Multi-scale Recurrence Quantification Analysis of Heartbeat Interval Series in Healthy vs. Heart Failure Subjects

Chengyu Huo, Yi Lu
School of Physics and Electronic Engineering
Changshu Institute of Technology
Changshu 215500, China
hcyy@cslg.edu.cn

Xiaolin Huang, Hongxing Liu, Xinbao Ning
Institute for Biomedical Electronic Engineering, School of Electronic Science and Engineering
Nanjing University,
Nanjing 210093, China

Abstract—Recurrence plot is a useful analysis method for nonlinear time series, and has been widely used in studies of heart rate variability in recent years. In this paper, recurrence plot and corresponding quantification analysis were utilized to analyze the heart rate variability data from healthy people and congestive heart failure sufferers. It was found that the measures for standard recurrence quantification analysis failed to distinguish between the two groups. Therefore, a multi-scale recurrence quantification analysis method was proposed, which is able to implement the quantitative analysis for recurrence structures of coarse-grained heartbeat interval series on multiple time scales. Experiment results showed that, on certain time scales, the recurrence structures of abnormal heart rate fluctuation from subjects with congestive heart failure could be reflected more clearly by multi-scale RQA parameters.

Keywords—recurrence plot; heart rate variability; multi-scale; recurrence quantification analysis

I. INTRODUCTION

Heart rate variability (HRV) is the beat-to-beat fluctuations of heart rate (or the heartbeat interval). This variability is affected by many factors, and under the regulation of autonomic nervous (sympathetic and parasympathetic nervous) system. Consequently, the autonomic nervous related physiological and pathological states of the cardiac system can be reflected by HRV [1, 2]. The heart is a complex nonlinear dynamical system, which determined the high complexity of HRV signals generated by the cardiac system. Therefore, in recent years, researchers have utilized various types of nonlinear methods to study the HRV signals (RR interval series) in phase space of the dynamic system [3-7].

Recurrence plot (RP) [8] is a popular and useful nonlinear method in analysis of the periodic and chaotic characteristics of the time series. With this method, the internal structures of the time series can be revealed, and the prior knowledge about the similarity and predictability can be obtained. Based on the RP, researchers have further proposed the recurrence quantification analysis (RQA) method [9, 10], which can make quantitative measures of the complex structures in RPs, so that the internal properties of the nonlinear system can be quantitatively reflected. The length and stability of time series is not critical for RQA method [11]. Therefore, the method is commonly used for the analysis of certain physiological signals. Since the physiological system often tend to change in multiple time scales, long-term stationary physiological time series are usually hard to be acquired.

Marwan et al [12] has studied the HRV data of ventricular tachyarrhythmia (VT) patients with RP and RQA methods. It was found that, values of some RQA parameters during VT state were significantly different from those in the control state. This study demonstrated the potential value of RP and RQA methods in the analysis of HRV signals.

In addition, recent studies have shown that, the heart rates of human beings are regulated by complex factors on different levels. And these regulatory mechanisms play their roles through multiple feedback loops that contain different delays [13-15]. Accordingly, on different time scales, the RR interval time series are expected to exhibit different dynamic characteristics. Therefore, in our study, the multi-scale RQA analysis was proposed to investigate the autonomic nervous modulation of heart rate in healthy and congestive heart failure (CHF) states.

II. METHODS AND MATERIALS

A. Recurrence Quantification Analysis

As a nonlinear analysis method for time series, the RP method is based on the concept of phase space reconstruction, and thus can reflect certain regularity of the trajectories in phase space. Recurrences take place in phase space of a system. One-dimensional time series cannot be directly analyzed by RPs; therefore, reconstruction of the phase space should be implemented. The commonly used method for phase space reconstruction is delay embedding, which needs the proper embedding parameters - the embedding dimension $m$ and the time delay $\tau$. Given $\{x_n\}_{n=1}^N$ as the time series ($N$ is the length of the sequence), the reconstruction of phase space by delay embedding can be expressed as:

$$\overline{X}_n = (x_n, x_{n-1}, \ldots, x_{n-(m-1)\tau}) \in R^m$$  \hspace{1cm} (1)
where the vector \( \overrightarrow{X} \) represents the point in the \( m \)-dimensional phase space of the system. Consequently, the state of the dynamical system at the time moment \( n \) can be represented. Thus we can use the recurrence matrix for expression of the RP:

\[
R_{i,j}(\varepsilon) = \Theta(\varepsilon - \| \overrightarrow{X}_i - \overrightarrow{X}_j \|), \ i, j = 1, \ldots, N
\]  

where \( N \) is the points number of vector \( \overrightarrow{X} \), \( \| \cdot \| \) is the norm for evaluating the distance between the two vectors, and the \( \varepsilon \) is a threshold, which can determine whether or not the two vectors are close to each other. Finally, \( \Theta() \) is the Heaviside function (if \( x < 0 \), the \( \Theta(x) = 0 \), else the \( \Theta(x) = 1 \)). The recurrence matrix can be plotted with two different colors (e.g., if \( R_{i,j} = 1 \), a black point on coordinates \((i,j)\) is plotted, else a white point on coordinates \((i,j)\) is plotted) [16].

The RPs are mainly used for qualitative analysis. For quantitative analysis, several measures of complexity have been proposed. These measures can quantify the structures in RPs, such as the density of recurrence points, the diagonal line structures and the vertical line structures. They are known as recurrence quantification analysis (RQA). Specifically, commonly used RQA measures include recurrence rate, determinism, average diagonal line length, the longest diagonal line length, recurrence entropy, laminarity, trapping time and maximal length of the vertical lines [16].

In our study, a subset of the above mentioned RQA measures was mainly used:

- **Determinism (DET):**
  
  \[
  DET = \frac{\sum_{l=l_{\min}}^{l_{\max}} IP(l)}{\sum_{l=1}^{N} IP(l)}
  \]  

  where \( l \) is the length of the diagonal line, \( P(l) \) is the number of diagonal lines with length \( l \), and \( l_{\min} \) is a preset threshold value of the diagonal lines. The determinism is the ratio of recurrence points forming diagonal structures to all recurrence points in the RP. With this measure, the recursive degree of the attractors and trajectories in phase space can be reflected. The higher \( DET \) value indicates the stronger predictability, whereas the lower \( DET \) value indicates the stronger randomness [16].

- **Average diagonal line length (L):**
  
  \[
  L = \frac{\sum_{l=l_{\min}}^{N} l \cdot IP(l)}{\sum_{l=1}^{N} IP(l)}
  \]  

  where \( l \) is the length of the diagonal line, \( P(l) \) is the number of diagonal lines with length \( l \), and \( l_{\min} \) is a preset threshold value of the diagonal lines. The average diagonal line length (L) can be interpreted as an average time, during which the two segments of the phase space trajectories keep close to each other [16].

- **The longest diagonal line length (L_{\text{max}}):**
  
  \[
  L_{\text{max}} = \max(l_{i,j}^{N})
  \]  

  where \( N_i = \sum_{l=l_{\min}}^{l_{\max}} P(l) \) is the total number of diagonal lines in the RP. The longest diagonal line length, \( L_{\text{max}} \), can reflect the divergence speed of the trajectories in phase space. If the trajectory segments in phase space diverge slowly, the diagonal lines will be longer and vice versa [16].

- **Recurrence Entropy (ENTR):**
  
  \[
  \text{ENTR} = -\sum_{l=l_{\min}}^{l_{\max}} p(l) \ln p(l)
  \]

  where \( l_{\min} \) is a preset threshold value of the diagonal lines, and \( p(l) \) is a probability, which can be calculated as: \( p(l) = P(l)/N_i \). Here \( P(l) \) is the number of diagonal lines with length \( l \), and \( N_i \) is the total number of diagonal lines in the RP. The recurrence entropy (ENTR) is the Shannon entropy of the probability \( p(l) \). With this measure, the complexity of the diagonal line structures in the RP can be reflected [16].

- **Laminarity (LAM):**
  
  \[
  \text{LAM} = \frac{\sum_{v=v_{\min}}^{N} vP(v)}{\sum_{v=1}^{N} vP(v)}
  \]

  where \( v \) is the length of the vertical line, \( P(v) \) is the total number of vertical lines with length \( v \), and \( v_{\min} \) is a preset threshold value of the vertical lines. The laminarity (LAM) is the ratio of the recurrence points forming the vertical lines to all recurrence points in the RP. This measure can reflect the states of vertical structures (i.e. the laminar states in the system). In the RP, if the vertical line structures are more obvious, the measure \( LAM \) will increase, and vice versa [16].

- **Trapping time (TT):**
  
  \[
  \text{TT} = \frac{\sum_{v=v_{\min}}^{N} vP(v)}{\sum_{v=v_{\min}}^{\infty} P(v)}
  \]

  where the meanings of symbols \( v \), \( v_{\min} \) and \( P(v) \) are all the same as those in the definition of Laminarity (LAM). The trapping time (TT) is the average length of vertical structures in the RP. With this measure, the mean time that the system will abide at a specific state (or how long the state of the system will be trapped) can be estimated [16].

### B. Multi-scale Recurrence Quantification Analysis

If the original RR interval time series of length \( N \) is denoted as \( \{X_i\}_{i=1}^{N} \), the consecutive coarse-grained time series \( \{Y_j\} \) is then constructed as:

\[
Y_{j}^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} X_{i}, \quad 1 \leq j \leq \frac{N}{\tau}
\]
where the integer $\tau$ represents the coarse-graining scale factor. For the coarse-grained series, $\{y_{i}^{(r)}\}$, the length will be $1/\tau$ of the original time series. Therefore, when scale $\tau=1$, the coarse-grained time series $\{y_{i}^{(1)}\}$ is just the same as the original time series [17, 18].

Finally, the RPs of these coarse-grained time series are plotted respectively, and the RQA measures are calculated as functions of the scale factor $\tau$. So that the dynamic behavior of heart rate fluctuations on different coarse-grained time scales can be observed.

C. Materials

In this research, the multi-scale recurrence quantification analysis was applied to the heartbeat interval series derived from HRV signals of healthy subjects with normal sinus rhythms (NSR), and subjects with CHF. All of the databases used in this study were obtained from PhysioNet [19]. Specifically, the data of the NSR group was obtained from the database of MIT-BIH Normal Sinus Rhythm. A total number of 18 subjects are included in this database. Specifically, there are 5 healthy males and 13 healthy females; the age (mean±SD) is 34±8 years, ranging 20-50 years. The CHF group was obtained from the database of BIDMC Congestive Heart Failure. A total number of 15 subjects with severe congestive heart failure (NYHA classes 3-4) are included in this database. Specifically, there are 11 males and 4 females, aged 56±11 years (mean±SD), ranging 22–71 years. Series of 18 subjects are included in this database. Specifically, there are 5 healthy males and 13 healthy females; the age (mean±SD) is 34±8 years, ranging 20-50 years. The CHF group was obtained from the database of BIDMC Congestive Heart Failure. A total number of 15 subjects with severe congestive heart failure (NYHA classes 3-4) are included in this database. Specifically, there are 11 males and 4 females, aged 56±11 years (mean±SD), ranging 22–71 years. Series of 18 subjects are included in this database. Specifically, there are 5 healthy males and 13 healthy females; the age (mean±SD) is 34±8 years, ranging 20-50 years. The CHF group was obtained from the database of BIDMC Congestive Heart Failure. A total number of 15 subjects with severe congestive heart failure (NYHA classes 3-4) are included in this database. Specifically, there are 11 males and 4 females, aged 56±11 years (mean±SD), ranging 22–71 years.

III. RESULTS AND DISCUSSION

A. Results

Firstly, the RPs of the RR interval series from NSR group and CHF group were plotted within scale range $\tau = 1-10$. The embedding dimension $m = 6$ [20]; the embedding delay $T = 6$ [21]; the distance threshold $\epsilon = 0.7\sigma$ ($\sigma$ is the RMS value of the RR interval series) [22]. For concise expression, only a typical example of the RPs from a healthy subject and a CHF sufferer on original scale ($\tau=1$) and middle scale ($\tau=5$) was presented on Fig.1.

Further, the graphs of the 6 RQA measures ($DET$, $L$, $L_{\text{max}}$, $ENTR$, $LAM$ and $TT$) vs. scale factor were presented in Fig.2.

$T$-tests for these RQA measures on scale 1-10 were done between the two groups of data. Results were shown in Tab. 1.
TABLE I. T-test of coarse-grained multi-scale RQA results for RR interval series from NSR group and CHF group (P represents the significance and ‘n.s.’ means P>0.05 (not significant))

<table>
<thead>
<tr>
<th>Scale factor (\tau)</th>
<th>(P) values</th>
<th>DET</th>
<th>(L)</th>
<th>(L_{max})</th>
<th>ENTR</th>
<th>LAM</th>
<th>TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>2</td>
<td>n.s.</td>
<td>0.02</td>
<td>n.s.</td>
<td>0.03</td>
<td>n.s.</td>
<td>0.02</td>
<td>n.s.</td>
</tr>
<tr>
<td>3</td>
<td>0.02</td>
<td>0.006</td>
<td>n.s.</td>
<td>0.006</td>
<td>n.s.</td>
<td>0.004</td>
<td>n.s.</td>
</tr>
<tr>
<td>4</td>
<td>0.004</td>
<td>0.002</td>
<td>0.004</td>
<td>0.001</td>
<td>0.01</td>
<td>0.002</td>
<td>0.004</td>
</tr>
<tr>
<td>5</td>
<td>0.005</td>
<td>0.002</td>
<td>0.003</td>
<td>0.001</td>
<td>0.007</td>
<td>0.002</td>
<td>0.004</td>
</tr>
<tr>
<td>6</td>
<td>0.002</td>
<td>0.002</td>
<td>6e-04</td>
<td>9e-04</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>7</td>
<td>0.002</td>
<td>0.003</td>
<td>0.004</td>
<td>0.001</td>
<td>0.002</td>
<td>0.002</td>
<td>0.004</td>
</tr>
<tr>
<td>8</td>
<td>0.002</td>
<td>0.004</td>
<td>0.003</td>
<td>0.001</td>
<td>0.002</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>9</td>
<td>0.003</td>
<td>0.004</td>
<td>0.005</td>
<td>0.001</td>
<td>0.003</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>10</td>
<td>0.003</td>
<td>0.01</td>
<td>0.005</td>
<td>0.001</td>
<td>0.003</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

From Fig.2 and Tab.I, we found that multi-scale RQA measures of the two groups varied differently with the increasing of the time scales. On the original scale, for all the 6 measures, no significant differences could be found between NSR and CHF groups. However, when scale \(\tau \geq 4\), all RQA measures showed significant differences \((P<0.05)\) between healthy and CHF groups. In general, standard RQA failed to reflect the differences of heart rate fluctuation between the two groups and thus had some limitations. While with the multi-scale RQA, these differences could be successfully observed on different time scales.

In order to set up reference values for the six parameters, we further made the receiver-operating-characteristics (ROC) analysis between the NSR and CHF group. In the ROC analysis, the area under the ROC curve (AUC) is usually calculated for the evaluation of diagnostic accuracy: a high AUC value indicates a high diagnostic accuracy; while a low AUC value indicates that the diagnostic accuracy is low. The maximum value of AUC is 1.0, which indicates that all the data in this diagnostic test have been properly classified. While the minimum value of 0.5 indicates that the classification result is random, so that the method has no diagnostic value [23]. In this study, the AUC values of each parameter under different scales \((\tau=4-10)\) were shown in Tab.II.

TABLE II. AUC VALUES OF EACH PARAMETER UNDER DIFFERENT SCALES \((\tau=4-10)\)

<table>
<thead>
<tr>
<th>Scale factor (\tau)</th>
<th>DET</th>
<th>(L)</th>
<th>(L_{max})</th>
<th>ENTR</th>
<th>LAM</th>
<th>TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.76</td>
<td>0.77</td>
<td>0.74</td>
<td>0.77</td>
<td>0.76</td>
<td>0.77</td>
</tr>
<tr>
<td>5</td>
<td>0.74</td>
<td>0.77</td>
<td>0.78</td>
<td>0.78</td>
<td>0.74</td>
<td>0.76</td>
</tr>
<tr>
<td>6</td>
<td>0.79</td>
<td>0.81</td>
<td>0.83</td>
<td>0.81</td>
<td>0.80</td>
<td>0.79</td>
</tr>
<tr>
<td>7</td>
<td>0.73</td>
<td>0.77</td>
<td>0.81</td>
<td>0.74</td>
<td>0.74</td>
<td>0.73</td>
</tr>
<tr>
<td>8</td>
<td>0.72</td>
<td>0.77</td>
<td>0.79</td>
<td>0.74</td>
<td>0.74</td>
<td>0.71</td>
</tr>
<tr>
<td>9</td>
<td>0.72</td>
<td>0.76</td>
<td>0.81</td>
<td>0.74</td>
<td>0.73</td>
<td>0.72</td>
</tr>
<tr>
<td>10</td>
<td>0.71</td>
<td>0.75</td>
<td>0.80</td>
<td>0.72</td>
<td>0.72</td>
<td>0.69</td>
</tr>
</tbody>
</table>

Combined with Tab.I and Tab.II, we believe that the RQA measures under scale \(\tau=6\) are most efficient in the actual diagnosis classification. Therefore, on scale \(\tau=6\), we calculated the maximum sum of specificity and sensitivity by the ROC curve. On this point of the ROC curve, the corresponding value of the parameter was set as the reference threshold. The reference threshold of each parameter and the corresponding accuracy in diagnosis tests were shown in Tab.III.

TABLE III. THE REFERENCE THRESHOLD OF EACH PARAMETER AND THE CORRESPONDING ACCURACY IN DIAGNOSIS TESTS \((\text{DATA LENGTH } N=2520)\)

<table>
<thead>
<tr>
<th>DET</th>
<th>(L)</th>
<th>(L_{max})</th>
<th>ENTR</th>
<th>LAM</th>
<th>TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold</td>
<td>0.68</td>
<td>3.24</td>
<td>81.00</td>
<td>1.34</td>
<td>0.73</td>
</tr>
<tr>
<td>Accuracy</td>
<td>76%</td>
<td>82%</td>
<td>79%</td>
<td>82%</td>
<td>76%</td>
</tr>
</tbody>
</table>

In addition, to investigate the influence of the data length on the RQA measures, RR interval series with different lengths \(N=1260\) and 5040 were also calculated. Results showed that, the distribution of multi-scale RQA curves were similar. On scale \(\tau=6\), the accuracy for each parameter in diagnosis tests under different data lengths were presented in Tab.IV.

TABLE IV. THE ACCURACY FOR EACH PARAMETER IN DIAGNOSIS TESTS UNDER DIFFERENT DATA LENGTHS

<table>
<thead>
<tr>
<th>Data Length</th>
<th>DET</th>
<th>(L)</th>
<th>(L_{max})</th>
<th>ENTR</th>
<th>LAM</th>
<th>TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1260</td>
<td>76%</td>
<td>79%</td>
<td>79%</td>
<td>76%</td>
<td>76%</td>
<td>76%</td>
</tr>
<tr>
<td>2520</td>
<td>76%</td>
<td>82%</td>
<td>79%</td>
<td>82%</td>
<td>76%</td>
<td>79%</td>
</tr>
<tr>
<td>5040</td>
<td>79%</td>
<td>79%</td>
<td>79%</td>
<td>79%</td>
<td>79%</td>
<td>76%</td>
</tr>
</tbody>
</table>

Taking the data length \(N=2520\) (used in most of the previous analysis) as a reference, double length of data did not bring significant improvement in accuracy of classification, and the calculation speed was greatly decreased. When data length decreased to 1/2, the accuracy of classification was just slightly decreased. Corresponding \(t\)-tests also proved that, the RQA measures of the two groups still have significant differences, though the significance declined slightly.
B. Discussion

The parameter DET can reflect the recursive degree of the phase space trajectories. The higher DET value indicates the stronger determinism (predictability), whereas the lower DET value indicates the stronger randomness. The multi-scale RQA results show that, when scale $\varepsilon \leq 3$, DET values of CHF group are significantly higher than those of NSR group. It indicates that the determinism and predictability of heart rate regulation system of CHF sufferers is significantly stronger than that of normal subjects, reflecting the diminished capacity of autonomic nervous regulation.

The parameters $L$ and $L_{\text{max}}$, can reflect the divergence speed of the trajectories in phase space. The multi-scale RQA results show that, when scale $\varepsilon \leq 4$, the $L$ and $L_{\text{max}}$ values of NSR group are remarkably lower than those of CHF group, indicating the average time of the trajectory segments closing to each other in reconstructed phase space is shorter, i.e., the unpredictability and complexity of healthy heart rate regulation system is relatively high.

Compared to the RPs of heartbeat interval series from healthy subjects, the large black rectangle structures in RPs of CHF sufferers are significantly expanded. The corresponding RQA measure $LAM$ and $TT$ for the CHF group should also increase significantly. However, the $LAM$ and $TT$ values on original scale do not effectively reflect this change. With multi-scale RQA, the increase of $LAM$ and $TT$ values of CHF sufferers are significant reflected when $\varepsilon \leq 4$, which indicates the abnormal increase of the laminar phases in this pathological state.

In addition, compared with the RQA results of ventricular tachyarrhythmia (VT) patients in [12], there are still some differences in details. It was reported that the RQA parameter maximal length of the vertical lines was most powerful in discriminating pathological state of VT from normal healthy state, which indicates that maximal length of vertical structures in RPs is greatly affected by VT. While in our study, for the pathological state of CHF, the parameters $LAM$ and $TT$ showed significant differences with normal healthy state, which indicates that the CHF state has a greater impact on total number and average length of vertical lines in RPs. These differences imply that, the physical nature of different pathological states can be detailedly reflected by RP and RQA methods.

In practice, by the analysis of ROC curve, the reference threshold of these RQA parameters for distinguishing healthy and CHF state can be established. The accuracy for each parameter in diagnosis tests is relatively high. Furthermore, higher classification accuracy is expected with multi-parameter joint classification.

IV. CONCLUSION

In this paper, a multi-scale recurrence quantification analysis method was proposed. With this method, RR interval series of healthy subjects and subjects with CHF were analyzed. Intuitively, on the original scale, there are differences in structures of the RPs from RR interval series of healthy subjects and CHF sufferers, but these differences are not sufficient to be significantly reflected by the standard RQA measures. The multi-scale RQA contributes to the decomposition of complex heart rate fluctuations caused by various internal and external factors, so that the recurrence properties of the cardiac system can be studied on multiple time scales. Multi-scale RQA results show that, on certain time scales, the abnormal heart rate fluctuation of subjects with CHF can be reflected more clearly, and the results are not highly dependent on the data length. Thus it may be helpful to the auxiliary diagnosis of heart failure.

REFERENCES


