Multiscale Entropy and Poincaré Plot-based Analysis of Pulse Rate Variability and Heart Rate Variability of ICU Patients

Rohit Parasnis, Akshay Pawar, M. Manivannan
Biomedical Engineering Group, Department of Applied Mechanics
Indian Institute of Technology Madras
Chennai, India
rohityparasnis@gmail.com

Abstract—It is now known that multiscale entropy has the potential to distinguish certain pathological time series clearly and reliably from the corresponding healthy series. However, the implications of this parameter for Heart Rate Variability (HRV) have not been studied extensively. Also, as reported by other studies, the Poincaré plots of the R-R interval series of a human subject’s ECG signal (which too function as a measure of HRV) could be more useful than the time-domain and frequency domain parameters of HRV for certain applications. Although the Poincaré plots of healthy R-R interval series and corresponding PPG-based interval series have been investigated by many researchers, they do not seem to have been examined for unhealthy subjects. Our goal in this study has been to assess the extent to which PRV can substitute for HRV in the determination of these two nonlinear parameters. We perform multiscale entropy analysis (MSE) on Pulse Rate Variability (PRV) and HRV of 20 ICU patients. We also obtain Poincaré plots associated with PRV from four PPG-based techniques and those characterizing HRV from the standard R-R interval technique. We then compare the resulting PRV data sets with their HRV counterparts. We observe that none of the PPG-based methods displays a satisfactory statistical agreement or even an acceptable statistical correlation with the standard ECG-based technique. Hence we conclude that as of now, one cannot rely on PRV as a convenient alternative to estimate the MSE values and Poincaré plots of HRV. However, further investigation on the lines suggested in the paper might yield fruitful insights.

Keywords—Poincaré plots; multiscale entropy; heart rate variability, pulse rate variability

I. INTRODUCTION

The concept of multiscale entropy analysis (MSE) was introduced in [1] because traditional measures of complexity (such as single-scale sample entropy) take higher values for the pathological time series that are somewhat random in nature, than for corresponding healthy series displaying “long-range correlations” that actually make the underlying healthy processes more complex. Since MSE takes into account the multiple time scales intrinsic to healthy physiological processes, unlike single-scale entropy, the multiscale entropy analysis of human heartbeat interval series was found to reliably separate healthy human subjects, Congestive Heart Failure patients and Atrial Fibrillation patients [1]. Motivated by this observation, the authors of [1] investigated MSE further and described in [2] how this new metric of physiological complexity could also distinguish between young and elderly healthy human subjects in both waking and sleeping periods. Since then, several studies have confirmed that MSE, either in its original form or after appropriate modifications, has applications that are many and varied [3–9]. Of these, its application to Heart Rate Variability (HRV), or the multiscale entropy analysis of human heartbeat interval series, is especially significant, as can be gleaned from [1, 2].

Another useful nonlinear parameter of HRV is what is known as “Poincaré plots”. These plots have been used traditionally to assess HRV. Poincaré plot as a tool to estimate HRV stands out from time-domain and frequency-domain HRV parameters because it provides information on the beat-to-beat variations in the behaviour of the heart [10]. For this reason, many researchers have already compared the Poincaré plots characterizing PRV with those characterizing HRV. M bolanos et al. have analyzed the Poincaré plots of average HRV and average PRV. They found the two plots to be remarkably similar [11], Selvaraj et al. too in [12] reported a good match between the Poincaré plot of RR interval series and PPI (systolic peak-to-peak interval) series. Furthermore, Rong-Chao Peng et al. in [13] compared the Poincaré plots of PRV obtained from smartphone-based PPG with their HRV counterparts. Even they have reported similar observations.

However, although Pulse Rate Variability (PRV) is generally expected to function as a more convenient substitute for HRV, to the best of our knowledge, no significant studies have compared these two non-linear parameters of PRV with those of HRV as far as subjects with cardiovascular injuries or diseases are concerned. As for MSE, none of the 105 studies that were reviewed in [14] includes any entropy measure computed over multiple temporal scales. Even the more recent studies that have sought to compare PRV with HRV [15–21] do not include MSE in any form. As for Poincaré plots, to the best of our knowledge, none of the existing studies report any comparative analysis of the PRV and HRV of subjects suffering from cardiovascular diseases or injuries. Therefore, our goal was to compare these two nonlinear parameters of PRV with their HRV counterparts and assess whether they could be used as surrogates for the estimation of HRV of ICU patients.*

* We had initially undertaken the task of comparing the Poincaré plots of signals obtained from healthy subjects. However, this has already been investigated thoroughly in [13] along with the assessment of the PPG-based methods we had earlier planned to evaluate.
II. METHODOLOGY

A. Data Acquisition

All of the analysis was performed on ECG and PPG signals of twenty ICU patients. The entire data was obtained from the MIMIC II Waveform Database accessible from the website of PhysioNet [22]. For every subject, synchronous signal segments corresponding to a selected window of 5 minutes duration were picked from both ECG and PPG signals. The criterion for window selection was the requirement of minimal signal distortion arising from motion artifacts, low-frequency baseline wander and high-frequency noise.

B. Preprocessing of ECG and PPG signals

Every PPG signal was smoothed using a third-order Savitzky-Golay filter having a frame-size of 49. It was further smoothed and de-noised using a smoothing function and wavelet decomposition techniques.

As for ECG signals, each of them was directly processed by a MATLAB code implementing Pan-TomPkin’s Algorithm, an algorithm used for the detection of R-wave peaks [23]. This algorithm uses a variety of preprocessing and filtering techniques in order to make the R-waves detectable.

C. Computation of Sample Entropy for MSE

The sample entropy, $SampEn(m,r,N)$ of a time series of length $N$ can be defined as the negative logarithm of a conditional probability: the probability that if the distance between two distinct $m$-length sequences (each of which has $m$ consecutive data points picked from the series) is less than $r$, it remains less than $r$ even when one more point is added to each sequence [1, 2]. The parameter $m$ is chosen to be 1 or 2 a priori in most cases, whereas $r$ is usually taken to be a fraction of the standard deviation, $\sigma$, of the time series. In most cases, this fraction is chosen to be between 0.1 to 0.25 [24]. For our purposes, the following values were used: $m = 2$ and $r = 0.2\sigma$.

The procedure that is usually followed for the computation of $SampEn(m,r,N)$ of a series $X=\{x_1, x_2, ..., x_N\}$ follows directly from the definition of sample entropy. The steps of this procedure are enumerated below [1, 2]:

- List all possible sequences of the form $\{x_i, x_{i+1}, x_{i+2}, ..., x_{i+m}\}$ where $i$ is an index such that $1 \leq i < i+m \leq N$. Clearly, there are $N-m+1$ such sequences.
- Compute the Chebyshev distance of the $j^{th}$ such sequence from every other sequence obtained in step 1. This is done as follows: treating both the sequences (the $j^{th}$ sequence and the other sequence) as vectors, say $u=\{u_1, u_2, ..., u_m\}$ and $v=\{v_1, v_2, ..., v_m\}$, find the maximum absolute difference between their vector components: $d(u,v) = \max[|u_k - v_k|]$. Now take $d(u,v)$ as the distance between the sequences.
- Count the number of such distances that are less than $r$ (or the number of sequences whose distance from the $j^{th}$ sequence evaluates to a value less than $r$). Let us denote this number by $n_j$. Repeat steps 2-3 for all $j \in \{1,2,...,N-m+1\}$.
- Repeat all the above steps for $m+1$ in place of $m$, so as to compute $n_{j,m+1}$ for every $j$ in $\{1,2,...,N-m+1\}$.
- Determine the sample entropy using the following formula:

$$SampEn(m,r,N) = \log \frac{\sum_{j=1}^{N-m+1} n_j}{\sum_{j=1}^{N-m+1} n_{j,m+1}}$$ (1)

In this study, the MATLAB code provided in [25], which executes the above procedure after taking the values of $m$ and $r$ as inputs, was used for computing all the sample entropies required for the multiscale entropy analysis.

D. Multiscale Entropy Analysis

The “multiscale entropy analysis” of a time series involves the following steps [1, 2]:

- The first step is the selection of a range of “scales factors”, say $\{m, m+1, ..., m+n-1\}$, from the set of natural numbers. The range selected for our study was $\{2, 3, ..., 41\}$ (1 was excluded because it corresponds to the already researched single-scale sample entropy).
- For every scale factor $\tau$ belonging to this range, a time series $X^\tau$ is generated from the original series $X$ by using the following equation:

$$X^\tau_j = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j} x_i, \quad 1 \leq j \leq N/\tau$$ (2)

Thus, every point in $X^\tau$ is the arithmetic mean of the corresponding $\tau$ consecutive points of $X$. This procedure of calculating $X^\tau$ from $X$ is called “coarse-graining”.
- For each scale factor $\tau$, the sample entropy, $SampEn(m,\tau,N/\tau)$ of $X^\tau$ is determined.*
- A plot of $SampEn(m,\tau,N/\tau)$ vs $\tau$ is generated. This plot shows the variation of the sample entropy of the time series $X$ over the chosen range of temporal scales.
- Finally, this plot is analyzed suitably in order to draw useful inferences and conclusions.

However, when coarse-graining is applied to short-term series, it causes some of the resulting values of MSE to be undefined [26]. Because the signals used in this study were of a relatively short duration (5 minutes), the resulting interval series turned out to be short-term series (consisting of 400 – 700 data points). Thus, coarse-graining led to the same problem in our study. Hence, the process of coarse-graining was replaced with that of a moving-average filter algorithm as

* In this study, multiscale entropy analysis was initially performed with approximate entropy (ApEn) in place of sample entropy. However, the metric of sample entropy was proposed as an improvement to approximate entropy [32]. Therefore, multiscale approximate entropy analysis is hardly performed on HRV, unlike multiscale sample entropy analysis. We thus decided to eliminate the former from our discussion.
proposed in [26], which defines $X^{(r)}$ as follows:

$$X_{j}^{(r)} = \frac{1}{\tau} \sum_{i=j}^{i+j-1} x_i, \quad 1 \leq j \leq N - \tau + 1 \quad (3)$$

Therefore, to be precise, our analysis should be called moving-average filter-based multiscale entropy analysis. For the sake of simplicity, however, we shall continue referring to it by the term “MSE”. The entire procedure was executed in MATLAB using appropriate functions and scripts.

E. Poincaré Plots

The Poincaré analysis of a time series generally gives us a lot of information about the variability of the time series data [27]. In the context of HRV, a Poincaré plot is a plot of the series of R-R intervals of the subject’s ECG vs the series of the immediately preceding R-R intervals i.e. a plot of RR(n) vs RR(n-1) where n is the index of the series and ranges from 2 to the total number of points present in the R-R interval series. When this concept is extended to the case of PRV, the series of the current systolic peak-to-peak intervals of the subject’s PPG signal is plotted against the series of the preceding systolic peak-to-peak intervals. In our study, we have also analyzed the Poincaré plots of the following interval series for each of the 20 subjects:

- Peak-to-peak intervals of the 1st derivative of the PPG signal
- Peak-to-peak intervals of the 2nd derivative of PPG
- Valley point-to-valley point intervals of PPG

Each of the above can be considered to be possible alternative methods of generating Poincaré plots characterizing HRV. For this reason, these four methods (including the systolic peak-to-peak interval method) were compared with the standard R-R interval method to determine which of these is the best suited for the generation of Poincaré plots of HRV.

Poincaré plots are often characterized by certain descriptors called $SD1$ and $SD2$ defined below [28]:

- $SD1$ is the standard deviation of the projection of the Poincaré plot on the line perpendicular to the line of identity (the line $y = -x$)
- $SD2$ is the standard deviation of the projection of the Poincaré plot on the line of identity itself ($y = x$).

In order to simplify our calculations, we use the following formulae proposed by Piskorski in [29]:

$$SD1 = \sqrt{Var \left( \frac{x^+ - x^-}{\sqrt{2}} \right)}, \quad SD2 = \sqrt{Var \left( \frac{x^+ + x^-}{\sqrt{2}} \right)} \quad (4)$$

where $x^+ = (x_2, x_3, \ldots, x_N)$ is the original time series $x$ with the last data point eliminated whereas $x^- = (x_2, x_3, \ldots, x_N)$ is the original series with the first point removed.

F. Methods of Comparison

For the purpose of assessing the overall statistical closeness between the data sets of HRV and PRV (which include both MSE values and Poincaré Plot descriptors), we examined the statistical correlation between these two data sets using the traditionally used Pearson correlation coefficient $\rho$, as well as the statistical agreement between them using the “concordance correlation coefficient” proposed by Lin in [30].

The concordance correlation coefficient of two data sets $X_1$ and $X_2$ is defined as follows:

$$\rho_c = \frac{2\sigma_{12}}{\sigma_1^2 + \sigma_2^2 + (\mu_1 - \mu_2)^2} \quad (5)$$

where $\sigma_1, \sigma_2, \mu_1$ and $\mu_2$ are the standard deviations and the arithmetic means of $X_1$ and $X_2$ respectively [30]. This measure provides a suitable judgment of the degree of the statistical agreement between two data sets resulting from linear correlation between them [14]. Pearson correlation coefficient is well-known and hence its definition is not reproduced here.

For each pair of data sets, Pearson correlation coefficients were computed in MATLAB using an appropriate function. All the concordance correlation coefficients were calculated using an online statistical calculator [31].

As a further check, all of the plots characterizing PRV were compared with the corresponding HRV plots by visual inspection.

III. RESULTS AND DISCUSSION

A. Comparison of MSE Results of HRV and PRV

The values resulting from the multiscale entropy analysis of R-R intervals, corresponding to the scale factors 2-41, were compared with the corresponding MSE values of the corresponding systolic peak-to-peak interval series. The values of the Pearson correlation coefficients pertaining to these comparisons for scales up to 15 are tabulated below.

<table>
<thead>
<tr>
<th>Temporal Scale Factor</th>
<th>Coefficient Value ($\rho$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.865492</td>
</tr>
<tr>
<td>3</td>
<td>0.8947</td>
</tr>
<tr>
<td>4</td>
<td>0.664903</td>
</tr>
<tr>
<td>5</td>
<td>0.626702</td>
</tr>
<tr>
<td>6</td>
<td>0.555996</td>
</tr>
<tr>
<td>7</td>
<td>0.614456</td>
</tr>
<tr>
<td>8</td>
<td>0.536038</td>
</tr>
<tr>
<td>9</td>
<td>0.53933</td>
</tr>
<tr>
<td>10</td>
<td>0.463935</td>
</tr>
<tr>
<td>11</td>
<td>0.463947</td>
</tr>
<tr>
<td>12</td>
<td>0.421157</td>
</tr>
<tr>
<td>13</td>
<td>0.427026</td>
</tr>
<tr>
<td>14</td>
<td>0.371537</td>
</tr>
<tr>
<td>15</td>
<td>0.35125</td>
</tr>
</tbody>
</table>

Likewise, the values of the concordance correlation coefficient were computed for each of the HRV-PRV pairs of MSE data sets. They are tabulated on the next page.
TABLE II. CONCORDANCE CORRELATION COEFFICIENTS BETWEEN MSE VALUES OF HRV AND PRV

<table>
<thead>
<tr>
<th>Temporal Scale Factor</th>
<th>Coefficient Value ($\rho$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.5288</td>
</tr>
<tr>
<td>3</td>
<td>0.2413</td>
</tr>
<tr>
<td>4</td>
<td>0.1803</td>
</tr>
<tr>
<td>5</td>
<td>0.4712</td>
</tr>
<tr>
<td>6</td>
<td>0.4408</td>
</tr>
<tr>
<td>7</td>
<td>0.5029</td>
</tr>
<tr>
<td>8</td>
<td>0.2559</td>
</tr>
<tr>
<td>9</td>
<td>0.4198</td>
</tr>
<tr>
<td>10</td>
<td>0.4658</td>
</tr>
<tr>
<td>11</td>
<td>0.4217</td>
</tr>
<tr>
<td>12</td>
<td>0.3327</td>
</tr>
<tr>
<td>13</td>
<td>0.4303</td>
</tr>
<tr>
<td>14</td>
<td>0.3821</td>
</tr>
<tr>
<td>15</td>
<td>0.4144</td>
</tr>
</tbody>
</table>

Table I indicates that the statistical correlation between the MSE values of HRV and PRV is far from satisfactory ($\rho$ is significantly less than 0.85). Similarly, Table II strongly implies that the statistical agreement between these two data sets is very poor ($\rho \leq 0.5$ at most scales). This clearly shows that the MSE analysis of systolic peak-to-peak interval series cannot replace that of R-R interval series for the purpose of estimating HRV. This observation is reiterated by the following example plots that result from our MSE analyses i.e. they show the variation of sample entropy with variation in temporal scale for one of the twenty cases that were examined.

Just like the above pair of plots, most of the nineteen other pairs too show significant differences between the multiscale entropies of HRV and PRV.

B. Comparison of Poincaré Plots

The values of the three Poincaré descriptors of R-R interval series, namely $SD_1$, $SD_2$ and the ratio $SD_1/SD_2$, were compared with their corresponding values of the four PPG-based interval series. The resulting values of Pearson correlation coefficients corresponding to these comparisons are tabulated below.

TABLE III. PEARSON CORRELATION COEFFICIENTS BETWEEN Poincaré DESCRIPTORS OF HRV AND PRV

<table>
<thead>
<tr>
<th>PPG-based Method</th>
<th>$SD_1$</th>
<th>$SD_2$</th>
<th>$SD_1/SD_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systolic Peak-Peak Interval Series</td>
<td>0.6736</td>
<td>0.4302</td>
<td>0.5014</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; Derivative Peak-Peak Interval Series</td>
<td>0.6545</td>
<td>0.5341</td>
<td>-0.0408</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; Derivative Peak-Peak Interval Series</td>
<td>0.4339</td>
<td>0.4105</td>
<td>-0.0132</td>
</tr>
<tr>
<td>Valley point-valley point Interval Series</td>
<td>0.1816</td>
<td>0.0781</td>
<td>-0.0503</td>
</tr>
</tbody>
</table>

The values of the concordance correlation coefficients pertaining to the above comparisons are tabulated below.

TABLE IV. CONCORDANCE CORRELATION COEFFICIENTS BETWEEN Poincaré DESCRIPTORS OF HRV AND PRV

<table>
<thead>
<tr>
<th>PPG-based Method</th>
<th>$SD_1$</th>
<th>$SD_2$</th>
<th>$SD_1/SD_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systolic Peak-Peak Interval Series</td>
<td>0.6636</td>
<td>0.4136</td>
<td>0.4821</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; Derivative Peak-Peak Interval Series</td>
<td>0.197</td>
<td>0.1837</td>
<td>-0.029</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; Derivative Peak-Peak Interval Series</td>
<td>0.0733</td>
<td>0.0789</td>
<td>-0.006</td>
</tr>
<tr>
<td>Valley point-valley point Interval Series</td>
<td>0.0606</td>
<td>0.0353</td>
<td>-0.047</td>
</tr>
</tbody>
</table>
Both these tables indicate that none of the PPG-based methods shows a satisfactory statistical correlation ($\rho \geq 0.85$) or a satisfactory statistical agreement ($\rho_c \geq 0.85$) with the standard R-R interval technique. This strongly implies that none of the PPG-based methods (that directly characterize PRV) can act as a surrogate for estimating the Poincaré descriptors of HRV. This is further confirmed by the Poincaré plots themselves, as can be seen from the example plots on the next page.

A quick inspection shows that none of the plots in Fig. 4 – Fig. 7 has a scatter even remotely similar to that of the ECG-based plot shown in Fig. 3. This reinforces our inference that no PPG-based method is reliable for the generation of Poincaré plots of HRV.

CONCLUSION

In this study, we compared two nonlinear parameters of PRV of ICU patients with their HRV counterparts: moving-average filter-based multiscale entropy and Poincaré plots and their descriptors. Our results show that neither of these parameters exhibits a satisfactory statistical agreement or even a reasonably good statistical correlation between HRV and PRV. The results regarding Poincaré plots are remarkably surprising because previous studies (such as [13]) have shown that as far as healthy subjects are concerned, the Poincare Plots of PRV are a useful indicator of HRV. This drastic change in the suitability of PRV for the estimation of HRV, due to a shift from healthy subjects to ICU patients, throws up a fundamental question: Could cardiovascular injuries and related disorders be responsible for altering the behavior of pulse rate in one way and that of heart rate in another? Pursuing this line of investigation could lead to novel insights into the nature of some broad categories of diseases that generally necessitate ICU treatment.

ICIIBMS 2015, T3A-20954, Okinawa, Japan
However, our observations allow us to safely conclude that further research is required before we can rely on PRV as a surrogate for the two crucial nonlinear parameters of HRV studied in this paper. As an example, one aspect that warrants attention is the optimal selection of the window length $m$ and the tolerance value $r$ associated with MSE. This can help minimize the relative error of MSE values. We also need to determine the values of $m$ and $r$ which could yield a reasonable agreement between the MSE data sets of PRV and HRV. We are currently working on these aspects and hope to acquire promising results in the near future.

REFERENCES


