Combination of equiprobable symbolization and time reversal asymmetry for heartbeat interval series analysis

Fengzhen Hou, Xiaolin Huang, Ying Chen, Chengyu Huo, Hongxing Liu, and Xinbao Ning

Combination of equiprobable symbolization and time reversal asymmetry for heartbeat interval series analysis

Fengzhen Hou,[1, 2] Xiaolin Huang,[1*] Ying Chen,[1] Chengyu Huo,[1] Hongxing Liu,[1] Xinbao Ning[1**]

[1] *Institute of Biomedical Electronic Engineering, School of Electronics Science and Engineering,*

Nanjing University, Nanjing, 210093, China

[2] *Information Management Teaching and Research Center, Department of Science, China*

Pharmaceutical University, Nanjing 210009, China

Abstract:

Symbolic dynamics method and time reversal asymmetry analysis are both important approaches in the study of heartbeat interval series. However, there is limited research work reported on combining these two methods. We provide a new method of time reversal asymmetry analysis which focuses on the differences between the forward and backward embedding 'm-words' after the operation of equiprobable symbolization. To investigate the total amplitude as well as the distribution features of the difference, four indices are proposed. Based on the application to simulation series, we found that these measures can successfully detect the time reversal asymmetry in chaos series. With the application to human heartbeat interval series (RR series), it is suggested that the distribution features of the forward-backward difference can sensitively capture the dynamical changes caused by diseases or aging. In particular, the index E_D, which reflects the random degree of the forward-backward difference distribution, can significantly discriminate healthy subjects from diseased ones. We conclude that RR series from the health shows more asymmetry in temporal structure in the original time scale from the perspective of equiprobable symbolization, whereas diseases account for the loss of this asymmetry.

---

[*] xlhuang@nju.edu.cn
[**] xbning@nju.edu.cn

1. Introduction

Time reversal asymmetry (TRA), which refers to the significant change of the statistical properties after time reversal [1-10], is closely related to the nonlinear dynamics or the non-Gaussian characteristics [1, 2].

The regulation of cardiac rhythm is a highly complex nonlinear process, which involves various inputs and multiple feedbacks [10, 11]. The heartbeat time series (RR series), which are defined as the sequence of time intervals between consecutive R peaks in the Electrocardiogram (ECG), have drawn much attention in the field of nonlinear dynamics analysis [12]. Recently, TRA has been demonstrated as one of the most important nonlinear characteristics of RR series by many researchers [3-9, 13, 14].

In the previous researches on TRA of RR series [3-9, 13, 14], people generally focused on evaluating the difference between the increments (denoted as $\triangle RR^+$ in the following) and the decrements (denoted as $\triangle RR^-$ in the following) of consecutive RR intervals. For example, Costa *et al.* suggested comparing the Shannon entropy of $\triangle RR^+$ and $\triangle RR^-$ over multiple scales [3]. Porta *et al.* advised to count the occurrences of $\triangle RR^+$ and $\triangle RR^-$ and value their percentage difference [4-6]. Guzik *et al.* provided an index to differentiate the square of $\triangle RR^+$ and the square of $\triangle RR^-$ [7-8]. Hou *et al.* proposed to simultaneously inspect both Porta's and Guzik's indices under multiple temporal scales and high-dimensional embedding [13-14].

In 2007, TRA of symbolized RR series was taken into analysis by Commarota *et al.* for the first time [9]. By setting the symbolization threshold to ± 10 ms, they firstly coded the one order differential RR series to ternary symbolic series with the three symbols of -1, 0, and 1. Subsequently, they compared the occurrence frequencies of sequence [1, 1, 1] and [-1, -1, -1] and found that the difference is significant, leading the conclusion that RR series is generally irreversible.

In Cammarota's method, 10 ms was intrinsically treated as the temporal resolution, and TRA of RR series was detected only by comparing between three successive accelerations and decelerations [9]. Commarota *et al.* took a valuable attempt to combine TRA analysis in RR series with the method of symbolic dynamics, and

demonstrated the great significance of this combination [9]. However, they ignored the contribution of other sequences to TRA of RR series, especially when the sequences [-1,-1,-1] and [1, 1, 1] account only for a small part of all. In addition, based on Kurths's research [15], when symbolic dynamics analysis is applied to RR series, the alphabet should be consisted of at least four symbols to preserve the intrinsic dynamics. Therefore, Commarota's symbolization method [9], which covers essentially only two states – the acceleration and deceleration of heart rates, might lose some useful information.

In essence, the symbolization of RR series is to convert the continuous-valued RR intervals into finite discrete symbols. It can simplify and speed up subsequent computation, as well as reduce the noise effects [16]. Meanwhile, the dominant dynamics features can be reserved as long as the symbolization method is appropriately applied.

Among symbolization methods of RR series, the most traditional one is proposed by Kurths et $al.$ [15]. It suggests to set partition boundaries at μ and $\mu \pm \alpha\mu$, where μ is statistic mean of the whole series and α is a pre-selected parameter [15]. Obviously this method will be affected by nonstationarity, and the selection of α might not be appropriate to other kinds of time series.

In 2007, Lin et $al.$ introduced a new symbolic representation of time series, which is named symbolic aggregate approximation (SAX) [17]. The method has received widespread attention in the field of data mining. Inspired by SAX, we propose the equiprobable symbolization (ES) method in this paper. Unlike the Kurths method [15], ES method divides the scalar measurement range into equiprobable regions so that all symbols in the alphabet have identical occurrence probabilities. Since it excludes the effects of mean value, standard deviation and the probability distribution of the original series, ES method can be applied to various time series.

In this paper, we firstly describe our general approach for symbolizing data and constructing forward and backward symbol-words. We also define the indices used to detect the TRA of time series. Then, we apply our method to several series which

have been proven irreversible for validation. Finally we apply it to RR series derived from different kinds of population: healthy young people, healthy old people, people suffering congestive heart failure (CHF), people suffering atrial fibrillation (AF), and people suffering sustained ventricular tachyarrhythmia (VT). Some meaningful results are presented and discussed in the following sections.

## 2．Time reversal asymmetry analysis based on equiprobable symbolization (TRAES)
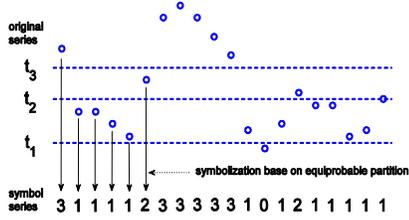
2.1. Equiprobable Symbolization (ES)

The original series given as $\{x_i : 1 \le i \le N\}$ are transformed into the symbolic series $\{s_i : 1 \le i \le N\}$ according to the rules as follows:

Firstly, the elements in the original series are sorted. Then, given the alphabet size n, the quantiles, denoted as $t_1, t_2, \cdots, t_{n-1}$ in ascending order, are determined by dividing the sorted elements into n equal-sized parts. Subsequently, the symbolization can be achieved through the following equation in the case of $n = 4$.

$$s_i = \begin{cases} 0 : x_i \le t_1 \\ 1 : t_1 < x_i \le t_2 \\ 2 : t_2 < x_i \le t_3 \\ 3 : t_3 < x_i \end{cases} \quad (1 \le i \le N). \quad (1)$$

Taking RR series for example, Fig. 1 indicates how ES method is developed.

(a)   (b)
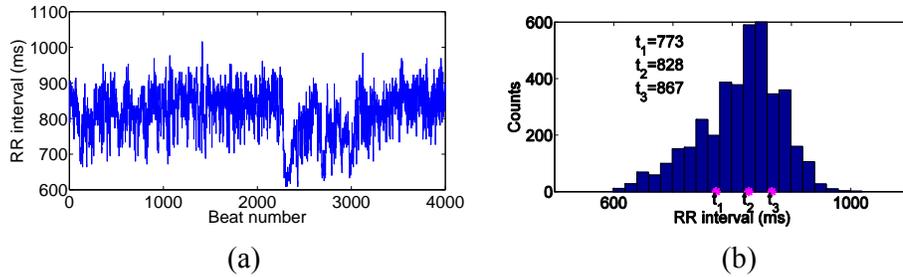
(c)

Fig. 1 (Color online) Diagram illustrating equiprobable symbolization. (a) An example of RR series with length 4000. (b) The histogram of RR intervals, in which 't_1', 't_2', 't_3' corresponding to the quartiles. (c) Symbolizing process based on the quartiles partition.

There are three key advantages in ES method. (1) It can directly reflect the dynamic features of the original series**.** Since the probability of each symbol in the alphabet becomes uniform in the original time scale, the frequency of each probable short sequence of certain length, which is also called 'word', depends solely on the time structure of the original series. (2) ES actually achieves the multi-resolution symbolization in the ranges of the original series. As shown in Fig.1 (b), more symbols are assigned to the dense regions to ensure enough resolution while fewer symbols are used to the sparse ones to reduce redundancy. (3) ES provides a widely applicable symbolization method since it is not affected by mean, standard deviation and probability distribution of the original series.

In ES, altering the number of symbols is very easy. According to the research of Kurths [15], at least 4 different symbols are necessary in order to preserve the essential and robust properties of the dynamics in RR sequences. Therefore, in this paper, 4 symbols are included in the alphabet.

2.2 Forward and backward symbolic words

In order to investigate the time asymmetry of the original series, as suggested by Daws *et al.* [2], we consider the forward symbolic-words (denoted as W_{fi}) and backward symbolic-words (denoted as W_{bi}). W_{fi} and W_{bi} are grouped together by

sequential symbols in nature (Equation 2) and reverse time (Equation 3), respectively.

$$W_{fi} = [s_i, s_{i+1}, \cdots, s_{i+m-2}, s_{i+m-1}], \quad (1 \leq i \leq N-m+1) \quad （2）$$

$$W_{bi} = [s_{i+m-1}, s_{i+m-2}, \cdots, s_{i+1}, s_i], \quad (1 \leq i \leq N-m+1) \quad （3）$$

This is conceptually similar to time-delay embedding with discrete symbols instead of continuously valued original measurements.

In the above equations, m is the word length, so the word is often called "m-word".

Each m-word, no matter it is constructed in nature time or reversed time, is a m-bit code of Quaternary, which can be converted to decimal as following

$$Q_{fi} = s_i \times 4^{m-1} + s_{i+1} \times 4^{m-2} + \cdots + s_{i+m-2} \times 4 + s_{i+m-1} + 1 \quad （4）,$$

$$Q_{bi} = s_{i+m-1} \times 4^{m-1} + s_{i+m-2} \times 4^{m-2} + \cdots + s_{i+1} \times 4 + s_i + 1 \quad （5）.$$

Where $Q_{fi}, Q_{bi} \in [1, 4^m]$.

We define $P_f(\lambda)$ and $P_b(\lambda)$ as the probable distribution of forward and backward m-words, respectively. The variable, λ, denotes the possible decimal code and $\lambda \in [1, 4^m]$. Therefore, the TRA of the original time series can be investigated through the difference between P_f and P_b, which are denoted as F-B difference from now on.

2.3 Measurements of TRA

In order to quantify the F-B difference, several measures are suggested in this paper.

(1) Euclidean Distance (D_E):

Taking P_f and P_b as two vectors in the multiple-dimensional space, the distance between them is calculated as

$$D_E = \sqrt{\sum_{\lambda=1}^{4^m} (P_f(\lambda) - P_b(\lambda))^2} \quad （6）.$$

In general, greater value of D_E corresponds to the more obvious temporal

asymmetry in the original series.

(2) Entropy of the difference (E_D)

The Euclidean Distance reflects only the total amplitude of the F-B difference, whereas it cannot express the distribution of the difference. Thus, we define the entropy of the difference as

$$E_D = - \sum_{P_d(\lambda) \neq 0} P_d(\lambda) * \log_2 P_d(\lambda) \qquad （7）,$$

where P_d stands for the normalized difference between P_f and P_b as

$$P_d(\lambda) = \frac{|P_f(\lambda) - P_b(\lambda)|}{\sum_{\theta=1}^{4^m} |P_f(\theta) - P_b(\theta)|} \qquad （8）.$$

The value of E_D reflects the random degree of the F-B difference distribution. Higher E_D means the difference distributes more uniformly, and vice versa. Different from the measurement D_E, E_D describes the time structure of the original series from a new aspect.

(3) The percentage of reversible words P_{RW} and the percentage of constant words P_{CW}

There are some special m-words with equal occurrences in the forward and backward sequences. Due to the limitation of logarithmic calculation, those words are not taken into account for the calculation of E_D. Therefore, we need to have a special treatment to those words.

There might be three cases for those words when word length is assigned to 4: (i) the words like "aaaa" which are constituted by four same symbols and are called constant words here; (ii) the words as "abba" which contain two different symbols but read same no matter forward or backward, and we call them symmetric words; and (iii) the rest words, which are called reversible words. The percentage of symmetric words combined with reversible words is denoted as P_{RW} in the following study. The value

of P_{RW} is directly related to the TRA of the original series, as the stronger asymmetry will decrease P_{RW} value.

The relationship between the constant words and the property of TRA is more complex. When the percentage of the constant words, denoted as P_{CW}, is not significantly great, the constant words can be treated as the reversible words. Otherwise, nontrivially large P_{CW} implies the substantial low-frequency rhythms, or trends, in current symbolizing resolution in the original time series. Therefore, we took the practice of considering P_{CW} as an independent parameter.

The measurements of E_D, P_{RW} and P_{CW}, constitute a full description to the distribution of the difference P_d.

2.4 The selection of word length

The word length can be seen as the length of the window from which we extract the underlying dynamical information, since the inner word temporal structure is mainly investigated. In general, the longer word-length set, the more information included. On the other hand, all embedding words of a series are deemed as a sample set and statistical analysis is then applied. In order to achieve reliable statistical results, the samples should be far more than possible patterns. Therefore, the selection of word length is a compromise of having both enough dynamical information and a sufficient good statistics to estimate the probability distribution. In this work, we have tried $m = 3$, $m = 4$, and $m = 5$, respectively, and it turns out that $m = 4$ is the optimal compromise. Consequently, 4 is used in the following sections.

2.5 The requirement of data length

Given alphabet size n and word length m, the amount of all possible m-words is n^m. In order to achieve reliable statistical results, the length of the original time series, denoted as N, should be taken far longer than n^m. In this paper, as the alphabet is consisted of four symbols, and the word length is assigned to 4, 2000 is considered enough for N. Moreover, both constant words and symmetric words may affect the

accuracy on the calculation of the index E_D, especially when they sum up to or even greater than 50%. Given above considerations, in this paper, the data length is allocated to 4000.

3. A numerical validation

As a validation, we applied the TRAES method to four kinds of chaotic maps [6, 13]: zeroth-order delayed Henon map (0-DHM), first-order delayed Henon map (1-DHM), zeroth-order delayed Tent map (1-DTM), and first-order delayed tent map (1-DTM). For each map, a series of length 4000 was obtained with the iterative computation. The four measurements suggested in TRAES method were calculated for those realizations and the results were listed in Tab. 1 (out of the brackets). For each series, 100 different surrogates were produced with the help of the iteratively refined amplitude-adjusted Fourier transform (IAAFT) algorithm [18, 19], and the surrogate data tests were used to check the significance of asymmetry [4]. The results of TRAES method applied to those surrogates were also listed in Tab. 1 (mean± standard deviation，in the brackets). As shown in Tab. 1, TRA was detected in all of the realizations through the measurements D_E, E_D, and P_{RW}, which is consistent with the previous researches [6, 13, 14].

Tab. 1 TRAES results of chaos series (out of the brackets) and their corresponding surrogates (mean±standard deviation, in the brackets). The mean and the standard deviation were obtained from 100 different surrogates.

	0-DHM	1-DHM	0-DTM	1-DTM
D_E	0.257(0.019±0.002)	0.122(0.019±0.002)	0.309(0.020±0.002)	0.215(0.020±0.002)
E_D	5.697(7.437±0.060)	6.907(7.429±0.059)	5.343(7.288±0.077)	6.401(7.396±0.068)
P_{RW}	0.001(0.100±0.026)	0.056(0.130±0.024)	0.024(0.119±0.028)	0.024(0.131±0.025)
P_{CW}	0.034(0.009±0.002)	0.020(0.009±0.002)	0.038(0.007±0.001)	0.018(0.009±0.002)

'0_DHM' stands for the zeroth-order delayed Henon map; '1-DHM' stands for the first-order delayed Henon map; '0-DTM' stands for the zeroth-order delayed Tent map; '1-DTM' stands for the

first-order delayed tent map.

As shown in Tab. 1, the differences between the original series and their surrogates can be successfully revealed by all the four indices. For each chaos series, the total amplitude of the F-B difference reflected by the parameter D_E is significantly greater than those of their surrogates, while the distribution of the difference reflected by the parameter E_D tends to be less random. Furthermore, it is noted that in all these series, either chaos or surrogates, the constant words only occupy tiny proportions. They can be treated as a reversal symmetry words and be analyzed by combining with P_{RW}. From the summations of P_{RW} and P_{CW}, it is found that the percentages of reversible words are less than those of their surrogates, which also reflects the TRA in the chaos series from another aspect. In a word, the results of the numerical validation suggest that the phenomenon of statistic symmetry is less noticeable in the original series than in their surrogates, showing the special temporal structures in the discussed chaos series.

## 4. TRAES analysis for RR intervals

The data analyzed here are five groups of RR series which come from different databases in the PhysioBank [20-22]: (i) the first group contains 22 records of healthy young subjects from nsrdb and nsr2db(11 female, 11 male, average age: 35), (ii) the second group contains 30 records of healthy elderly subjects from nsrdb and nsr2db （average age: 65, 19 female, 11 male）, (iii) the third group contains 36 records of subjects suffering congestive heart failure from chfdb and chf2db except one unknown subjects (6 female, 15 male, 15 unknown gender,, average age: 54), (iv) the fourth group contains 24 records of subjects suffering atrial fibrillation from afdb, and (vi) the last contains 22 records of subjects suffering a sustained ventricular tachyarrhythmia from sddb except 4 unknown subjects (7 female, 13 male, 2 unknown gender, average age: 59).

In the data preprocessing of RR series, the extreme intervals exceeding 2000 ms,

which might be introduced by the accumulation of missing records, were eliminated firstly. And then in order to exclude "outliers" from the records, the detection algorithm proposed in Ref. [23] and Ref. [24] has been applied, i.e., for each set of five contiguous intervals, if the local mean, excluding the central interval, is less than half of the central interval, this central interval is excluded from further analysis. No interpolation was done for eliminated intervals.

As mentioned above, the data length analyzed in this paper is 4000 (about 40 minutes of ECG sampling). Therefore, from each processed series, we extracted a segment of 4000 intervals according to the following two principles: the selected segments do not include any extreme intervals and they correspond to the waking state of subjects.

We applied the TRAES method to all of the selected segments. Noting that the symbols in the alphabet were still {0, 1, 2, 3} and the word length was also assigned to 4. As shown in Fig. 2, the four measurements suggested in TRAES method were calculated and the results were illustrated in the form of error bars. Meanwhile, unpaired t-tests were carried out to check the differences between different groups. The results of t-tests were listed in Tab. 2. For comparison, we applied the time and frequency domain analyses [25] to all the same segments. Table2 shows the results of t-tests.
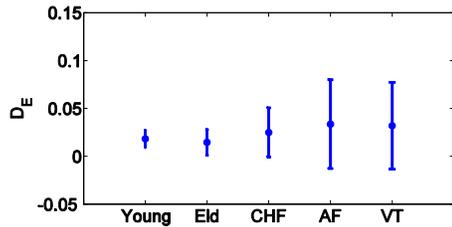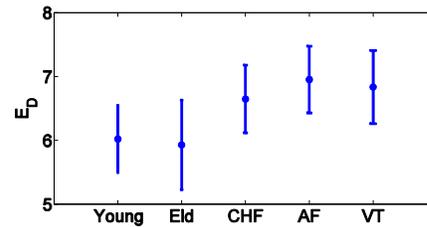
(a)   (b)
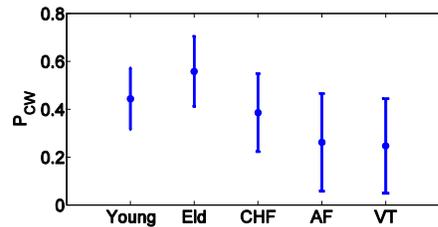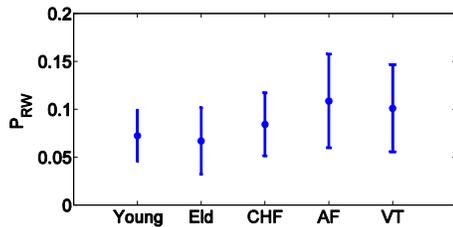
(c)   (d)

Fig. 2 (Color online) TRAES analysis of RR series from healthy young subjects (Young), healthy elderly subjects (Eld), congestive heart failure sufferers (CHF), atrial fibrillation sufferers (AF), and ventricular tachyarrhythmia sufferers (VT) using indices (a) D_E, (b) E_D, (c) P_{RW}, and (d) P_{CW}, where series length $N = 4000$, word length $m = 4$, and alphabet size $n = 4$.

Tab. 2 Statistic significance tests for the differences between the young healthy (Yng) and the elderly healthy (Eld), as well as healthy and sick subjects. The series length $N = 4000$, the word length $m = 4$ and alphabet size $n = 4$. 'n.s.' stands for no significant difference at the 1% significance level.

	Yng vs. Eld	Yng vs. CHF	Yng vs. AF	Yng vs.VT	Eld vs. CHF	Eld vs. AF	Eld vs. VT
D_E	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
E_D	n.s.	10^{-5}	10^{-7}	10^{-5}	10^{-5}	10^{-7}	10^{-6}
P_{RW}	n.s.	n.s.	10^{-3}	n.s.	n.s.	10^{-4}	10^{-3}
P_{CW}	10^{-3}	n.s.	10^{-4}	10^{-4}	10^{-5}	10^{-8}	10^{-8}
mRR	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
SDNN	n.s.	10^{-5}	n.s.	n.s.	10^{-4}	n.s.	n.s.
LF	n.s.	10^{-4}	n.s.	n.s.	10^{-3}	n.s.	n.s.
nLF	n.s.	10^{-4}	10^{-4}	10^{-4}	n.s.	10^{-3}	n.s.
HF	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
nHF	n.s.	10^{-3}	10^{-4}	10^{-3}	n.s.	10^{-3}	n.s.

Note: mRR— mean RR intervals from ECG, SDNN —standard deviation of RR intervals,

LF—low-frequency spectral component of HRV, HF—high-frequency spectral component of HRV,

nLF—, normalized LF, and nHF—normalized HF.

From Fig. 2 and Tab. 2, we found that:

(1) The significant differences of index D_E are not observed between any two different groups. It suggests that after equiprobable symbolization, the total amplitude of F-B difference is not enough to elucidate the dynamical changes brought by diseases or aging effects in RR series.

(2) E_D of the healthy groups are significantly lower than those of the CHF, AF and VT groups. As a rational deduction, we speculated that the diseases may lose time asymmetry in cardiac rhythm dynamics. Moreover, the results of t-tests indicated that there is also significant difference ($p < 0.05$) between CHF and AF sufferers showing the different effects on cardiac dynamics of different diseases. When compared to conventional analysis, such as mean heart rate, standard deviation and frequency domain parameters, the index E_D provides a better ability to distinguish different groups.

(3) The index P_{RW} has the similar trends as E_D, except that the discrimination ability is weaker. The values of P_{RW} from AF and VT groups are significantly higher than the healthy. It indicates that the asymmetry reduces for the two kinds of cardiac diseases even from the point of the reversal words merely.

(4) As shown in Fig. 2(d), there are considerable constant words in RR series, especially in those from the healthy subjects. There are two possible reasons for the formation of a large number of constant words: the alphabet size n is not large enough to give a fine amplitude resolution, or the time series have substantial low frequency trends. With these suspicions, we examine the effects brought by changes of alphabet size n. As we know, the choice of n is a trade-off between the cost and effectiveness -- greater n leads to finer amplitude resolution and longer data length but increases the computation. In addition, the improvements in resolution are restricted by the original resolution of experimental data. Particularly, for RR series analyzed in this paper, the original amplitude resolution is limited by ECG sampling rates (128Hz and 250Hz).

Therefore, during the symbolization, if some difference between the adjacent quantiles gets less than 10 ms, which approaches to the RR resolution, it will lead to series noise contamination instead of finer resolution. We found that, when alphabet size n is set to 4, only 3 records have such close adjacent quantiles, while the total of such records increases to 29 when n is increased to 8. As a result, we will not consider the situations when n is greater than 8. To ensure the statistical validity, when the word length $m = 4$, the data length are set as $N = 20000$. Because two VT records cannot meet the requirement of data length, 132 records in total are calculated in this step. We only give the illustrations of indices E_D and P_{CW} in Fig. 3, since the other two indices are in substantial agreements with Fig. 2. The results for $n = 5$ and 7 are also omitted for the similarities.
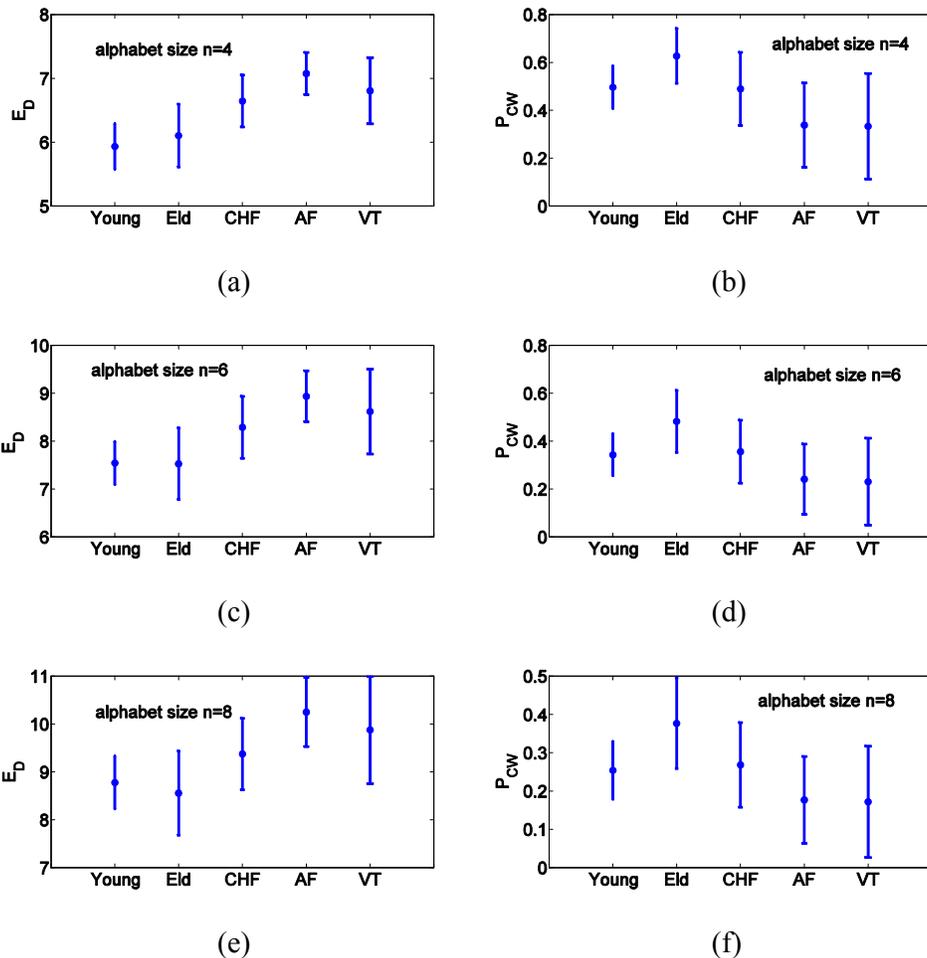

(a)

(b)

(c)

(d)

(e)

(f)

Fig. 3 (Color online) TRAES analysis of RR series with the alphabet size of (a)-(b) $n = 4$, (c)-(d) $n = 6$, and (e)-(f) $n = 8$. The series length $N = 20000$ and the word length $m = 4$.

As shown in Fig. 3 (a), (c), and (e), the discrimination ability between groups for E_D does not get better as n increases, and the most excellent discrimination ability is still achieved when $n = 4$. In addition, though P_{CW} does decrease as n increases, there are still considerable constant words even as n is increased to 8. It suggests that there are substantial low frequency components in cardiac rhythm, and the percentages in the elderly healthy subjects are significantly higher than those in the young healthy ones. However, the AF and VT groups have lower values. It implies that purely aging might increases the low frequency elements in the healthy cardiac dynamic systems, while the diseases of AF and VT might decrease them on the contrary.

The low frequency trends of RR series have no determinate correspondence with time asymmetry, since current sampling is somewhat over-sampling to such trends. As we known, given the word length invariant, over-sampling will lead to false linear approximation for any system. We also attempt to increase the word length m to 6, but find little difference. In order to find out TRA of these trends, we conjecture that multi-scale analysis [3] should be taken into consideration. Nevertheless this also raises demands for much longer data length.

5. Conclusions

In this paper, time reversal asymmetry analysis combined with equiprobable symbolization is applied to chaos series and human RR intervals in different physiological and pathologic statuses. To measure both the total amplitude and the distribution features of the forward-backward differences of embedding symbolic

'm-words', four indices are introduced.

By using the measurements we proposed, temporal reversal asymmetry has been confirmed in the Henon maps as well as in Tent maps, suggesting that the proposed method is a reliable tool for the TRA detection of time series.

When applying to the RR series, it is found that the distribution features of the forward-backward difference can sensitively capture the dynamical changes brought by diseases or aging. E.g., the index E_D, which stands for the random degree of F-B difference distribution, can significantly differentiate RR series derived from the health and disease populations. The percentage of reversible words, P_{RW}, is significantly greater in patients with CHF, AF and VT than in the healthy people. Both indices reflect the TRA descent in RR series in the original time scale caused by cardiac diseases. Considering the good discrimination ability between groups, the index E_D may be helpful for clinical diagnostics of heart diseases. What is more, in the proposed method, the embedding word length 4 determined that the temporal structure of successive 5 heart beats was mainly investigated. Physiologically, this temporal scale mostly corresponds to respiratory sinus arrhythmia and it is greatly affected by vagal regulation. Therefore, according to above results, we speculate that these diseases have correlation with changes of vagal-function and the indices, especially E_D, are sensitive to these changes. In addition, by comparing to the classical frequency domain analysis, it is found that the proposed index E_D can discriminate the health and the diseases more significantly than traditional parameters such as LF and HF, which are considered as the major indicators of sympathetic and vagal activities [25]. The proposed method may be appropriate for elucidating the neural pathophysiological mechanisms occurring during the mentioned diseases. However, in order to prove whether it can be used as a probe to investigate cardiac autonomic modulation or not, experimental models for sympathetic activation, such as tilt test and high-dose atropine administration, or parasympathetic activation, such as phenylephrine and low-dose atropine administration, should be designed in the

future[26].

It is also found that the rhythm of low-frequency, which is represented by the index P_{CW}, is considerably presented in the cardiac dynamics of healthy human. In addition, it significantly increases with aging but decreases with the occurrence of diseases. Although the physiological essence of this phenomenon requires further studies, it still suggests that aging affects cardiac rhythm regulation differently from diseases studied in this paper.

Acknowledgement

References

[1]   C. Diks, J. C. van Houwelingen, F. Takens, and J. DeGoede, Phys. Lett. A **201**, 221 (1995).

[2]   C. S. Daw, C. E. A. Finney, and M. B. Kennel, Phys. Rev. E **62** 1912 (2000).

[3]   M. Costa, A. L. Goldberger, and C.-K. Peng, Phys. Rev. Lett. **95**, 198102 (2005).

[4]   A. Porta, K. R. Casali, A. G. Casali, T. Gnecchi-Ruscone, E. Tobaldini, N. Montano, S. Lange, D. Geue, D. Cysarz, and P. Van Leeuwen, Am. J. Physiol. Regul. Integr. Comp. Physiol. **295**, 550 (2008).

[5]   A. Porta, S. Guzzetti, N. Montano, T. Gnecchi-Ruscone, R. Furlan, and A. Malliani, Comput. Cardiol. **33**, 77 (2006).

[6]   A. Porta, G. Daddio, T. Bassani, R. Maestri, and G. D. Pinna, Phil. Trans. R. Soc. A **367**, 1359 (2009).

[7]   P. Guzik, J. Piskorski, T. Krauze, A. Wykretowicz, and H. Wysocki, Biomed. Tech. **51**, 272 (2006).

[8]   J. Piskorski, and P. Guzik, Physiol. Meas. **28**, 287 (2007).

[9]   C. Cammarota, and E. Rogora, Chaos Soliton. Fract. **32**, 1649 (2007).

[10] K.R. Casali, A.G. Casali, N. Montano, M.C. Irigoyen, F. Macagnan, S. Guzzetti, and A. Porta, Phys. Rev. E **77**, 066204 (2008).

[11] R. T. Baillie, A. A. Cecen, and C. Erkal, Chaos **19**, 028503 (2009).

[12] A. Voss A, A. Schulz, R. Schroeder, M. Baumert, and P. Caminal, Phil. Trans. R. Soc. A **367**, 277 (2009).

[13] F. Z. Hou, J. J. Zhuang, C. H. Bian, T. J. Tong, Y Chen, J Yin, X. J. Qiu, and X.B. Ning, Physica A **389**, 754 (2010).

[14] F. Z. Hou, X. B. Ning, J. J. Zhuang, X. L. Huang, M. J. Fu, and C. H. Bian, Med. Eng. Phys. **33**, 633 (2011).

[15] J. Kurths, A. Voss, A. Witt, P. Saparin, H. J. Kleiner, and N. Wessel, Chaos **5**, 88 (1995).

[16] C. S. Daw, C. E.A Finney, and E. R. Tracy, Rev. Sci. Instrum. **74**, 915 (2003).

[17] J. Lin, E. Keogh, L. Wei, and S. Lonardi, Data Min. Knowl. Disc. **15**, 107 (2007).

[18] T. Schreiber, and A. Schmitz, Physica D **142**, 346 (2000).

[19] T. Schreiber, and A. Schmitz. Phys. Rev. Lett. **77**, 635 (1996).

[20] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P.Ch. Ivanov, R.G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, Circulation **101**, 215 (2000) (see also http://www.physionet.org)

[21] D. S. Baim, W. S. Colucci, E. S. Monrad, H. S. Smith, R. F. Wright, A. Lanoue, D. F. Gauthier, B. J. Ransil, W. Grossman, and E. Braunwald, J. Am. Coll. Cardiol. **7**, 661 (1986)

[22] S. Petrutiu, A. V. Sahakian, and S. Swiryn, Europace **9**, 466 (2007)

[23] P. C. Ivanov, L. A. N. Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik, and H. E. Stanley, Nature **399**, 461 (1999).

[24] P. A. Varotsos, N. V. Sarlis, E. S. Skordas, and M. S. Lazaridou, Appl. Phys. Lett. **91**, 064106 (2007).

[25] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A. Malliani, A.J. Moss and P. J. Schwartz, Europ. Heart J. **17**, 354 (1996).

[26] S. Guzzetti, E. Borroni, P. E. Garbelli, E. Ceriani, P. D. Bella, . N. Montano, C.Cogliati, V. K. Somers, A. Mallani, A. Porta, Circulation **112**, 465(2005).