DINCON 2017
CONFERÊNCIA BRASILEIRA DE DINÂMICA, CONTROLE E APLICAÇÕES
30 de outubro a 01 de novembro de 2017 – São José do Rio Preto/SP

Simple wavelet-based features for arrhythmia identification from HRV signals based on an artificial immune system

Caio Cesar Enside de Abreu[1]
Department of Computing, UNEMAT, Alto Araguaia, MT
Bruno Rodrigues de Oliveira[2]
Department of Electrical Engineering, UNESP, Ilha Solteira, SP
Marco Aparecido Queiroz Duarte[3]
Department of Mathematics, UEMS, Cassilândia, MS
Jozue Vieira Filho[4]
Department of Electrical Engineering, UNESP, Ilha Solteira, SP
Francisco Villarreal[5]
Department of Mathematics, UNESP, Ilha Solteira, SP

Abstract. This paper presents two simple wavelet-based features for automatic arrhythmia identification from heart rate variability (HRV) signals. Decomposition provided by the stationary wavelet transform is used in order to filter possible changes in the HRV signal baseline and highlight abrupt fluctuations between two consecutive RR-intervals. From detail wavelet coefficients, two simple features named Range and Peaks are derived and used as input to a classifier. The classifier implemented is an artificial immune system (AIS) that has not yet been applied to the related problem. Results showed that proposed features, as well as the AIS are suitable for arrhythmia identification from HRV series, pointing that future research should be conduced based on new combinations of features encompassing the proposed ones.

Keyword. Arrhythmia Identification, HRV Analysis, Stationary Wavelet Transform, Artificial Immune System.

1 Introduction

Currently the leading cause of death in the world is related to heart disease. The main way to make prognostics of health heart is based on electrocardiogram (ECG) analysis [8]. In this sense, an important measurement related to the ECG is the time interval fluctuation between two consecutive beats (RR-intervals). Commonly referred as HRV analysis, it is

[1]caioenside@unemat.br
[2]bruno@cerradosites.com
[3]marco@uems.br
[4]jozue@dee.feis.unesp.br
[5]villa@mat.feis.unesp.br

often used by researchers in order to identify or classify arrhythmias [5,8]. An HRV signal is a time series containing RR-intervals from the whole ECG signal.

Automatic detection of abnormal cardiac rhythms from normal heart activity became an important task for clinical reasons. Through the analysis of HRV signals, researchers aims to train classifiers able to predict if an HRV observation is normal or abnormal. Therefore, besides the choice of the classifier, a challenge is to achieve consistent features capable to discriminate between normal and abnormal HRV observations. Current approaches have focused on time and frequency domain analysis. Classical time domain features are statistical, geometric and energetic [8]. Frequency domain analysis is based mainly on power spectrum density estimation [7]. Additionally, other HRV series transformations are used. Examples are Hilbert-Huang transform [6] and wavelet transform [4]. Wavelet domain extracted features are wavelet entropy measures [4].

The motivation for this paper is the proposition of simple wavelet-based features for automatic arrhythmia identification, by means of an AIS. Researchers in the AIS field look to the biological immune system (BIS) for inspiration on how to solve problems in engineering and computer science [2]. The AIS used in this work is the negative selection algorithm (NSA) [3]. NSA is a relatively recent binary classifier that has been applied successfully to several areas; some examples are: anomaly detection and pattern recognition [1]. Among its characteristics, NSA is attractive for its simplicity of implementation and high accuracy in pattern recognition [1,2]. NSA is mainly based on simple comparisons between patterns by means of an affinity measure, e.g. a distance measure.

In [8], authors used a large set of features extracted in both, time and time-frequency domain. Then, simulations encompassing several combinations of features were done in order to find the best result. In fact, due to the high number of possible features, researchers usually need some feature reduction approach as in [5]. Therefore, this work proposes two simple wavelet-based features for arrhythmia identification from HRV signals, contributing with the related area. Results presented in this paper enables future investigations on combinations of the proposed features with classical time, time-frequency or frequency domain features. As a secondary objective, NSA performance on arrhythmia identification will be investigated.

The remainder of this paper is organized as follows: Section 2 presents the proposed features. NSA in its original form is described in Section 3; Section 4 presents simulation details; and, finally, concluding remarks are presented in Section 5.

2 Proposed features

Consider a length N HRV observation $y[n] = [y_0, y_1, \ldots, y_{i-1}, y_i, \ldots, y_{N-1}]$, the following normalization is adopted

$$\bar{y}_i = \frac{y_i}{\max\left(|y[n]|\right)}. \tag{1}$$

Afterward, the undecimated wavelet transform, also known as stationary wavelet transform (SWT), is applied over $\bar{y}[n]$. SWT provides J octave bands on the HRV observation, where J is the number of decomposition levels, previously defined. It is possible since detail coefficients, provided by wavelet decomposition, correspond to the frequency situated

approximately between $(2^{-j}f_s, 2^{-j-1}f_s)$, where $j = 1, \ldots, J$ is the current decomposition level and f_s is the signal sampling rate.

In the proposed feature extraction scheme, SWT with $J = 1$ decomposition level is used in order to filter possible changes in the HRV signal baseline and highlight abrupt fluctuations between two consecutive RR-intervals. This fact is showed in Figure 1.

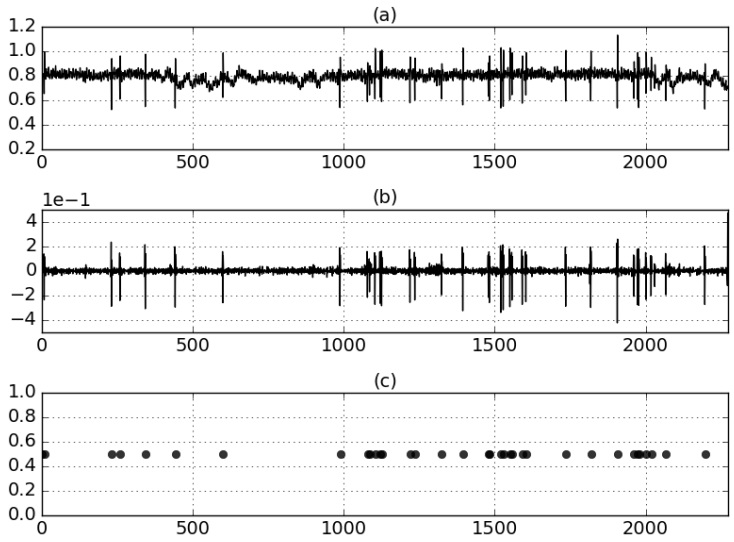

Figure 1: Decomposition of an HRV signal observation by the SWT using Coiflets 2 wavelet: (a) HRV signal; (b) detail coefficients at scale $j = 1$; (c) arrhythmia markers.

Note, from Figure 1, that the HRV signal baseline changes are removed. This is possible because these changes represent low frequency components, present in approximation coefficients and not detail coefficients. Detail coefficients are related to high frequency components. In addition, arrhythmias are highlighted from HRV signal.

From SWT detail coefficients $d[j, n]$ at scale $j = 1$, two features named Range and Peaks are proposed in equations (2), (3) and (4). Range feature is defined as follows:

$$\text{Range} = \max\left(d[1, n]\right) - \min\left(d[1, n]\right). \tag{2}$$

It is noteworthy that Range feature in (2) only is viable due to decomposition provided by SWT. In equations (3) and (4) Peaks feature is presented.

$$p[n] = \begin{cases} 1, & \text{if } |d[1, n]| > \lambda \\ 0, & \text{if } |d[1, n]| \leq \lambda. \end{cases} \tag{3}$$

$$\text{Peaks} = \sum_{n=0}^{N-1} p[n]. \tag{4}$$

λ threshold was empirically defined and it is evaluated adaptively on each HRV observation as $\lambda = 0.3 \max\left(d[1, n]\right)$. Analyzing equation (3), it can be seen that $p[n]$ is an auxiliary vector used to count the occurrence of values above λ.

For a better understanding of the proposed features, boxplot and scatter plot for both features evaluated on training data are respectively shown in Figures 2 and 3.

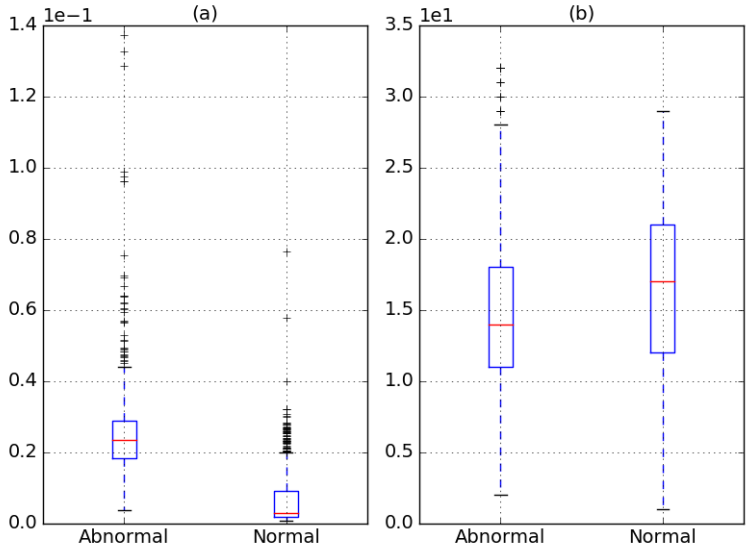

Figure 2: Proposed features boxplot: (a) Range; (b) Peaks.

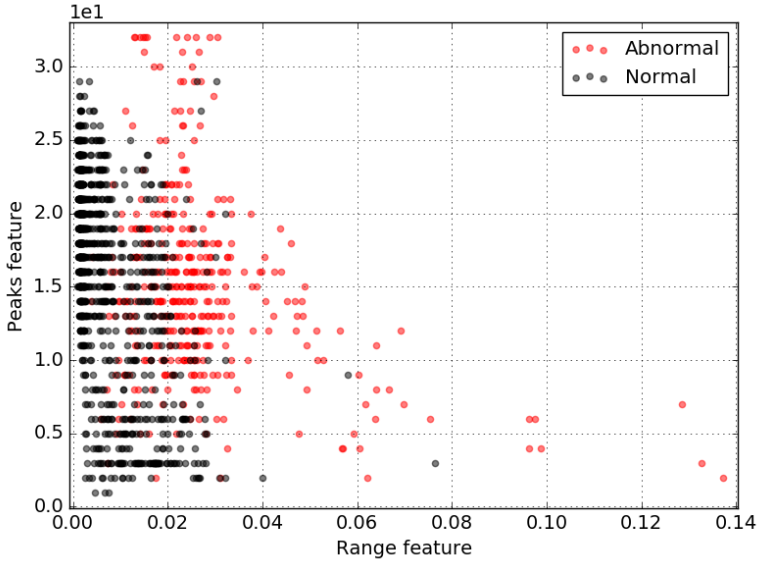

Figure 3: Scatter plot of the proposed features.

Note from Figure 2 that Range feature is suitable for the discrimination between normal and abnormal classes, separating near fully arrhythmic from normal HRV signal observation. Despite Peaks feature does not provide a discrimination as accurate as Range, it is

still possible to distinguish both classes. Together, the proposed features are consistent. This fact is pointed out in Figure 3. It can be seen from Figure 3 that Range and Peaks features, together, can discriminate between two classes, being useful for HRV analysis.

3 Negative selection algorithm

For the suitable functioning and thus preventing of autoimmune disease development, BIS needs to be able to distinguish between molecules of our own cells, called self molecules and foreign molecules, generally referred to as self and non-self, respectively [3]. The elimination of any cell with receptors able to recognize self molecules, named auto-reactive cells, characterizes the negative selection concept.

NSA was developed based on the negative selection of T-cells within the thymus. The thymus is responsible for T-cell maturation. Thus, only T-cells that do not recognize self-molecules are allowed to survive [2]. Based on this biological concepts, NSA is defined as follows [1, 3]:

1. Define two different datasets $\mathbf{P}$ and $\mathbf{C}$. $\mathbf{P}$ contains only self patterns and $\mathbf{C}$ contains both self and non-self patterns. Compute the affinity (match) between each element $c_j \in \mathbf{C}$ and each element in $\mathbf{P}$. If an element in $\mathbf{C}$ is recognized by an element in $\mathbf{P}$, in other words, if the affinity score between c_j and an element in $\mathbf{P}$ exceeds an established threshold, reject c_j. Otherwise, store c_j in a set of detectors $\mathbf{R}$. This phase of the NSA is commonly called the Censoring phase.

2. After the generation of $\mathbf{R}$, the current phase consists of system monitoring in order to detect non-self patterns. A set $\mathbf{P}_*$ is defined to be protected and the affinity between each element in $\mathbf{P}_*$ and each $r_j \in \mathbf{R}$ is evaluated. If such an affinity is superior to a predefined threshold, then a non-self element is identified. This phase is called the Monitoring phase.

In the present implementation, an HRV signal normal observation is referred to as self pattern. Thus, several kind of cardiac arrhythmias are referred to as non-self patterns. The censoring phase aims to build a set of arrhythmia detectors $\mathbf{R}$. Upon the building of $\mathbf{R}$, monitoring phase will act in the task of arrhythmia identification on test data in the following way: if an HRV signal observation matches with a $r_j \in \mathbf{R}$, it is classified as arrhythmic. In both NSA phases, the match is checked based on an affinity measure, commonly defined as a distance measure [2].

4 Simulation and results

In order to evaluate the proposed features and NSA performance, simulations were carried out using the MIT-BIH arrhythmia database. MIT-BIH arrhythmia database consists of 48 ECG records with half-hour each taken from 47 patients. Only records 201 and 207 were not used in the simulations for technical reasons. Data were split into training and test sets with 23 recordings each. The HRV signal was taken from each record

and segmented in 32-points observations, as well as in [8]. Each 32-points observation was marked as normal or arrhythmic according to database annotations: if an observation contains more than 29 normal beats, then the segment is considered normal; otherwise, the segment is arrhythmic. As well as in [8], 13 kinds of arrhythmia are considered for identification.

In order to calibrate the classifier, Canberra distance was selected as affinity measure. In the NSA, the matching between two patterns can be perfect or partial [2]. In perfect matching, the two patterns are equal. However, in real signal applications the concept of partial matching is often employed [2]. Partial matching is achieved by defining a matching threshold δ. Therefore, for any two instances x and y , if $d(x; y) < \delta$, the match occurs.

The matching threshold δ was defined empirically according to the classifier performance. Performance was checked based on four objective measures: accuracy, sensitivity, specificity and predictive. Upon several simulations encompassing the matching threshold δ and several wavelet families, better results for all objective measures were found with $\delta = 0.09$ and rbio1.3 wavelet function. Results and comparisons with current works are presented in Tables 1 and 2.

Table 1: Main results encompassing different wavelet functions.

Wavelet	Accuracy (%)	Sensitivity (%)	Predictive (%)	Specificity (%)
sym6	85.36	85.54	85.23	85.18
sym13	85.36	82.86	87.22	87.86
coif2	86.07	85.18	86.73	86.96
bior4.4	85.27	84.82	85.59	85.71
rbio1.3	**86.52**	**85.71**	87.11	87.32
rbio1.5	86.34	84.29	**87.90**	**88.39**

Table 2: Current studies on arrhythmia classification from HRV series.

Method	Features	Segments classified	Acc/Sens/Pred/Spec, (%)
Tsipouras and Fotiadis (2004)	63	2000	- /89.95/ - /92.91
Jovic and Jovic (2017)	17-138	-	91,20/91.20/ - /97.10
Proposed Features	2	1120	86.52/85.71/87.11/87.32

It can be seen from Tables 1 and 2 that the proposed features are consistent and provide satisfactory results with two simple features only. Note that the obtained results are close to the ones considered as state of art; in particular, they are very close to the results in [8]. Regarding the NSA, results show that this classifier is suitable for arrhythmia identification due to its intrinsic characteristics of pattern recognition.

5 Conclusion

This paper has presented two simple wavelet based features for HRV analysis. Stationary wavelet transform was used to remove changes from the HRV signal baseline and

highlight abrupt fluctuations between two consecutive RR-intervals. In addition, NSA performance on the related problem was investigated. Results showed that proposed features, as well as NSA are suitable to arrhythmia identification from HRV series, pointing that future research encompassing several combinations of the proposed features with traditional time, time-frequency and frequency domain analysis are necessary.

Acknowledgments

Authors would like to thank Coordination for the Improvement of Higher Education Personnel (CAPES) - Brazil.

References

[1] C. C. E. Abreu, M. A. Q. Duarte, F. Villarreal, An immunological approach based on the negative selection algorithm for real noise classification in speech signals, *AEÜ - Int. J. Electron. Commun.*, vol. 72, 125-133, (2017), DOI:10.1016/j.aeue.2016.12.004.

[2] D. Dasgupta and F. Nino, *Immunological computation: theory and applications*, CRC Press, Boca Raton, (2008).

[3] S. Forrest, et al., Self-nonself discrimination in a computer, In *IEEE Computer Society Symposium on Research in Security and Privacy*, Oakland, California, United States of America, 202-212, (1994), DOI:10.1109/RISP.1994.296580

[4] Y. Isler and M. Kuntalp, Combining classical HRV indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure, *Comput. Biol. Med.*, vol. 37, 1502-1510, (2007), DOI:10.1016/j.compbiomed.2007.01.012.

[5] A. Jovic and F. Jovic, Classification of cardiac arrhythmias based on alphabet entropy of heart rate variability time series, *Biomed. Signal Process. Control*, vol. 31, 217-230, (2017), DOI:10.1016/j.bspc.2016.08.010.

[6] H. Li, et al., Hilbert-Huang transform for analysis of heart rate variability in cardiac health, *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 8, 1557-1567, (2011), DOI:10.1109/TCBB.2011.43.

[7] G. Panina, et al., Role of spectral measures of heart rate variability as markers of disease progression in patients with chronic CHF not treated with angiotensin-converting enzymeinhibitors, *Am. Heart J.*, vol. 131, 153–157, (1996), DOI:10.1016/S0002-8703(96)90064-2.

[8] M. G. Tsipouras and D. I. Fotiadis, Automatic arrhythmia detection based on time and time-frequency analysis of heart rate variability, *Comput. Methods Programs Biomed.*, vol. 74, 95-108, (2004), DOI:10.1016/S0169-2607(03)00079-8.