35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Multi-scale Tone Entropy in Differentiating Physiologic and Synthetic RR Time Series

Chandan K. Karmakar, *Member, IEEE*, Ahsan H. Khandoker, *Senior Member, IEEE*, and Marimuthu Palaniswami, *Fellow, IEEE*

***Abstract*— Heart rhythm is extrinsically modulated by the autonomic nervous system and recently, the Tone-Entropy (T-E) measurement was reported as a measure of autonomic balance and activity in time domain HRV analysis. Current algorithm for T-E measurement describes only beat-to-beat or influence of a heart beat on a train of succeeding beats on a single scale. Therefore, conventional T-E analysis has often not been able to discern various physiological conditions using heart rate variability (HRV) signal. In this study, we will present a mathematical framework to define multi-scale T-E analysis, apply this in differentiating physiological and synthetic RR time series. Finally, we compare the performance of proposed parameters with conventional T-E measurements.**

I. Introduction

The RR interval (interbeat) time series is used for studying heart rate variability (HRV), which is a popular and noninvasive tool to study cardiac autonomic activity [1-3]. The heart rate changes over time with no linear relationship between HR and time. Thus, the underlying mechanism involved in human heart rate control has been reported to be complex and nonlinear [4]. Therefore, analysis of the dynamic behavior of RR interval time series has opened up a new approach towards the assessment of normal and pathological cardiovascular behavior.

The conventional Tone-Entropy (T-E) method of quantifying heart rate variability (HRV) uses single-scale based successive R-R intervals. The physiological interpretations of single-scale T-E in various experimental settings were previously reported [5, 6]. Lower Tone values (negative) indicate that vagal activity predominates in the sympatho-vagal balance in a healthy population at rest. On the other hand, Entropy represents the total sympatho-vagal activity which means that entropy increases with increasing sympatho-vagal activity and vice versa [6-8]. The lagged T-E method was reported by Karmakar et. al. and the authors have used multi-lag T-E method to identify increased risk of post-infarct people with diabetes [9]. In previous studies, Khandoker et. al. [5] and Karmakar et. al. [10] reported the use of conventional and multi-lag T-E method in classifying cardiac autonomic neuropathy and monitoring severity of cardiovascular autonomic neuropathy (CAN) in diabetic subjects .

C. Karmakar, A. Khandoker and M. Palaniswami are with the Electrical & Electronic Engineering Department, University of Melbourne, Parkville, Melbourne, VIC 3010, Australia (Phone: +61-3-9035-7614; email: karmakar@unimelb.edu.au, ahsank@unimelb.edu.au; palani@unimelb.edu.au).

A. Khandoker is also with Dept of Biomedical Engineering, Khalifa University, Abu Dhabi, UAE.

Conventional and lagged T-E measurement both uses single scale RR time series to assess cardiac autonomic system. However, no study has reported the use of T-E in multiple scale, despite that fact that biological systems are complex and operate across multiple spatial and temporal scales [11]. Multi-scale entropy analysis of HRV signal was first proposed by Costa et. al. which have been used in a wide range of physiological conditions [11-13]. Motivated by multi-scale properties of cardiac autonomic system, in this study we first developed the multi-scale T-E measurement technique and then applied this approach to differentiate physiological (real) and synthetic RR interval data. We also compared to performance of multi-scale T-E analysis with conventional T-E measurement to prove that multi-scale T-E provides additional information about the underlying system than conventional T-E measurement.

II. Data and Methods

A. Data

We used 50 RR interval series obtained from the PhysioNet/CinC Challenge 2002 dataset [14]. Each series was extracted from a 20–24h electrocardiogram (ECG) signal (approx. 70,000–130,000 points). 26 of these signal was derived from ambulatory ECG recordings of people aged from 20 to 50 years who had no known cardiac abnormality and 22 consisted of synthetic data that had been generated to emulate healthy RR intervals using 11 different models, and the remaining two were time-reversed physiological series [15]. In this study, the last two time-reversed series were reverted to real physiological signal. Hence, there were 28 real physiological (real) and 22 synthetic series used for the study. In this study, we used a segment of 40000 consecutive RR intervals from the beginning for all subjects.

B. Multi-scale T-E analysis of HRV signal

In this study, we have used the scaling approach proposed by Costa et. al. [11]. In brief, for a given time series $\{y_1, y_2, y_3 \dots, y_N\}$ the coarse-grained time series at scale m is constructed by averaging m number of data points in non-overlapping windows. j-th element of the coarse-gained time series at scale m is calculated using eq. (1).

$$y_j^m = \frac{1}{m} \sum_{i=(j-1)*m+1}^{j*m} y_i \qquad (1)$$

A RR interval or period is defined as the time difference between two consecutive R peaks of the electrocardiogram (ECG) signal. Let the RR intervals time series **RR** be defined as:

$$\boldsymbol{RR} \equiv \{RR_1, RR_2, \dots . RR_N\}$$

where, N is the number RR intervals. Heart rate acceleration and inhibition can be determined from the difference of consecutive RR intervals. If RR_{i+1} become shorter than RR_i then it is an acceleration of heart rate. Therefore, acceleration of the heart is expressed as a plus difference and inhibition as a minus difference of RR intervals. However, to reduce the impact of heart rate variation over a wide range of time and different subjects, normalized variation in RR interval is preferred to monitor the variability. In conventional T-E analysis, percentile change of the successive RR intervals with respect to the first RR interval is expressed as the percentage index (PI) and defined as:

$$PI(i) = \frac{RR_i - RR_{i+1}}{RR_i} \times 100 \qquad (2)$$

The Tone is defined as a first order moment (arithmetic average) of this PI time series as:

$$Tone = \frac{1}{N-1}\sum_{i=1}^{N-1} PI(i) \qquad (3)$$

Tone is the balance between accelerations (PI > 0) and inhibitions (PI< 0) of the heart rate and represents the sympatho-vagal balance faithfully as shown in previous studies [6, 16]. Entropy is defined from the probability distribution of PI by using Shannon's formula [17]:

$$Entropy = -\sum_{i=1}^{n} p(i) log_2 p(i) \qquad (4)$$

where, $p(i)$ is a probability of PI having values in the range $i < PI < i+1$, where i is an integer. The entropy evaluates total acceleration–inhibition activities, or total heart period variations, in a familiar unit of bit.

For multi-scale T-E analysis, we have introduced the scale m in equation (2), used to derive the PI time series from the RR time series signal. Hence, in the multi-scale T-E analysis PI is expressed as the percentile change of the i-th and $i+1$-th RR inverval at scale m and is defined as:

$$PI^m(i) = \frac{RR_i^m - RR_{i+1}^m}{RR_i^m} \times 100 \qquad (5)$$

where, m is an integer and $m = 1$ represents the conventional T-E analysis. The detailed methodology of conventional T-E analysis has been described in previous reports [6, 16].

B. Classification of groups

In this study, we used the Quadratic Discriminant (QD) classifier rather than the more traditional linear classifier to test the ability of Tone and Entropy values individually and together over multiple scales to differentiate real RR time series from synthetic. The classification at any scale m represents the classification is performed using corresponding parameter (Tone, Entropy or both together) at that scale only. On the other hand, the scale limit k for any parameter (Tone, Entropy or both together) represents all parameters upto scale k. For example, $k = 5$ for $Tone$ indicates that classification is performed using all $Tone$ parameters from scale 1 to 5. A leave-one-out cross-validation scheme was adopted to evaluate the generalization ability of the classifiers. Cross-validation procedures have been used in a number of classification evaluations, particularly for limited data sets [18]. In this scheme the data set was uniformly divided into 50 subsets with one used for testing (unknown to classify) and the remaining 49 records used to train the classifiers. This was repeated for the remaining subsets so that all subsets were used as the testing sample.

The following three measures of accuracy, sensitivity and specificity were used to assess the performance of the classifiers [19, 20]:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100$$

$$Specificity = \frac{TN}{TN + FP} \times 100$$

Where, TP is the number of true positives, i.e., the classifier identifies a RR interval time series that was labeled as synthetic; TN is the number of true negatives, i.e., the classifier identifies a RR interval time series that was labeled as real; FP is false synthetic identifications; and FN is false real identifications. Accuracy indicates overall detection accuracy, sensitivity is defined as the ability of the classifier to accurately recognize a synthetic RR time series whereas specificity indicates the classifier's ability not to generate a false negative.

C. Statistics

The non-parametric Mann-Whitney U-test was performed to allow for pairwise testing for significant differences of HRV parameters between the two groups. Since, the number of subjects are small and their distribution is not normal non-parametric test is more appropriate than parametric test.

III. Results & Discussion

Mean ± SD (standard deviation) $Tone$ and $Entropy$ values for both real and synthetic RR interval time series are given in Table 1. Mean $Tone$ value is lower for synthetic RR time series at scale $m = 1{\sim}2$, however it is higher for the other scales. The reason behind lower mean $Tone$ value at scale $m = 1$ due to the very low $Tone$ value of two synthetic RR time series that is clearly visible in Figure 1 (Scale 1). Even with such exception, there is an opposite trend in mean $Tone$ value for real and synthetic RR time series. With increasing scale mean $Tone$ decreases for real RR time series whereas, it increases for synthetic RR time series. Since $Tone$ represents the sympatho-vagal balance [5, 21], this results indicates that the sympatho-vagal balance of synthetic RR time series tends to symmetricity ($Tone = 0$) with increasing scale. Therefore, the higher asymmetricity (lower sympatho-vagal balance) of synthetic RR time series at lower scales may be due randomness rather than complexity of underlying structures. This supports the findings reported by Khandoker et al. [5], in which author found higher Tone (lower sympatho-vagal balance) values for subjects with

definite cardiac autonomic neuropathy (CAN). The $Tone$ values are significantly different among two groups at scale $m = 4{\sim}8$, which supports the need for multi-scale study over conventional single scale study [11].

TABLE I. MEAN ± SD VALUES OF TONE AND ENTROPY PARAMETERS OF REAL AND SYNTHETIC RR TIME SERIES FOR SCALE M=1~8.

Parameter	Scale m	Real (Mean ± SD)	Synthetic (Mean ± SD)
Tone	1	-0.12 ± 0.10	-0.47 ± 1.16
	2	-0.14 ± 0.08	-0.17 ± 0.20
	3	-0.16 ± 0.10	-0.15 ± 0.19
	4	-0.18 ± 0.11	-0.13 ± 0.15*
	5	-0.19 ± 0.10	-0.12 ± 0.11**
	6	-0.19 ± 0.10	-0.11 ± 0.09**
	7	-0.19 ± 0.09	-0.10 ± 0.08**
	8	-0.18 ± 0.09	-0.10 ± 0.09*
Entropy	1	3.91 ± 0.57	3.95 ± 0.84
	2	4.21 ± 0.42	4.00 ± 0.87
	3	4.31 ± 0.42	3.87 ± 0.96
	4	4.40 ± 0.41	3.85 ± 0.92*
	5	4.45 ± 0.38	3.83 ± 0.86*
	6	4.48 ± 0.35	3.82 ± 0.82*
	7	4.48 ± 0.33	3.78 ± 0.80**
	8	4.48 ± 0.31	3.79 ± 0.80**

* p<0.05; ** p<0.01

The $Entropy$ of a RR time series represents the overall sympathetic-parasympathetic activity. Therefore, higher $Entropy$ values indicate increase in total activity and lower values indicate decline. Although $Entropy$ is a measure of complexity it is related to the degree of randomness as well [11], therefore the higher mean $Entropy$ value of synthetic RR time series at scale $m = 1$ than real RR time series found in this study (Table 1) may be because of randomness rather than underlying system complexity. This is also supported by the findings that with increasing scale the mean $Entropy$ values decline for synthetic RR time series, whereas the opposite/no trend is found for real RR time series. This indicates that the $Entropy$ value of real RR time series is due to complexity of underlying structure rather than randomness. Similar to $Tone$ values, the $Entropy$ values are also significantly different among two groups at scales $m = 4{\sim}8$ rather than at scale $m = 1$. Therefore, both $Tone$ and $Entropy$ value supports the necessity of multi-scale analysis to differentiate between real and synthetic RR time series signal.

The $Tone - Entropy$ $(T - E)$ space of RR time series is a visual representation of regulation or dysregulation of RR interval time series. Higher $Tone$ and lower $Entropy$ are indicative of autonomic nervous system dysregulation and vice versa. In a previous study, Khandoker et. al. have reported that the lower $Tone$ and higher $Entropy$ in healthy subjects in supine resting condition compared to parasympathetic perturbed conditions (atropine infusion and 70^0 head-up tilt) [5]. In this study we have found similar findings where the real RR time series have lower mean $Tone$ and higher mean $Entropy$ values than synthetic RR time series. However, at lower scale $m = 1{\sim}4$ the $T - E$ spaces of two groups are highly overlapping, which reduces with increasing scales (Figure 1). The important finding in this study is that at higher scales $m > 1$ the $T - E$ space has shown better difference among two groups (real and synthetic RR time series) than at scale $m = 1$.

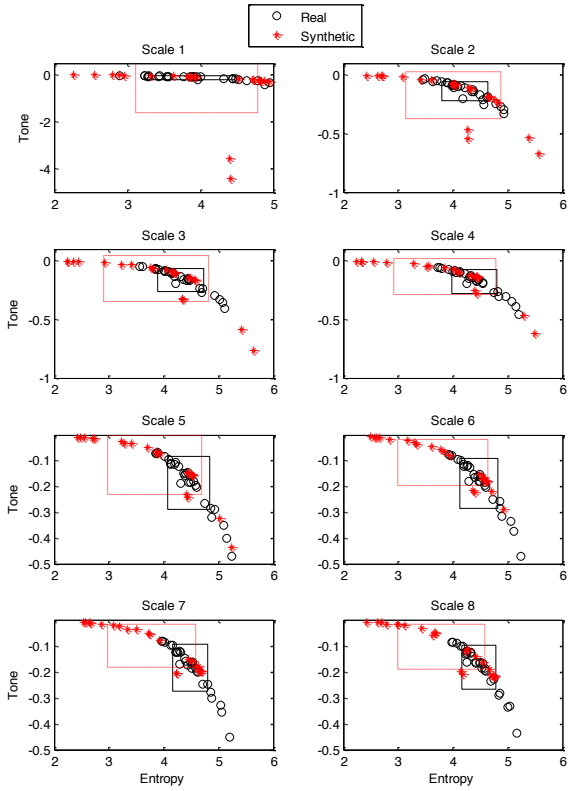

Figure 1. Evaluated Tone and Entropy in T-E space, ensemble averages by open rectangles (mean ± SD) and indivduals, by symbols. Dotted red rectangles are for synthetic RR time series and solid black rectangles represent real RR time series.

The sensitivity, specificity and accuracy of individual $Tone$ and $Entropy$ as well as both parameters together using QD classifier at individual scale $m = 1{\sim}8$ and scale limit k (ranging from 1 to k) are shown in Figure 2. At individual scale $Tone$ has shown lowest accuracy at all scale, except m=3, compared to $Entropy$ and both parameters together. $T - E$ together has shown highest accuracy at all scale except $m = 5$, where Entropy has shown the maximum accuracy. Finally, the highest accuracy in individual scale analysis was found 84% using $Tone$ and $Entropy$ at scale $m = 7$. Therefore, we can conclude that $T - E$ together has performed better to classify real and synthetic RR time series than individual $Tone$ or $Entropy$ parameter.

For scale limit k, again $Tone$ has shown lowest accuracy at all limits $k = 1{\sim}8$. $T - E$ together has shown maximum accuracy at all limits except $k = \{1, 3, 8\}$. However, the highest accuracy 96% was found for both $Entropy$ and $T - E$ parameters at scale limit $k = 8$. Therefore, we can conclude that $Entropy$ and $T - E$ parameters over a scale

limit perform best in classifying real and synthetic RR time series. This findings support the findings reported by Costa et. al. in which the authors have used the pattern over multiple scale to differentiate real and synthetic RR time series [12]. However, the authors have used a qualitative or manual approach in differentiating both groups rather than automatic classification. From the results of this study, we can conclude that real and synthetic RR time series can be better differentiated using multi-scale $T-E$ analysis than conventional single scale $T-E$ analysis.

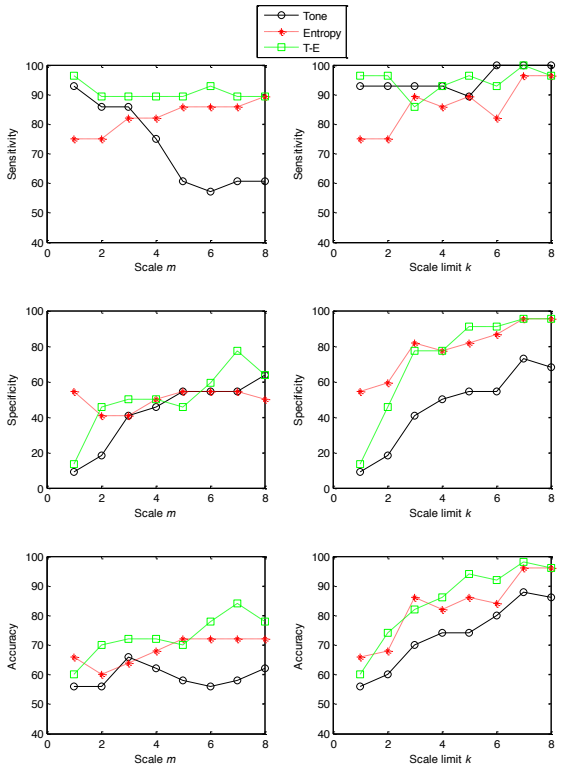

Figure 2. Sensitivity, Specificity and Accuracy of individual Tone and Entropy as well as both parameters together using QD classifier at individual scale *m=1~8* and scale limit *k* (ranging from 1 to k)

IV. Conclusion

In this paper, we studied the benefit of using multi-scale $Tone-Entropy$ analysis to differentiate between real and synthetic RR time series. The results of the investigation indicate the multi-scale $T-E$ analysis provides additional information to conventional analysis. The conventional single scale T-E could not differentiate between real and synthetic RR time series, whereas multi-scale T-E has shown 96% accuracy in differentiating the two groups. In future, it will be interesting to see how multi-scale T-E performs in discriminating various pathological conditions.

Acknowledgment

This study was partly funded by the University of Melbourne Early Career Researcher Grant.

References

[1] H. V. Huikuri, T. H. Mäkikallio, K. E. Airaksinen, T. Seppanen, P. Puukka, I. J. Raiha, and L. B. Sourander, "Power-law relationship of heart rate variability as a predictor of mortality in the elderly," *Circulation*, vol. 97, pp. 2031-2036, 1998.

[2] M. Brennan, M. Palaniswami, and P. Kamen, "Do existing measures of Poincare plot geometry reflect nonlinear features of heart rate variability?" *IEEE Trans. Biomed. Eng.*, vol. 48, no. 11, pp. 1342-1347, 2001.

[3] M. Brennan, M. Palaniswami, and P. Kamen, "Distortion Properties of the Interval Spectrum of IPFM Generated Heartbeats for Heart Rate Variability Analysis," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 11, pp. 1251- 1264, November 2001.

[4] P. C. Ivanov and M. G. Rosenblum, "Scaling behaviour of heartbeat intervals obtained by wavelet-based time-series analysis," *Nature*, vol. 383, pp. 323 – 32, 1966.

[5] A. H. Khandoker, H. Jelinek, T. Moritani and M. Palaniswami, "Association of cardiac autonomic neuropathy with alteration of sympatho-vagal balance through heart rate variability analysis," *Medical Engineering & Physics*, vol. 32, pp. 161-167, 2010.

[6] E. Oida, T. Moritani and Y. Yamori, "Tone–entropy analysis on cardiac recovery after dynamic exercise," *J Appl Physiol*, vol. 82, pp. 1794–1801, 1997.

[7] M. Bootsma, C. A. Swenne, H. H. Van Bolhuis, P. C. Chang, V. M. Cats and A. V. Bruschke, "Heart rate and heart rate variability as indexes of sympathovagal balance," *Am J Physiol*, vol. 266, pp. H1565–H1571, 1994.

[8] M. P. Tulppo, R. L. Hughson, T. H. Makikallio, K. E. Airaksinen, T. Seppanen and H. V. Huikuri, "Effects of exercise and passive head-up tilt on fractal and complexity properties of heart rate dynamics," *Am J Physiol Heart Circ Physiol*, vol. 280, pp. H1081–H1087, 2001.

[9] C. Karmakar, H. Jelinek, A. Khandoker, M. Tulppo, T. Makikallio, A. Kiviniemi, H. Huikuri and M. Palaniswami, "Identifying increased risk of post-infarct people with diabetes using multi-lag Tone-Entropy analysis," *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE* , pp. 25-28, 2012.

[10] C. K. Karmakar, A. H. Khandoker, H. F. Jelinek and M Palaniswami, "Risk stratification of cardiac autonomic neuropathy based on multi-lag Tone-Entropy", *Medical & Biological Engineering & Computing*, Accepted on December 18, 2012.

[11] M. Costa, A. L. Goldberge and C. K. Peng,."Multiscale entropy analysis of complex physiologic time series," *Phys Rev Lett* , vol. 89:068102, 2002.

[12] M. Costa M, A. L. Goldberger AL and Peng CK, "Multiscale entropy to distinguish physiologic and synthetic RR time series", *Comput Cardiol*, vol. 29, pp. 137-140, 2002.

[13] P. R. Norris, S. M. Anderson, J. M. Jenkins, A. E. Williams and J. A. Morris Jr, "Heart rate multiscale entropy at three hours predicts hospital mortality in 3,154 trauma patients,", *Shock*, vol. 30, pp. 17–22, 2008.

[14] Physionet, http://www.physionet.org/challenge/2002/dataset/.

[15] G. B. Moody, " RR interval time series modeling: the PhysioNet/Computers in Cardiology challenge 2002," *Comp Cardiol*, vol. 29, pp. 125–128, 2002.

[16] M. Amano, E. Oida and T. Moritani, "Age-associated alteration of sympatho-vagal balance in a female population assessed through the tone–entropy analysis," *Eur J Appl Physiol*, vol. 94, pp. 602–610, 2005.

[17] C. E. Shannon, "A mathematical theory of communication," *Bell Syst Tech J*, vol. 27, pp.379–423, 1948.

[18] B. D. Ripley, " Pattern Recognition and Neural Networks," *Cambridge Univ Press*, Cambridge, UK, 1996.

[19] K. Chan, T. W. Lee, P. A. Sample, M. H. Goldbaum, R. N. Weinreb and T. J. Sejnowski, "Comparison of machine learning and traditional classifiers in glaucoma diagnosis," *IEEE Trans Biomed Eng*, vol. 49, pp. 963–974, 2002.

[20] C. C. C. Pang, A. R. M. Upton, G. Shine and M. V. Kamath, "A comparison of algorithms for detection of spikes in the electroencephalogram," *IEEE Trans Biomed Eng*, vol. 50, pp. 521–526, 2003.

[21] E. Oida, T. Kannagi, T. Moritani and Y. Yamori, "Aging alteration of cardiac vagosympathetic balance assessed through the tone–entropy analysis," *J Gerontol,* vol. 54A, pp. M219–M224, 1999.