Proceedings of the 17th Iranian Conference of Biomedical Engineering (ICBME2010), 3-4 November 2010

Utilizing Occurrence Sequence of Heart Rate's Phase Space Points in order to Discriminate Heart Arrhythmia

S Moharreri, S Parvaneh, N Jafarnia Dabanloo
Islamic Azad University, Science and Research Branch
Tehran, Iran
sadaf.moharreri@gmail.com; parvaneh@ieee.org
n_jafarnia@yahoo.com

A M Nasrabadi
Shahed University
Tehran, Iran

Abstract— **Poincare plot analysis of RR time series allows a beat-to-beat approach to Heart Rate Variability (HRV), detecting patterns associated with nonlinear processes. Since the measurement of standards descriptors of Poincare plot is based on the point's distribution in relation to the line of identity ($y=x$), we have concentrated on it and evaluated the points behavior related to this line. For this purpose, we test two global and local analyses of points against the identity line. For evaluating these two novel features of Poincare plot, we try to use them for distinguishing four groups of subjects (Arrhythmia, Congestive Heart Failure (CHF), Atrial Fibrillation (AF) and Normal Sinus Rhythm (NSR)). Kruskal-Wallis test was used to define the level of significance of features. The results show that global feature discriminate different groups by $p<6E-7$, and local feature discriminate them by $p<2E-7$.**

Keywords- poincare plot; heart rate variability; temporal variations; occurrence

I. INTRODUCTION

Heart rate is an indicator of heart's condition [1]. Assessment of heart rate has been shown to aid clinical diagnosis and intervention strategies. It has been proved that nonlinear analysis of it might provide more valuable information for the physiological interpretation of heart rate fluctuations [2]. However, the variety of contradictory reports in this field indicates that there is a need for a more rigorous investigation of methods as aids to clinical evaluation [2]. The nonlinear analysis of Heart Rate Variability (HRV) is a valuable tool in both clinical practice and physiological research reflecting the ability of the cardiovascular system.

The Poincare plot is a tool developed by Henry Poincare for analyzing complex systems [1]. It has found its use in such diverse fields as physics and astronomy, geophysics, meteorology, mathematical biology and medical sciences. In the context of medical sciences it is mainly used for quantifying HRV and proves to be quite an effective measure of this marker [3]. Poincare plot is a geometrical representation of RR time series to demonstrate patterns of heart rate dynamics resulting from nonlinear processes. Poincare plot analysis of RR time series allows a beat-to-beat approach to HRV, detecting patterns associated with nonlinear processes. It is a familiar method for the analysis of two-dimensional nonlinear dynamic systems[4]. Tulppo et. al. [5] fitted an ellipse to the distribution of the poincare plot and defined two standard descriptors *SD1* and *SD2* for quantification of the poincare plot geometry. These standard descriptors represent the minor axis and the major axis of the ellipse (Fig. 1) and guide the visual inspection of the distribution. In case of HRV, it reveals a useful visual pattern of the RR interval data by representing both short and long term variations of the signal [6]. But standard descriptors *SD*1 and *SD*2 are linear statistics and hence the measures do not directly quantify the nonlinear temporal variations in the time series contained in the poincare plot. Moreover, the limitations of the *SD*1/*SD*2 analysis are important to understand when attempting to investigate the physiological mechanisms in a time series, or when analyzing data where the occurrence of nonlinear behavior may be a distinguishing feature between health and disease[6].

The identity line (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[7].Since the measurement of standards descriptors of poincare plot is based on the point's distribution in relation to the line of identity ($y=x$), we have concentrated on it and evaluated the points behavior related to

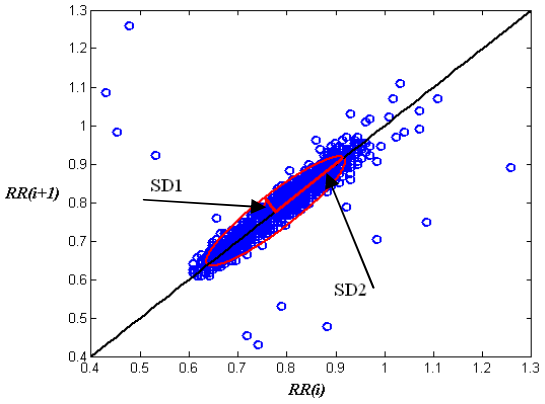

Figure 1. Poincare plot of RR intervals of a healthy person with its standard descriptors *SD1* and *SD2*

978-1-4244-7484-4/10/$26.00 ©2010 IEEE

this line. For this purpose, we defined two global and local analyses of points against the identity line. In global method, we contemplate all points in Poincare plot and constructed a *3×1* vector by counting the points above, under, and on the identity line. In local method, we focused on point to point variations against the identity line. So the temporal aspects of poincare plot have been detected. For this purpose, we constructed a *3×3* matrix by counting the number of points which have the same temporal variations against the identity line.

For evaluating these two novel features of Poincare plot, we try to use them for distinguishing four groups of subjects (Arrhythmia, Congestive Heart Failure (CHF), Atrial Fibrillation (AF) and Normal Sinus Rhythm (NSR)).

II. POINCARE PLOT

A. Standard Descriptors

A standard poincare plot of RR interval is shown in figure 1. Given a time series RR = $\{RR_1, RR_2, \ldots, RR_n, RR_{n+1}\}$ the standard poincare plot is a scattergram constructed by locating points from the time series on the coordinate plane according to the pairing (x_i, y_i) in which,

$$x = \{x_1, x_2, \ldots, x_n\} = \{RR_1, RR_2, \ldots, RR_n\} \quad (1)$$

$$y = \{y_1, y_2, \ldots, y_n\} = \{RR_2, RR_3, \ldots, RR_{n+1}\} \quad (2)$$

and i = 1, 2, 3, …, n and *n* is the number of points in the poincare plot which is one less than the length of the RR time series [7].

As mentioned above, *SD1* and *SD2* are two standard descriptors of poincare plot. *SD2* is defined as the standard deviation of the projection of the poincare plot on the line of identity ($y = x$), and *SD1* is the standard deviation of projection of the poincare plot on the line perpendicular to the line of identity ($y = -x$) [3]. So we may define them as:

$$SD1 = (Var(d_1))^{1/2}, \quad SD2 = (Var(d_2))^{1/2} \quad (3)$$

whereVar(d) is the variance of *d*, and

$$d_1 = (x-y) / (2)^{1/2}, \quad d_2 = (x+y) / (2)^{1/2} \quad (4)$$

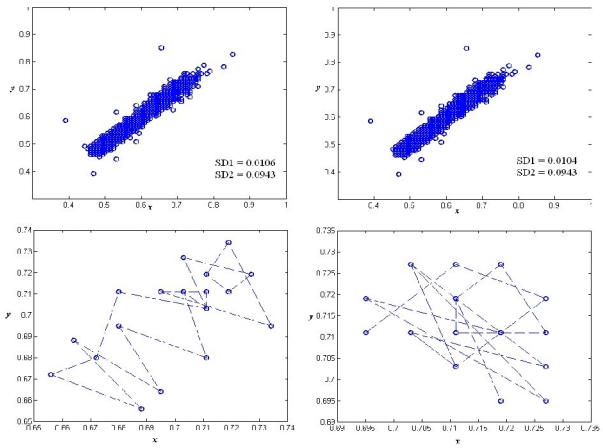

Figure 2. Temporal dynamics of poincare plots with similar *SD1* and *SD2*

B. Limitations of Standard Descriptors

Time is an implicit parameter in a poincare plot, so time-sequence information is lost and onlydistributional information is represented. So the lack of temporal information is the primary limitation of the standard descriptors of the poincare plot. *SD1* and *SD2* represent the distribution of signal in 2D space and carries only shape information [6]. It should be noted that many possible RR interval series result in identical plot with exactly similar *SD*1 and *SD*2 values in spite of different temporal structure [6]. Thus, the same plot can be generated by data sets with differentunderlying dynamics[8].In Fig. 2, two signals with similar *SD*1 and *SD*2 values are shown to be different in terms of temporal structure.

Temporal information is important for the detection of non-stationary behavior [2]. A simple but effective techniqueto include temporal features is to animate the construction of the poincare plot [1]. Hence, to reflect temporal variation, we developed a new method to incorporate temporal information in relation to the line of identity which can be used in quantification of the temporal dynamics of the system.

III. NOVEL FEATURES

In this section, we introduced our new features: Global Occurrence Matrix (*GOM*) and Co-Occurrence Matrix (*COM*). Both of them are defining in the base of point's distribution in relation to the line of identity. For this purpose, firstly, the theoretical development of each feature has been given and then they have been used for distinguishing different groups of subjects which is followed by statistical analysis.

A. Global Occurrence Matrix (GOM)

As mentioned earlier, the line of identity in the poincare plot is defined as the line that passes through the origin at an angle of 45° with *x*-axis [9]. We have defined our new features in a poincare plot dependent of the line of identity i. e. decision about a point is made based on its position with respect to the line of identity on the 2D poincare plot. In the proposed *GOM*, the points of the plots are partitioned into three parts (Fig. 3):

- Points which are up the line of identity (U);
- Points which are on the line of identity (O);
- Points which are down the line of identity (D).

The decision about a point as to whether it belongs to one of the above three classes is made based on the point's distance to the line of identity.For measuring the distance of a point to the line, we used:

$$Dist = (ax + by + c) / (a^2 + b^2)^{1/2} \quad (5)$$

in which for the line of identity and points in Poincare plot, we have: $a = -1;\ b = 1;\ c = 0$. So for points $P_i\ (x_i, y_i)$ of the poincare plot, we have:

$$Dist_i = (y_i - x_i) / (2)^{1/2} \quad (6)$$

In which *x* and *y* are as mentioned in (1) and (2) and $i = 1, 2, \ldots, n$. Therefore the status of the point P_i with respect to its distance to the line of identity is defined as follows:

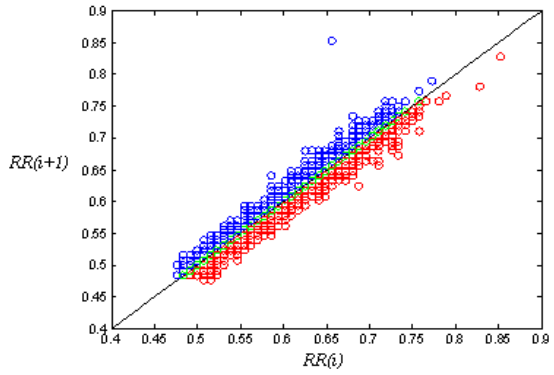

Figure 3. Three classes in calculating *GOM*

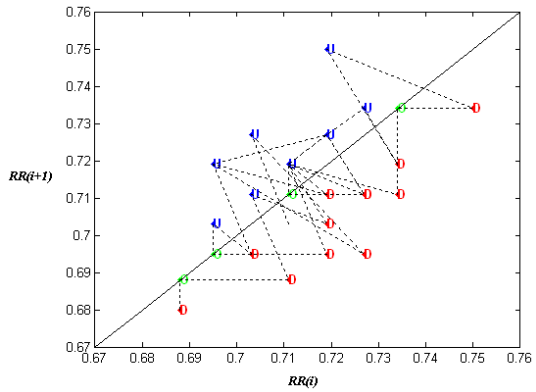

Figure 4.Temporal Variations in *COM*

- *If* $Dist_i > 0$ *then* $P_i \in (U)$
- *If* $Dist_i = 0$ *then* $P_i \in (O)$
- *If* $Dist_i < 0$ *then* $P_i \in (D)$

These three classes are shown in Fig. 3. After defining the classes of all the points, we counted the members of each class for constructing *GOM*.

GOM is a 3×1 matrix which elements are defined as follows:

$$GOM = [N_U \; N_O \; N_D] \qquad (7)$$

In which N_U is the number of points in class *U*, N_O is the number of points in class *O*, and N_D is the number of points in class *D*.

B. Co-Occurrence Matrix (COM)

For defining the second feature, we focused on local temporal behavior of the points in relation to each other dependent on the line of identity. For this purpose, we used the same definitions which were mentioned in previous section such as three different classes U, O, and D. But the difference is that in *COM*, we considered two following points P_i and P_{i+1} and so the analysis corresponds to at least three consecutive RR intervals of the RR interval time series. Therefore, in *COM* we should count nine different behaviors dependent to points' classes in relation to each other and line of identity which are defined as follows (Fig. 4):

- *If* $(P_i \in U)$ *&* $(P_{i+1} \in U)$ *then* $UU = UU + 1$
- *If* $(P_i \in U)$ *&* $(P_{i+1} \in O)$ *then* $UO = UO + 1$
- *If* $(P_i \in U)$ *&* $(P_{i+1} \in D)$ *then* $UD = UD + 1$
- *If* $(P_i \in O)$ *&* $(P_{i+1} \in U)$ *then* $OU = OU + 1$
- *If* $(P_i \in O)$ *&* $(P_{i+1} \in O)$ *then* $OO = OO + 1$
- *If* $(P_i \in O)$ *&* $(P_{i+1} \in D)$ *then* $OD = OD + 1$
- *If* $(P_i \in D)$ *&* $(P_{i+1} \in U)$ *then* $DU = DU + 1$
- *If* $(P_i \in D)$ *&* $(P_{i+1} \in O)$ *then* $DO = DO + 1$
- *If* $(P_i \in D)$ *&* $(P_{i+1} \in D)$ *then* $DD = DD + 1$

Hence, *COM* is a 3×3 matrix which elements are defined as follows:

$$COM = \begin{bmatrix} UU & UO & UD \\ OU & OO & OD \\ DU & DO & DD \end{bmatrix} \qquad (8)$$

IV. Discrimination of Heart Arrhythmia

In order to validate the proposed features, *GOM* and *COM*, we have used them to discriminate four groups of subjects (Arrhythmia, Congestive Heart Failure (CHF), Atrial Fibrillation (AF) and Normal Sinus Rhythm (NSR)).For each groups, we calculate *GOM* and *COM* separately.

The data from MIT-BIH Physionet database [10] are used in the experiment. In this study, we have used 15 long-term ECG recordings of subjects in normal sinus rhythm from Physionet Normal Sinus Rhythm database [10]. Furthermore, we have also used NHLBI sponsored Cardiac Arrhythmia Suppression Trial (CAST) RR-Interval Sub-study database for the arrhythmia data set from Physionet. Subjects of CAST database had an acute myocardial infarction (MI). The database is divided into three different study groups among which we have used the Encainide (e) group data sets for our study. From that group we have chosen 15 subjects belong to subgroup baseline (no medication)[10]. Also, we have used 15 long-term ECG recordings of subjects with CHF from Physionet Congestive Heart Failure database along with 15 ECG recordings of subjects with Atrial Fibrilation from Physionet Atrial Fibrilation database [10].The original long term ECG recordings in every four groups were digitized at 128 Hz[10].

Table 1.*p*-Value Results for *COM* Parameters

Groups	*COM* Parameters								
	UU	*UO*	*UD*	*OU*	*OO*	*OD*	*DU*	*DO*	*DD*
NSR, CHF	0.003*	0.2321	0.5502	0.5810	0.0169*	0.0273*	0.8541	0.3010	0.0731
NSR, CAST	0.3345	0.4906	0.2322	0.5198	0.6792	0.4482	0.1902	0.4762	0.4212
NSR, AF	0.0094*	8.29E-6*	0.0345*	6.66E-6*	6.60E-6*	8.29E-6*	0.0130*	6.68E-6*	7.76E-5*
CHF, CAST	0.0203*	0.0891	0.2701	0.2797	0.0386*	0.0564	0.3825	0.1826	0.0934
CHF, AF	0.0009*	1.02E-5*	0.0365*	2.86E-5*	6.60E-6*	6.66E-6*	0.0430*	6.66E-6*	5.25E-5*
CAST, AF	0.0006*	0.0011*	0.9268	0.0001*	1.55E-5*	1.36E-4*	0.6791	5.25E-5*	8.31E-6*
Total	9.13E-5*	4.43E-6*	0.1256	4.42E-6*	1.91E-7*	6.61E-7*	0.0929	1.13E-6*	2.50E-6*

Table 2.*p*-Value Results for *GOM* Parameters

Groups	*GOM Parameters*		
	U	*O*	*D*
NSR, CHF	0.0024*	0.0535	0.5970
NSR, CAST	0.8003	0.7827	0.5813
NSR, AF	0.0019*	6.70E-6*	1.02E-5*
CHF, CAST	0.0076*	0.0730	0.1542
CHF, AF	0.0001*	6.68E-6*	3.53E-5*
CAST, AF	0.003*	4.29E-5*	0.0002*
Total	2.82E-5*	5.30E-7*	6.91E-6*

V. RESULTS

For comparing the results and evaluate the proposed parameters, we have used statistical analysis which are explained in details in next section.

A. *Statistical Analysis*

In this study, we have used Kruskal-Wallis test to define the level of significance of our proposed features.

Kruskal-Wallis test is a nonparametric version of the classical one-way ANOVA, and an extension of the Wilcoxon rank sum test to more than two groups [2]. The assumption behind this test is that the measurements come from a continuous distribution, but not necessarily a normal distribution. The test is based on an analysis of variance using the ranks of the data values, not the data values them.

In our study, this test has been used to evaluate the hypothesis for *GOM* and *COM* separately, once for each two groups and then for all four groups as mentioned total in Table 1 and Table 2. The *p* values obtained from Kruskal-Wallis analysis are shown in Table 1for *COM* andin Table 2 for *GOM*.In case of $p< 0.05$ to be considered as significant, we can see that *COM* and *GOM* would show the significant difference between groups which *p* value is shown by '*' in Table 1 and Table 2.

The results show that *GOM* discriminate CHF from NSR by $p<3E-3$; AF from NSR by $p<7E-6$; CHF from arrhythmia by $p<8E-3$; CHF from AF by $p<7E-6$; and arrhythmia from AF by $p<5E-5$. Hence, this feature is able to discriminate this four groups by $p<6E-7$.

In the same way, *COM* discriminate CHF from NSR by $p<3E-3$; AF from NSR by $p<7E-6$; CHF from arrhythmia by $p<3E-2$; CHF from AF by $p<7E-6$; and arrhythmia from AF by $p<9E-6$.Hence, this feature is able to discriminate this four groups by $p<2E-7$.

B. *Comparison with SD1 and SD2*

In this section, we have used Kruskal-Wallis test for *SD1* and *SD2*, which *p* values are displayed in Table 3, to compare these results with those were obtained from proposed features *GOM* and *COM*. The comparison between Table 1, Table 2, and Table 3 show that *SD1* and *SD2* are able to distinguish arrhythmia from normal subjects whereas our novel proposed features not only discriminate arrhythmia from normal, but also they are able to differentiate different arrhythmia from each other.

Table 3. *p*-Value Results for Standard Descriptors

Groups	*Standard Descriptors*	
	SD1	*SD2*
NSR, CHF	0.0013	0.0595
NSR, CAST	0.0002	0.0596
NSR, AF	0.0131	0.0661
CHF, CAST	0.4347	0.4906
CHF, AF	0.0115	0.0169
CAST, AF	0.0028	0.0058
Total	4.6145E-5	0.0076

VI. DISCUSSION

The main motivation for using poincare plot is to visualize the variability of any time series signal [6]. In order to this qualitative approach, we proposed two novel features, GOM and COM, to extract temporal variations in a poincare plot. The proposed features have been able to demonstrate useful information about the temporal variation of the Poincare plot and unlike SD1 and SD2; they are not dependent to the kinds of arrhythmia and are able to classify different arrhythmia. For example, parameters such as U, UU, and OO were able to distinguish five of six subjects in this study. So in the future, they may be used as efficient features for temporal and point to point analysis of poincare plot.

REFERENCES

[1] N. Jafarnia Dabanloo, S. Moharreri, S. Parvaneh, A. M. Nasrabadi, "Application of Novel Mapping for Heart Rate Phase Space and Its Role in Cardiac Arrhythmia Diagnosis," 37th Annual Computers in Cardiology, IEEE Computer Society Press, 2010

[2] N. Jafarnia Dabanloo, S. Moharreri, S. Parvaneh, A. M. Nasrabadi, "New Representation of Heart Rate and Evaluation of Geometric Features Extracted From It," 37th Annual Computers in Cardiology, IEEE Computer Society Press, 2010

[3] J. Piskorski, P. Guzik, "Filtering Poincaré plots," Computational methods in science and technology, vol. 11, pp. 39-48, 2005.

[4] M. Brennan, M. Palaniswami, P. Kamen,"Do existing measures of Poincaré Plot geometry reflect nonlinear features of Heart Rate Variability?," IEEE transactions on biomedical engineering, vol. 48, pp. 1342-1347, 2001.

[5] M. Tulppo, T. Makikallio, T. Takala, T. Seppanen, H. Huikuri, "Quantitative beat-to-beat analysis of heart rate dynamics during exercise," American Journal of Physiology- Heart and Circulatory Physiology, vol. 271, p. H244, 1996.

[6] C. Karmakar, A. Khandoker, J. Gubbi, M. Palaniswami, "Complex Correlation Measure: a novel descriptor for Poincaré plot," BioMedical Engineering OnLine, vol. 8, p. 17, 2009.

[7] J. Piskorski, P. Guzik, "Geometry of the Poincaré plot of RR intervals and its asymmetry in healthy adults," Physiological measurement, vol. 28, p. 287, 2007.

[8] K. Hnatkova, X. Copie, A. Staunton, M. Malik, "Numeric processing of Lorenz plots of RR intervals from long-term ECGs: Comparison with time-domain measures of heart rate variability for risk stratification after myocardial infarction," Journal of Electrocardiology, vol. 28, pp. 74-80, 1995.

[9] C. Karmakar, A. Khandoker, J. Gubbi, M. Palaniswami, "Defining asymmetry in heart rate variability signals using a Poincaré plot," Physiological measurement, vol. 30, p. 1227, 2009.

[10] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, H. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," Circulation, vol. 101, p. e215, 2000.

[11] H. Huikuri, T. Makikallio, C. Peng, A. Goldberger, U. Hintze, M. Moller, "Fractal correlation properties of RR interval dynamics and mortality in patients with depressed left ventricular function after an acute myocardial infarction,"*Circulation,* vol. 101, p. 47, 2000.

[12] P. Kamen, H. Krum, A. Tonkin,"Poincare plot of heart rate variability allows quantitative display of parasympathetic nervous activity in humans,"*Clinical science,* vol. 91, pp. 201-208, 1996.