. The achieved results of our one-feature method are comparable with other authors of more complicated and computationally expensive methods [9], [10]. Fast a simplicity of our method is useful for example in application where the time is limitation and also for mobile devices, where the low computational complexity is important.

In addition, our ECG experts found that in often used testing databases of ECG signals are errors in annotation (in sense of AF) and also in other pathologies. Testing algorithms on poorly annotated databases cannot bring reliable results and algorithms useful in real medical practice. There is a place for their correction or creation of new ones.

Acknowledgments

This work has been funded by the United States Office of Naval Research (ONR) Global, award number N62909-19-1-2006.

References

[1] B. F. Kusumoto, "ECG interpretation: from pathophysiology to clinical application," *Springer*, pp. 107, 2009.

[2] A. Martínez, R. Alcaraz, J.J. Rieta, "Application of the phasor transform for automatic delineation of single-lead ECG fiducial points," *Physiological Measurement* 31, pp. 1467-85, 2011.

[3] L. Maršánová, et al., "Automatic detection of P wave in ECG during ventricular extrasystoles," *Proceedings of the World Congress on Medical Physics and Biomedical Engineering*, pp. 381-85, 2018.

[4] L. Maršánová, A. Němcová, R. Smíšek, et al., "Advanced P wave detection in ECG signals during pathology: Evaluation in Different Arrhythmia Contexts," *Scientific Report*, no. 9, 2019.

[5] L. Maršánová, et al., "Automatic detection of P wave in ECG during ventricular extrasystoles," *World Congress on Medical Physics and Biomedical Engineering 2018, IFMBE Proceedings*, pp. 381-385, 2018.

[6] X. Zhou, et al., "A Real-Time Atrial fibrillation detection algorithm based on the instantaneous state of heart rate," *PLoS ONE*, vol. 10, no. 9, 2015.

[7] X. Zhou, et al., "Automatic online detection of atrial fibrillation based on symbolic dynamics and Shannon entropy", *BioMedical Engineering OnLine*, vol. 13 no. 1, 2014.

[8] A. Afdala, et al., "Automatic Detection of atrial fibrillation using basic Shannon entropy of RR Interval feature," *Journal of Physics: Conference Series*, vol. 795, 2017.

[9] Y. Li, X. Tang, A. Wang,. et al., Probability density distribution of delta RR intervals: a novel method for the detection of atrial fibrillation. *Australas Phys Eng Sci Med*, vol. 40, pp. 707–716, 2017.

[10] M. Lown, et al., Machine learning detection of atrial fibrillation using wearable technology, *PLoS ONE*, vol. 15, no. 1, 2020.

Address for correspondence:

Ing. Lucie Maršánová
Department of Biomedical Engineering
Faculty of Electrical Engineering and Communication
Brno University of Technology
Technická 12
Brno 616 00
Czech Republic.
marsanova@vutbr.cz