HRV Signal Dynamic Extraction in the Poincare Plot by analyzing the Extended U-Sequences for Cardiac Arrhythmia Classification

Pouria Sarlak[1], Amir homayoon Jafari[2], Gholamreza Attarodi[1], Nader Jafarnia Dabanloo[1],
Seyed Kamaledin Setarehdan[4], Nazanin Hemmati[1]

[1]Science and Research Branch, Azad University, Tehran, Iran
[2]Tehran Medical University, Tehran, Iran
[4]Tehran University, Tehran, Iran

Abstract

This article presents a new method for extracting the dynamic of heart rate variability signal by studding the position of the points in the Poincare plot with respect to the line of identity for cardiac arrhythmia classification.

In this method the points of HRV signal in Poincare plot are changed in to digital codes. Then, sequences of the codes are formed. At the end, by analyzing the manner of the repetitions and the continuance of these sequences in different arrhythmia classes, it could be possible to differentiate the arrhythmias from each other.

One of the major advantages of this method is to discriminate arrhythmia by applying only 10 samples of RR intervals.

To evaluate the performance of the proposed method, 3 different types of arrhythmia classes are utilized and they are classified with an accuracy of 95.6%, 93.04% and 97.21%, respectively.

1. Introduction

The heart rate is evaluated by measuring the time interval between the successive R-peaks (R–R interval) of the ECG waveform [1]. The variation in the time series of consecutive heartbeats is referred as heart rate variability (HRV) [2]. HRV signal, is a nonlinear and non-stationary signal that represents the autonomic activity of the nervous system and the way it influences the cardiovascular system [3]. HRV believed to be a good marker of the individual's health condition and heart diseases. Therefore, HRV analysis became a critical tool in cardiology for the diagnosis of heart diseases [4].

RR intervals are extracted samples of ECG signals obtained by Poincare plot in the embedding space. If each interval $RR(i+1)$ is plotted as a function of the previous interval $RR(i)$, then the resulting plot is known as the Poincare plot [7]. Poincare plot is a geometrical representation of RR time series to demonstrate patterns of heart rate dynamics resulting from nonlinear processes [7]. This plot can be quantitatively analyzed by calculating the standard deviations of the distances of the points $RR(i)$ from the lines $y = x$ and $y = x + 2RR_m$, where RR_m is the mean of all $RR(i)$ values. These standard deviations are denoted by $SD1$ and $SD2$, respectively[7]. Some studies have been done for extracting these features from HRV signal [5-3] whereas SD1 and SD2 are linear statistics and unable to extracting the dynamic of HRV signal.

The line of identity (LOI) ($y = x$) in the Poincare plot has a simple physiological interpretation: the points on this line correspond to equal consecutive RR intervals, the points above it correspond to increasing heart rate and the points below this line to decreasing heart rate[6].

In [7], Mohareri et al. proposed 2 new features of the Poincare plot for classifying cardiac arrhythmia by analyzing HRV signal. The first one, is a 3×1 matrix which its elements are defined as the number of points above, on and bellow the line of identity. The second one, is a 3×3 matrix consisting of nine different behaviors that 2 points have in relation to each other and line of identity in Poincare plot. Their method was only on the based of counting number of the cases. Hence, the signal dynamic has been ignored.

This study presents a new method for classification of cardiac arrhythmia by analyzing the HRV signal dynamic. In this method, each point in the Poincare plot is coded respecting its position to the LOI, following which, a series of triple code sequences are formed. These sequences will keep the codes corresponding the points in succession and at the end, the manner of the repetitions and the continuance of these sequences in various arrhythmia classes will be analyzed and indicates the HRV signal dynamic.

The paper is organized as follows: In section 2,

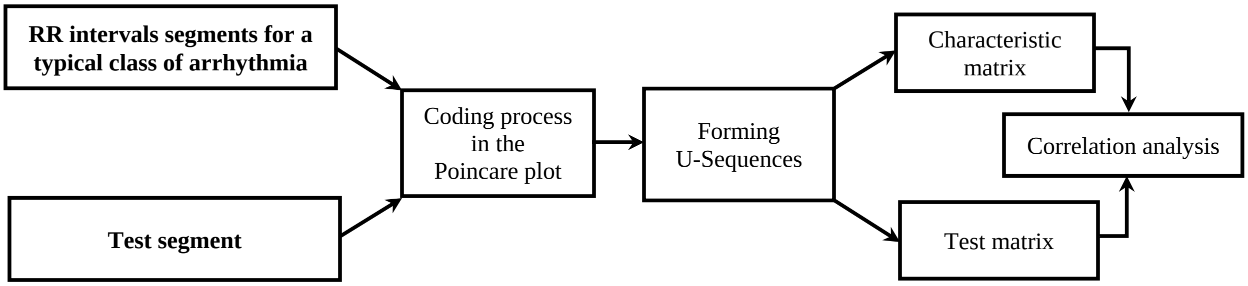

Figure 1. Block diagram of proposed method

proposed method is presented. Results will be given in section 3, and finally discussion and conclusion are in section 4.

2. Proposed method

In this method, by using all the train segments of each group of cardiac arrhythmia, a characteristic matrix is constituted for each group, according to the block diagram, shown in Fig. 1. Then, to distinguish an unknown segment selected from test data, first it is transformed in to the form of above mention characteristic matrix called “Test matrix”. Then, the correlation of the test matrix and each one of the characteristic matrix of different arrhythmia classes will be computed. The unknown segment belongs to the group that results the highest amount of the correlation.

In continue, each block is described in more details.

2.1. Coding process in the Poincare plot

First, a segment with 10 samples of RR intervals is plotted in the Poincare plot. Each point in this plot is shown by $p_i(RR_i, RR_{i+1})$. Then, three different algorithms are proposed for coding process as follows:

2.1.1. First algorithm

In this algorithm each point is coded just in accordance with its position with respect to the LOI. The coding process in this algorithm is defined as follow:

$$p_i(RR_i, RR_{i+1}) = \begin{cases} 0, & RR_i > RR_{i+1} \\ 1, & RR_i \le RR_{i+1} \end{cases} \quad (1)$$

2.1.2. Second algorithm

For coding the points in this algorithm, as shown in Fig. 2, in addition to the position of the points with respect to the LOI, the distance between the points and LOI is taken into consideration as follow:

$$p_i(RR_i, RR_{i+1}) = \begin{cases} 0, & d < -\text{dis}_1 \\ 1, & -\text{dis}_1 \le d < 0 \\ 2, & 0 \le d < \text{dis}_2 \\ 3, & \text{dis}_2 \le d \end{cases} \quad (2)$$

Where dis_1 and dis_2 are the distances between 2 parallel line with LOI bellow and above it, respectively as they are shown in Fig. 3, and d is the distance of a point in the Poincare plot and LOI which is defined as :

$$d = \frac{RR_i - RR_{i+1}}{\sqrt{2}} \quad (3)$$

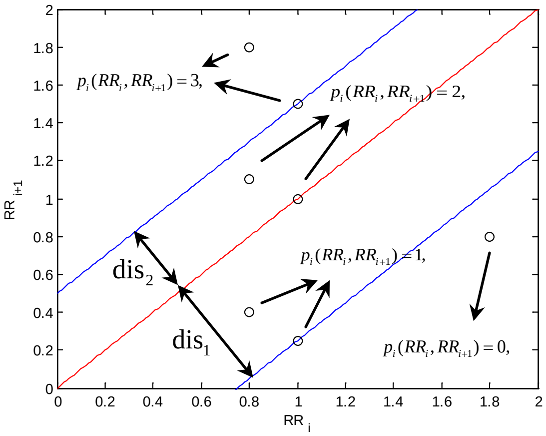

Figure 2. Coding process in second algorithm

2.1.3. Third algorithm

In the third one, more information regarding the distance between the points above the LOI and the line itself is taken into consideration, but in order to keep the simplicity of the computation, the information related to the distance of the points below the LOI is omitted as follow:

$$p_i(RR_i, RR_{i+1}) = \begin{cases} 0, & d < 0 \\ 1, & 0 \le d < \mathrm{dis}_1 \\ 2, & \mathrm{dis}_1 \le d < \mathrm{dis}_2 \\ 3, & \mathrm{dis}_2 \le d \end{cases} \quad (4)$$

Where dis_1 and dis_2 are the distances between 2 parallel line above the LOI shown in Fig.3 and d is defined as per equation (3).

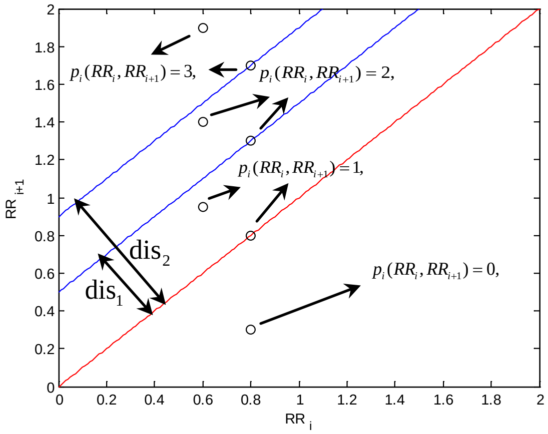

Figure 3. Coding process in third algorithm

2.2. Forming U-Sequence

To analyse the signal dynamic, a series of triple code sequences are formed as per Table.1, so the U-Sequences are formed.

Table 1. Forming U-Sequences

PP Points	Code	U-Sequence	V
(RR_1, RR_2)	c_1		
(RR_2, RR_3)	c_2	$c_1c_2c_3$	v_1
(RR_3, RR_4)	c_3		
(RR_4, RR_5)	c_4	$c_2c_3c_4$	v_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$
(RR_n, RR_{n+1})	c_n		
(RR_{n+1}, RR_{n+2})	c_{n+1}	$c_nc_{n+1}c_{n+2}$	v_n
(RR_{n+2}, RR_{n+3})	c_{n+2}		
(RR_{n+3}, RR_{n+4})	c_{n+3}	$c_{n+1}c_{n+2}c_{n+3}$	v_{n+1}

For further analyzing and processing, the decimal value of each U-Sequence is computed as per equation (5) and saved in a 8×1 matrix which is called "V".

$$v_i = c_i \times x^2 + c_{i+1} \times x^1 + c_{i+2} \times x^0 \quad (5)$$

Where, with consideration to value of the codes (c_i), the highest and lowest amount of V elements (v_i) are shown in Table. 2.

Table 2. c_j, v_i and x limit

	c_j	v_i	x
1st algorithm	0,1	$0 \le v_i \le 7$	2
2nd algorithm	0,1,2,3	$0 \le v_i \le 63$	4
3rd algorithm	0,1,2,3	$0 \le v_i \le 63$	4

2.3. Characteristic matrix

For analyzing the manner of the repetitions and the continuance of U-Sequences, the characteristic matrix is called "P_r" matrix will be formed as per equation (6). In this equation, N_t is number of 10 RR intervals segments which belongs to train data. T_v is a $(8 \times N_t)$ matrix having a V matrix for each of its column, formed by a segment from train data. $v_{i\max}$ is the highest value of v_i as presented in Table. 2 for each algorithm, and $v_i(c_j)$ is defined by omitting the highest place value in the U-Sequence corresponding v_i and replacing the code c_j in its lowest place value, that is:

$$\begin{aligned} v_i &= c_{j-3} \times x^2 + c_{j-2} \times x^1 + c_{j-1} \times x^0, \\ v_i(c_j) &= c_{j-2} \times x^2 + c_{j-1} \times x^1 + c_j \times x^0, \end{aligned} \quad (7)$$

Where v_i and c_j are defined in Table. 2. By considering v_i and c_j, P_r is a 8×2 matrix obtained from the first and 64×4 matrix from the second and third algorithm.

2.4. Correlation analysis

To distinguish an unknown segment selected from test data, first it should be changed into a matrix "V" being defined in section 3.2. Then another matrix is formed similar to matrix P_r called matrix "H", corresponding that matrix H is constructed just from matrix V obtained from one test segment. The size of matrix H is the same as matrix P_r.

$$(8)$$

$$\mathrm{H}(v_i + 1, c_j + 1) = \mathit{Number\ of}\ \{\mathrm{V}(x) = v_i \ \wedge \ \mathrm{V}(x+1) = v_i(c_i)\}$$

$$\forall\ x = 1, 2, \ldots, v_{i\max} + 1,$$

$$\mathrm{P}_r(v_i+1, c_j+1) = \frac{Number\ of\ \left\{\ T_v(x,y)=v_i\ \wedge\ \mathrm{T}_v(x+1,y)=v_i(c_j)\ \right\}}{N_t \times 7}, \qquad \begin{array}{l} \forall x = 1,2,...,v_{i\max}+1, \\ \forall y = 1,2,...,N_t, \end{array} \tag{6}$$

Table 3. Classification result

	1st algorithm			2nd algorithm			3rd algorithm		
	Sensitivity %	**Specificity %**	**Accuracy %**	**Sensitivity %**	**Specificity %**	**Accuracy %**	**Sensitivity %**	**Specificity %**	**Accuracy %**
NSR	55.00	86.19	71.67	95.60	98.83	97.09	94.87	96.38	95.68
AFIB	55.30	76.75	67.11	87.80	95.75	92.20	90.20	95.36	93.04
B	98.92	77.09	78.99	95.07	94.82	91.31	97.18	97.22	97.21

At the end, the correlation between matrix H and each one of matrix P_r defined for different groups of arrhythmia will be computed. The unknown segment belongs to the group resulting the highest amount of the correlation.

The correlation between matrix H and each one of matrix P_r is defined as follows:

$$S_k = \sum_m \sum_n \mathrm{H}(m,n) \times \mathrm{P}_{rk}(m,n), \tag{9}$$

Where k is a class of arrhythmia and P_{rk} is its matrix P_r.

3. Results

To evaluate the performance of the proposed method and comparing 3 algorithm, 3 groups of HRV signals are applied which includes Normal sinus rhythm (NSR), Atrial fibrillation (AFIB) and Ventricular bigeminy (B). The HRV data of these arrhythmia provided by the MIT-BIH Arrhythmia Database [8]. By dividing these data into 10 RR intervals segments, finally, it is acquired 1100 NSR segments, 1071 AFIB segments and 206 B segments. The HRV signals at each class are randomly divided into the train and test sets in an approximate ratio of 2/3 and 1/3, respectively.

The results of specificity, sensitivity and accuracy parameters are shown in Table. 3 for 3 algorithms.

4. Discussion and conclusion

In this article, a new method for analysing HRV signal dynamic was presented that it was applied for classification of 3 groups of cardiac arrhythmia.

With consideration to privilege of second and third algorithm results with respect to first algorithm, shows in Table. 3, it is clear that by considering just the position of the points with respect to LOI, they can not be classified from each other, but their distance from LOI is also important.

By using this method, the classification is performed only by applying 10 samples of RR intervals, and due to short processing time, it can be used in real-time application.

References

[1] Rajendra Acharya U, Subbanna Bhat P. Classification of heart rate data using artificial neural network and fuzzy equivalence relation. Pattern Recognition 2003; 61 – 68

[2] McLernon D, Dabanloo N. A new nonlinear model for generating RR tachograms. Computers in Cardiology 2004:481-4.

[3] Babak Mohammadzadeh Asl, Seyed Kamaledin Setarehdan. Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. Artificial Intelligence in Medicine 2008; 44, 51-64.

[4] Tsipouras MG, Dimitrios I. Fotiadis. Automatic arrhythmia detection based on time and time-frequency analysis of heart rate variability. Computer Methods and Programs in Biomedicine 2004; **74**, 95-108.

[5] Thalange AV, Mergu RR. HRV Analysis of Arrhythmias Using Linear-Nonlinear Parameters. International Journal of Computer Applications 2010;75-80.

[6] Piskorski J, Guzik P. Geometry of the Poincar´e plot of RR intervals and its asymmetry in healthy adults. Physiological measurement 2007; 287-300.

[7] S Moharreri, S Parvaneh,. Utilizing Occurrence Sequence of Heart Rate’s Phase Space Points in order to Discriminate Heart Arrhythmia. Proceedings of the 17th Iranian Conference of Biomedical Engineering (ICBME2010) 2010

[8] Goldberger A, Amaral L, Glass L, Hausdorff J, Ivanov P, Mark R, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation. 2000;101(23):e215.

Address for correspondence.
Amir homayoon Jafari
Tehran University of Medical Sciences, Poorsina St., Tehran, IRAN.