Robust Time Series Processing for Heart Rate Variability Analysis in Daily Life

L Y Ji, Y J Yang, A G Li, S F Wang, J K Wu

Graduate University of Chinese Academy of Sciences, Beijing, China

Abstract

Heart Rate Variability (HRV) analysis can be heavily affected by slow trends and outliers in inter-beat interval data, especially when recorded during exercise. In this paper, a trend and outlier removal method is presented. Inter-beat interval data are considered as time series and decomposed into trend, outlier and fluctuation components using an iterated algorithm. Trend components reflect heart rate variation due to activity or exercise and are estimated based on Empirical Mode Decomposition (EMD). Outlier components reflect noise and other abrupt interference and are estimated using non-parameter method. The fluctuation components are obtained by removing trend and outlier components from original heart rate data, and thus have pure heart rate variation caused by autonomous nervous system function and are used for HRV analysis. This iterative method can handle heart rate data contaminated by large percentage of outliers and slow trends simultaneously, which is usually the case during exercise.

1. Introduction

Heart rate variability (HRV) is a property that beat to beat time intervals of heart rate vary over time. HRV can be used to evaluate function of the autonomic nervous system (ANS), which has a significant relationship with cardiovascular mortality [1]. For example, reduction of HRV is a predictor of mortality after myocardial infarction (MI), and also related to diabetic neuropathy, cardiac transplantation, myocardial dysfunction, tetraplegia, renal failure, drugs, etc [2].

One of difficulties to utilize HRV is that it can be affected by several factors such as activity, emotion, sex, and age. Among all the non-disease factors affecting HRV, activity is one of the biggest. To avoid interference of activity, HRV is usually measured at rest. Sometimes light intensity activity is allowed for long-term data analysis. However, more and more heart monitor data are recorded in people's daily life, where people are more likely doing activities with various intensity. Thus, HRV analysis on these dynamic data is valuable and unavoidable.

Unfortunately, sparse experiments are conducted to draw reliable conclusion about HRV during exercise [3], and results are heterogeneous. HRV of children during light exercise are not reliable compared with HRV at rest [4]. It is concluded that other factors other than ANS dominate the HRV during heavy exercise [5]. Same conclusion was drawn for persons with down syndrome during exercise accompanied findings that all HRV variables decreased significantly [6]. In older adults, higher vagal-related HRV indices are associated with moderate low-energy expenditure compared with sedentary subjects [7]. With high frequency (HF) increased and LF/HF ratio decreased, it is believed early exercise training improves sympathovagal balance and may reduce the arrhythmogenic substrate following MI [8]. We could see that HRV during exercise is sensitive and not stable. One reason of this is that there are compound factors such as heart rate variation due to activity. Another reason is more errors in RR interval data. Although there are detrend and outlier removal operation in HRV analysis [9], it is not realized that this is more important for HRV during exercise. Current methods are not designed for middle, intensive or random exercise and thus not robust enough. Most of time visual examination and manual correction are needed even when applying to not so intensive activity scenarios [10].

In this paper, our primary goal is to investigate reliable HRV analysis under exercise while people are taking daily life exercise. By proposing a robust trend, outlier detection and removal method for RR interval series, we are trying to remove the variation and artifacts caused by exercise. With our approach, HRV indices during exercise show more reliable relationship with HRV at rest.

2. Methods

When people are doing exercise, heart rate will increase to a target heart rate and drift higher if the exercise continued. If exercise stops or intensity of exercise lowers, the heart rate will drop down. The rising and dropping of heart rate due to exercise are called heart rate kinetics. The kinetics show as the long term trend of heart rate data. Thus, the variation of heart rate due to exercise will mixed up with variation due to "background noise" caused by activity of ANS, which is the main part of HRV at rest.

The "mix-up" effect influence the HRV analysis especially when the intensity of exercise is not stationary. Therefore, the heart rate variation due to exercise should be removed from the heart rate data before conducting HRV analysis. Because the heart rate kinetics demonstrates itself as a long term trend, a trend estimation and removal method based on Empirical Mode Decomposition (EMD) will be introduced to do this. By selecting proper intrinsic mode functions (IMF) of EMD, one can reconstruct the trend related to exercise in heart rate data. Because higher IMFs in EMD has lower frequency component, trend can be defined as

$$trend = \sum_{m^* \leq k \leq M} (IMF(k)) \qquad (1)$$

where m^* is the lowest order IMF which make the scale of the estimated $trend$ smaller than a predefined scale. Thus, we can initialize $trend$ as residue of EMD denoted by $residue_{EMD}$, and iteratively add lower IMFs to it until its scale is larger than the predefined scale. Approximately, intervals between two successive extrema or cross-zeros can be used to represent the scale of trend. In our HRV analysis application, definitions of HRV indices during exercise are the same as at rest. Therefore, the maximum interval between two successive cross-zero's of trend of heart rate data at rest, denoted by czi_{rest}, is used as the predefined scale of trend of heart rate data during exercise.

$$scale\_of\_trend_{during\ exercise} = czi_{rest} \qquad (2)$$

where czi_{rest} is interval between two successive points where heart rate data at rest cross its mean value.

Also, there are more outliers in ambulatory heart rate data due to the noise and artifacts when recorded during exercise. The outliers will affect trend removal and HRV analysis. Vice visa, the trend estimation will affect detection of the outliers. Therefore, we have to insert the outlier detection and correction procedure into every iteration of trend estimation. The complete algorithm is defined in Alogrithm 1.

The outlier detection in Algorithm 1 is conducted by setting a numerical threshold for the distance between candidate data and pattern trend. If the distance of one sample of the heart rate series is outside the following interval, then it is determined as an outlier.

$$[Q_1 - 1.5IQR, Q_3 + 1.5IQR] \qquad (3)$$

In above formula,

$$
\begin{aligned}
Q_1 &= median(\text{lowest 25\% of data}) \\
IQR &= Q_3 - Q_1 \qquad (4) \\
Q_3 &= median(\text{highest 25\% of data})
\end{aligned}
$$

This method is utilized because it is more robust and makes no assumption about the distribution of time series.

Algorithm 1 Detrend and Outlier Removal

1:   Compute czi_{rest} of heart rate data at rest
2:   Compute EMD of signal and obtain M order IMFs.
3:   Initialize $trend = residue_{EMD}$ and $m = M$.
4:   Compute largest cross-zero intervals of trend czi_{trend},
5:   **if** $czi_{trend} > czi_{rest}$ **then**
6:      trend = trend + IMF(m)
7:      $m = m - 1$;
8:      Go to step 4
9:   **else**
10:      Go to step 12
11:   **end if**
12:   Compute fluctuation = RR intervals - trend
13:   Detect outliers of fluctuation
14:   **if** number of outliers > 0 **then**
15:      Go to step 19
16:   **else**
17:      **return** m and fluctuation
18:   **end if**
19:   Replace outliers by interpolating nearest non-outlier heart rate data.
20:   Go to step 2.

After applying Algorithm 1, the heart rate data are actually decomposed into trend, outlier and fluctuation component. The fluctuation component are RR series without slow trend and outliers, and can be used to do HRV analysis directly.

3. Results and discussion

3.1. Experiment

The ambulatory electrocardiography(ECG) was recorded using our own uCare ECG recorder. It is a Holter-like 3-lead ECG recorder with 3-axis accelerometer. The device can be worn by the subjects around the chest. The ECG data and acceleration data are stored on microSD card. The sampling rate of ECG is 300Hz, while the sampling rate of acceleration is 100Hz. Both ECG and acceleration data have 8-bit resolution.

Ten young healthy subjects (aged 25±5; 8 male, 2 female) were asked to do the experiments wearing the uCare device. The possible risks and benefits of the tests were fully explained to the subjects before they gave their consent. The subjects were asked to sitting down for 30minutes with body relaxed but free for the hands movements. Without the removal the devices, the subjects are asked to do the jogging with middle exercise for another 30 minutes. The intensity was not rigorously controlled, which made the data more like being recorded from daily life exercise. The place for sitting is indoor with temperature of

$23 \sim 26°C$. The place for jogging is outdoor with temperature of $27 \sim 31°C$.

Both ECG at rest (sitting) and during exercise (jogging) was processed using the same program to obtain heart rate data. The processing program was a wavelet based ECG annotation method in [11]. The data recorded was divided into segments at rest and during exercise manually according to the acceleration data. After processing the heart rate data during exercise using the proposed method, three groups of heart rate data are obtained: the data at rest, the raw data during exercise, and the processed data during exercise. Then, time domain and frequency domain HRV indices are computed on these three groups of data. Specific parameters of the computation are as follows: the standard deviation of NN interval (SDNN) and standard deviation of average NN interval means (RMSSD) and percentage of differences between adjacent NN intervals that are greater than some milliseconds (pNNx) were computed using 30 minutes long data. The standard deviation of average NN interval means (SDANN) was computed by dividing data into multiple 5-minute long segments. A 50-millisecond threshold was used in pNNx. For the frequency domain indices, Power Spectral Density (PSD) was computed using Welch method on the 30 minutes data, where RR interval data were resampled using 2Hz sampling rate. Four frequency indices are proportions of very low frequency, low frequency, and high frequency from PSD, i.e., pVLF, pLF, pHF, and ratio of low frequency to high frequency, namely, LF/HF. Note it that the 30 minutes data include the beginning and recovery phase of heart rate.

3.2. Results and disscussion

The outlier and trend removal effect can be illustrated by processing results of subject No.6 shown in Fig. 1.

The time domain HRV analysis results are shown in Fig. 2, from which we could see that with our approach, all the HRV indices during exercise show more reliable relationship with HRV indices at rest. The relationship is that time domain HRV indices during exercise is smaller than at rest. This is due to that our approach removes the variation and outliers caused by exercise interference, which may make HRV measurement during exercise wrongly bigger than real value.

The reduction of time domain HRV indices agreed with the increase of activity of sympathetic nervous system during exercise. HRV at rest is controlled by parasympathetic/vagus nervous system, while sympathetic nervous system is in dominant position while the activity of parasympathetic/vagus nervous system drops down relatively during exercise [12].

Frequency domain HRV indices comparison is shown in Fig. 3. With our approach, the dominant changes of HRV during exercise are clearer. For example, the proportion of
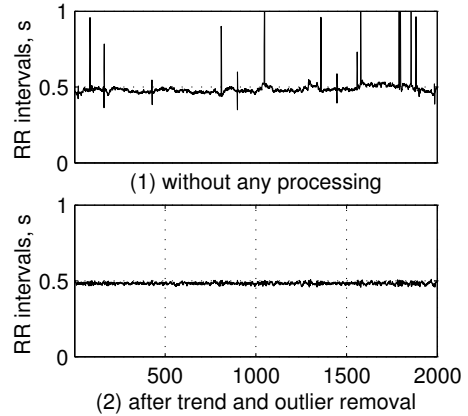

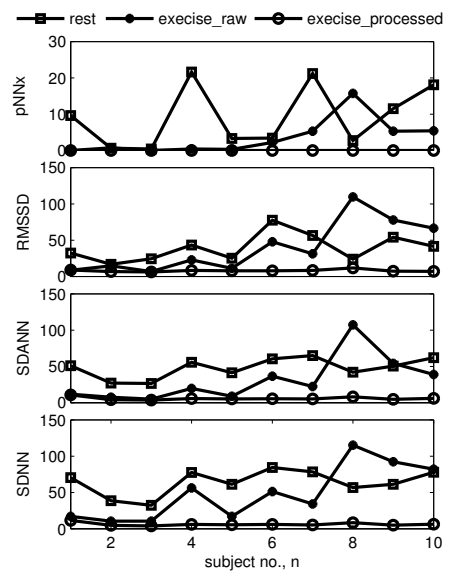Figure 1. Trend and outlier removal results in RR intervals of subject No.6

Figure 2. Time domain indices comparison at rest and during exercise

HF components increases are easier to observe.

However, the frequency domain indices during exercise show relatively obscure relationship with the one at rest compared with time domain indices. This is because details of PSD can change dramatically during exercise, which makes the proportion of different frequency change more complexly. The PSD pattern change can be illustrated by one example of the PSD of heart rate data shown in Fig. 4.

As we know, the low frequency domain is relevant to sympathetic while the high frequency domain is relevant to parasympathetic. What the PSD reflects is the degree of ANS regulatory but not the activity of ANS. During exercise, although the activity drops, the regulatory degree of
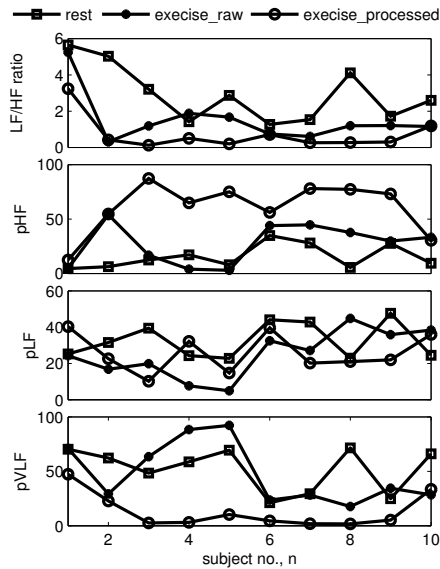
Figure 3. Frequency domain indices at rest and during exercise comparison
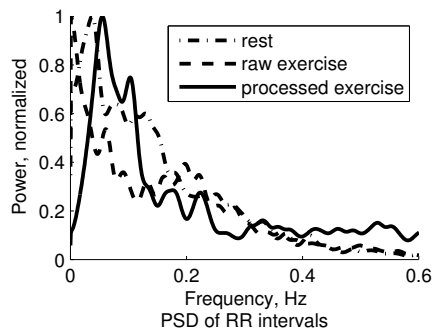

Figure 4. Power spectral density of RR intervals at rest and during exercise of subject No.2

parasympathetic nervous system enhanced. So, the power of HF is bigger while the power of LF is smaller [12].

4. Conclusions

In this paper, we have investigated making HRV analysis during exercise more reliable by introducing robust trend and outlier removal method. The more stable relationship between HRV at rest and during exercise than the traditional method can at least verify that the trend and outlier removal method is one valid way to remove HRV-irrelevant factors caused by exercise or activity.

Acknowledgements

References

[1] Heart rate variability-standard of measurements, physiological interpretation and clinical use. European Heart Journal 1996;17:354–381.

[2] Acharya UR, Joseph KP, Kannathal N, Lim CM, Suri JS. Heart rate variability: a review. Medical and Biological Engineering and Computing 2006;44:1031–1051.

[3] Brodie DA, Hodges LD, Sandercock GRH. Exercise, fitness and health: research strategies for a changing society. Journal of Exercise Science and Fitness 2004;2(2):85–93.

[4] Winsley RJ, Armstrong N, Bywater K, Fawkner SG. Reliability of heart rate variability measures at rest and during light exercise in children. Br J Sports Med 2003;37(6):550–552.

[5] Casties JF, Mottet D, Gallais DL. Non-linear analyses of heart rate variability during heavy exercise and recovery in cyclists. Int J Sports Med 2006;27(10):780–785.

[6] Baynard T, Pitetti KH, Guerra M, Fernhall B. Heart rate variability at rest and during exercise in persons with down syndrome. Archives of Physical Medicine and Rehabilitation 2004;85(8):1285–1290.

[7] Buchheit M, Simon C, Charloux A, Doutreleau S abd Piquard F, Brandenberger G. Heart rate variability and intensity of habitual physical activity in middle-aged persons. Med Sci Sports Exerc 2005;37(9):1530–1534.

[8] Russoniello C, Mahar M, DiNallo J, McCammon M, Skalko T, Rowe D. Effects of a physical activity program on heart rate variability in obese children. Applied Psychophysiology and Biofeedback 2002;27(4):299–320.

[9] Tarvainen M, Ranta-aho P, Karjalainen P. An advanced detrending method with application to hrv analysis. IEEE Trans Biomed Eng 2001;49(2):172–175.

[10] Buccelletti F, Gilardi E, Scaini E, Galiuto L, Persiani R, Biondi A, Basile F, Gentiloni Silveri N. Heart rate variability and myocardial infarction: systematic literature review and metanalysis. European Review for Medical and Pharmacological Sciences 2009;13:299–307.

[11] Chesnokov Y, Nerukh D, Glen R. Individually adaptable automatic qt detector. Computers in Cardiology 2006; 33:337–341.

[12] De La Cruz Torres B, López López C, Naranjo Orellana J. Analysis of heart rate variability at rest and during aerobic exercise: a study in healthy people and cardiac patients. Br J Sports Med 2008;2(9):715–20.

Address for correspondence:

Lianying Ji
Room 212, Keyan Building, Yuquan Road 19A, Graduate University of Chinese Academy of Sciences, Shijingshan District, Beijing, China
jilianying@gucas.ac.cn