Fluctuation and Noise Letters
Vol. 5, No. 4 (2005) L557–L563

ESTIMATING THE COMPLEXITY OF HEART RATE FLUCTUATIONS — AN APPROACH BASED ON COMPRESSION ENTROPY

MATHIAS BAUMERT*, VICO BAIER and ANDREAS VOSS

University of Applied Sciences Jena, Department of Medical Engineering,
Carl-Zeiss-Promenade 2, 07745 Jena, Germany
**mathiasbaumert@web.de*

LARS BRECHTEL

Humboldt University, Institute for Sports,
Fritz-Lesch-Str. 29, 13053 Berlin, Germany

JENS HAUEISEN

Friedrich Schiller University, Department of Neurology, Biomagnetic Center Jena,
Erlanger Allee 101, 07747 Jena, Germany

Received 12 April 2005
Revised 23 July 2005
Accepted 25 July 2005
Communicated by Bela Suki

Heart rate exhibits spontaneous fluctuations that are mainly modulated by control loops within the autonomic nervous system. Assessing the dynamics of heart rate fluctuations can provide valuable information about regulatory processes and patho-physiological behavior. In this paper heart rate fluctuations and its entropy are assessed using an algorithmic information theoretic concept applying a data compressor. First, the beat-to-beat fluctuations of heart rate are binary coded for decreases and increases, respectively. Subsequently, those symbol sequences are compressed using the LZ77 algorithm. The ratio of the length of the compressed sequences to the original length is used as an estimate of entropy. We investigated the compressibility of heart rate fluctuations in athletes before, during, and after a training camp. Heart rate time series were obtained from ECGs recorded over 30 minutes under supine resting conditions. We found a significant entropy reduction during the training camp, reflecting the effects of physical fatigue. In conclusion, the compression entropy seems to be a suitable approach to assess the complexity of heart rate fluctuations.

Keywords: Heart rate variability; data compression; entropy; symbolic dynamics; nonlinear dynamics.

1. Introduction

Heart rate exhibits spontaneous fluctuations that are largely modulated by control loops within the autonomic nervous system. Normal heart rate is internally generated via rhythm-generating neurons in the sinus node of the heart, accelerations and decelerations are accomplished via sympathetic and parasympathetic nerves from the brain stem. The investigation of heart rate variability (HRV) provides valuable information about regulatory processes. HRV can be obtained non-invasively from the ECG and is therefore easy to process. Although HRV has already been used for heart diagnosis in traditional Chinese medicine [1], the emergence of digital signal recording/processing tools has stimulated enormously research in clinical science [2]. The analysis of HRV is usually performed using linear signal processing techniques applying statistical measures such as mean and standard deviation and power spectra [3]. There is, however, convincing evidence of the nonlinear nature of HRV. Approaches that have been applied to quantify nonlinear features include symbolic dynamics [4–6], fractal scaling, Lyapunov exponents and dimension measures [7].

In this paper we assess beat-to-beat fluctuations of heart rate. In this context beat-to-beat fluctuations are understood as consecutive heart rate changes within two possible states: increasing or decreasing. To assess whether these beat-to-beat fluctuations contain useful information we estimate their entropy. The introduced methodology is applied to investigate heart rate fluctuations in athletes before, during, and after a two-week training camp. Excessive physical training can lead to the overtraining syndrome (i.e. neural and physical fatigue) which is a serious problem for both professional and recreational athletes [8]. The overtraining syndrome is considered as an imbalance between training/competition versus recovery and is often associated with frequent infections and depressions that occur following hard training and competition. Overreaching, i.e. short-term overtraining, which resolves within two weeks of adequate rest, must be distinguished from long-term overtraining and can be seen as a normal part of athletic training or peaking for performance.

2. Methods

2.1. *Data and pre-processing*

Ten healthy experienced athletes participated in this study. The athletes underwent a two-week training camp that included a daily program of endurance running, cycling, and an exhaustion test on a cycle ergometer. To assess the effects of training on autonomic control, three measurements were carried out: M1 — one week prior to the training camp, M2 — after one week in the training camp, and M3 — after four days of recovery. Exhaustion during M2 was confirmed by psychological tests (Profile of Mood States).

High-resolution ECG (1600 Hz) was recorded from the body surface over 30 minutes under resting conditions in supine position. In this way a quasi-stationary system state is accomplished. Thereafter, time series of beat-to-beat intervals (BBI) were automatically extracted from the ECG using a pattern-matching algorithm based on the cross-correlation function. One of the athletes showed ectopic beats during the measurements (M1: 5 beats; M2: 10 beats; M3: 8 beats) that were filtered out and interpolated based on local variance estimation. Finally, the time series of beat-to-beat fluctuations (S) were obtained by transforming BBI according to Eq.(1).

$$S(n) = \begin{cases} 0 : (\mathrm{BBI}(n) - \mathrm{BBI}(n+1) \leq 0) \\ 1 : (\mathrm{BBI}(n) - \mathrm{BBI}(n+1) > 0) \,. \end{cases} \tag{1}$$

2.2. *Entropy estimation*

2.2.1. *Shannon entropy*

The Shannon entropy, H_s, was computed as the classical measure of information [10]. The formula is given in Eq.(2), where M is the amount of elements in the alphabet of a given text and $p(i)$ denotes the probability of the ith element. H_s characterizes the number of bits that are necessary to code each symbol in a given text on average.

$$H_s = -\sum_{i=1}^{M} p(i) \log_2 p(i) \,. \tag{2}$$

2.2.2. *Compression entropy*

An alternative approach to describe the entropy of a text was introduced in the framework of algorithmic information theory, developed in the 1960s by Kolmogorov, Solomonov and Chaitin [11]. Here, the entropy (also called Kolmogorov–Chaitin complexity, algorithmic complexity) of a given text is defined as the smallest algorithm that is able to generate the text. Although it is theoretically impossible to develop such an algorithm, data compressors might represent a sufficient approximation. In this study we applied the LZ77 algorithm for loss less data compression, introduced by Lempel and Ziv in 1977 [12]. This algorithm is widely used and implemented in many data compressors. The algorithm is based on a sliding window technique (see Appendix). If the length L of the text to be compressed is large ($L \to \infty$) the ratio of compressed to original text length represents the entropy (rate) H_c per symbol. To assess how large L should be in order to obtain a reliable estimate, we compressed sequences with different lengths that were obtained from the M1 recordings, starting from 200 bits up to the maximum data length, with an increment of 200 bits.

3. Results

With increasing sequence length the compressibility increases and seems to approach asymptotically the entropy of the heart rate fluctuations (Fig. 1). Considering the group-averaged graph in the plot, the 30 minutes recording length might be sufficient for the compression entropy estimation. Comparing the compression entropy values with those computed with Shannon's formula it was found that they were fundamentally different (Table 1). H_s was approximately one for all data, i.e. one bit is necessary to code one heart rate fluctuation, suggesting that heart rate fluctuations appear almost randomly. In contrast, entropies computed with the compression technique were notably less than one bit for all measurements, suggesting that there are repetitive structures within heart rate fluctuations. Since one bit of the original data had on average an entropy of $\approx \frac{1}{6}$ bits, the repetitive structures have an average length of approximately six bits. To assess whether there are dominant

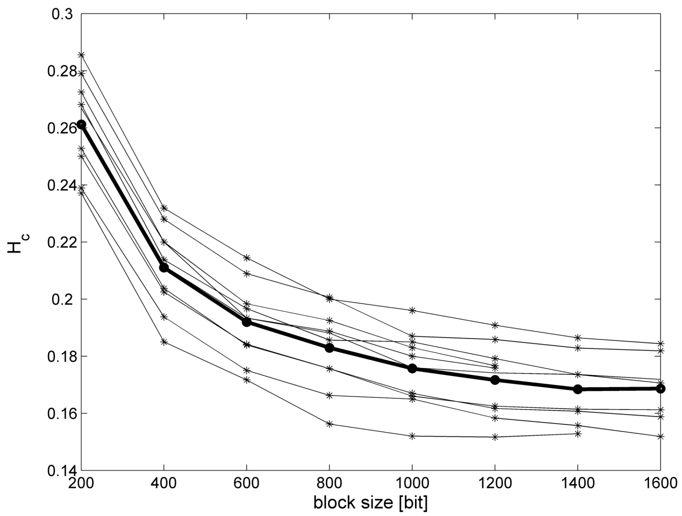

Fig. 1. Compression entropy (H_c) dependent of the length of the compressed binary sequence computed for ten athletes (*thin lines*) and group averaged (*bold line*) for measurement M1.

word types and to re-examine Shannon's approach, we computed additionally the Shannon entropies over six-bit sequences (see Table 1).

The comparison of heart rate fluctuation entropies before, during, and after the training camp revealed a reduction during and after the training period in the compression entropy, but with no changes in the Shannon entropy.

Table 1. Shannon entropy (H_s) and Compression entropy (H_c) of binary-coded heart rate fluctuations as well as Shannon entropy of six-bit word sequences (H_{s6bit}) before (M1), during (M2), and after training camp (M3) presented as median [interquartile range] and Wilcoxon test values for the group comparisons. T1: M1 versus M2; T2: M2 versus M3; T3: M1 versus M3.

	M1	M2	M3	T1	T2	T3
H_s	0.997 [0.997–0.999]	0.999 [0.992–1.000]	0.999 [0.996–1.000]	0.65	0.20	0.45
H_c	0.171 [0.156–0.177]	0.159 [0.151–0.166]	0.159 [0.145–0.166]	0.08	0.48	0.04
H_{s6bit}	5.33 [5.07–5.56]	5.39 [5.04–5.51]	5.23 [4.94–5.55]	0.88	0.51	0.31

4. Discussion

Our study shows that beat-to-beat fluctuations of heart rate are not purely random but follow some deterministic structures. This is consistent with other approaches based on symbolic dynamics [6, 13–15]. The Shannon entropy, however, seems unsuited to assess these heart rate fluctuations since it does not consider any order within the sequence. The estimated entropy value of approximately one is not surprising, considering that the heart rate generating system is stable. To maintain the mean heart rate, accelerations must be compensated with the same amount of heart rate decelerations on average. In contrast to the Shannon formula, the

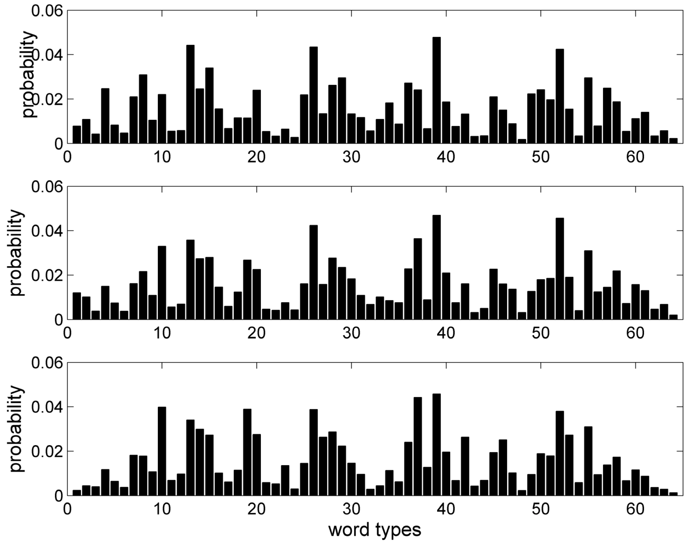

Fig. 2. The group-averaged probability distribution of six bits word types shows similar characteristics before (*top*), during (*middle*), and after (*bottom*) the training camp. The binary words are presented in decimal system for clarity.

compression-based entropy estimation takes also previous values of the system's output into account. Therefore, it is also able to consider some of the dynamics underlying the text-generating system. The LZ77 algorithm applied in this study seems to be suited to map short-term dynamics of heart rate fluctuations within seven consecutive heart beats on average. The histograms of the six-bit word types (see Fig. 2) indicate that certain word types are more likely than others. Consequently, the Shannon entropy over six-bit word types is notably less than six bits. In contrast to the compression entropy, however, the Shannon entropy over those word types was not significantly altered between the three measurements of the training camp study. Thus, the consideration of dynamical matching lengths rather than a simple static coding of six-bit sequences seems to be the valuable feature of the compression entropy. Although the 30-minute recordings seem to give an appropriate estimate of the entropy, this assumption has to be validated with longer recordings.

The entropy reduction of heart rate fluctuations during and after the training camp indicate an effect of the training camp on the autonomous nervous system, i.e. changes in the sympathetic and/or parasympathetic heart rate modulation, that remained even after a few days of recovery. Studies have shown that the early stage of an overtraining syndrome is paralleled by a sympathetic activation [9]. Although no conclusions about an occurring overtraining syndrome can be drawn in the framework of our study, this would suggest that the found entropy reduction reflects a change in sympathetic/parasympathetic heart rate control, probably an increase of the sympathetic influence. Several studies have shown that certain cardiovascular diseases are accompanied by less complex heart rate dynamics applying various techniques [4,16]. Other studies investigating heart rate fluctuations in congestive heart failure patients [17] and heart rate dynamics in healthy controls during exercise [18] have found that the complexity (assessed by detrended fluctuation analysis) increases with suppressed parasympathetic control. Obviously, different analysis techniques assess different signal features and underlying (patho-)physiological mechanisms, respectively, such as the pulsus alternans phenomenon, for example [19]. Future works have to therefore relate the compression entropy with other complexity measures.

In conclusion, the compression entropy is suited to measure heart rate fluctuations and their alterations under different physiological conditions and might be a useful approach for complexity estimation in time series analysis [20].

Acknowledgments

This study was supported by grants from the Deutsche Forschungsgemeinschaft (DFG VO505/4-2), the German Federal Ministry of Education and Research (BMBF – 0312704C) and the German Federal Ministry of Economics and Labor (BMWA – PRO INNO, KF 0317302KLF3).

Appendix LZ77 algorithm

A sequence $\mathbf{x} = x_1, x_2, \ldots$ of length L of a given alphabet $\boldsymbol{\Theta}$ with the amount of $\Phi = |\boldsymbol{\Theta}|$ symbols has to be compressed. Subsequences of $\mathbf{x}$ $(x_m, x_{m+1}, \ldots, x_n)$ are denoted as x_m^n. The algorithm keeps the w recently coded symbols in a memory called a *sliding window*. The symbols to be coded next are stored in a lookahead buffer of size b. The compressor, positioned at p, searches for the longest match of length n between the not-yet-coded string x_p^{p+n-1} in the lookahead buffer and the already-coded string in the sliding window $x_{p-w+\nu}^{p-w+\nu+n-1}$ beginning at position ν. The matching string is therefore fully described by n and ν (i.e. a pointer to the prior occurrence of this sequence).

References

[1] T. O. Cheng, *Decreased heart rate variability as a predictor for sudden death was known in China in the third century A.D.*, *Eur. Heart J.* **21** (2000) 2081–2082.

[2] F. Lombardi, *Clinical implications of present physiological understanding of HRV components*, *Card. Electrophysiol. Rev.* **6** (2002) 245–249.

[3] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, *Heart rate variability. Standards of measurement, physiological interpretation, and clinical use*, *Eur. Heart J.* **17** (1996) 354–381.

[4] J. Kurths, A. Voss, P. Saparin, A. Witt, H. J. Kleiner and N. Wessel, *Quantitative analysis of heart rate variability*, *Chaos* **5** (1995) 88–94.

[5] A. Voss, J. Kurths, H. J. Kleiner, A. Witt, N. Wessel, P. Saparin, K. J. Osterziel, R. Schurath and R. Dietz, *The application of methods of non-linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death*, *Cardiovasc. Res.* **31** (1996) 419–433.

[6] M. Baumert, T. Walther, J. Hopfe, H. Stepan, R. Faber and A. Voss, *Joint symbolic dynamic analysis of heart rate and blood pressure interactions in normal pregnancy*, *Med. Biol. Eng. Comput.* **40** (2002) 241–245.

[7] A. L. Goldberger, *Non-linear dynamics for clinican: Chaos theory, fractals, and complexity at the bedside*, *Lancet* **347** (1996) 1312–1314.

[8] R. Budgett, *Fatigue and underperformance in athletes: The overtraining syndrome*, *Brit. J. Sports Med.* **32** (1998) 107–110.

[9] M. Lehmann, C. Foster, H. Dickhuth and U. Gastmann, *Autonomic imbalance hypothesis and overtraining syndrome*, *Med. Sci. Sports. Exerc.* **30** (1998) 1140–1145.

[10] C. E. Shannon, *A mathematical model of communication*, *The Bell System Technical J.* **27** (1948) 379–423, 623–656.

[11] M. Li and P. Vitnyi, *An introduction to Kolmogorov complexity and its applications*, Springer (1997).

[12] A. Lempel and A. Ziv, *Universal algorithm for sequential data-compression*, *IEEE Trans. Inf. Theor.* **23** (1977) 337–343.

[13] A. C. Yang, S. S. Hseu, H. W. Yien, A. L. Goldberger and C. K. Peng, *Linguistic analysis of the human heartbeat using frequency and rank order statistics*, *Phys. Rev. Lett.* **90** (2003) 108103.

[14] V. Baier, M. Baumert, P. Caminal, M. Vallverdu, R. Faber and A. Voss, *Hidden Markov models based on symbolic dynamics for statistical modeling of cardiovascular control in hypertensive pregnancy disorders*, *IEEE Trans. Biomed. Eng.* in press.

[15] M. Baumert, V. Baier, A. Schirdewan, S. Truebner and A. Voss, *Short- and long-term joint symbolic dynamics of heart rate and blood pressure in dilated cardiomyopathy*, *IEEE Trans. Biomed. Eng.* **52** (2005) 2112–2115.

[16] M. G. Signorini, S. Cerruti, S. Guzzetti, R. Parola, *Non-linear dynamics of cardiovascular variability signals*, *Methods Inf. Med.* **33** (1994) 81–84.

[17] Y. Ashkenazy, P. C. Ivanov, S. Havlin, C. K. Peng, A. L. Goldberger and H. E. Stanley, *Magnitude and Sign Correlations in Heartbeat Fluctuations*, *Phys. Rev. Lett.* **86** (2001) 1900–1903.

[18] R. Karasik, N. Sapir, Y. Ashkenazy, P. C. Ivanov, I. Dvir, P. Lavie and S. Havlin, *Correlation differences in heartbeat fluctuations during rest and exercise*, *Phys. Rev. E.* **66** (2002) 062902.

[19] U. Leder, H. P. Pohl, V. Baier, M. Baumert, M. Liehr, J. Haueisen, A. Voss and H. R. Figulla, *Alternans of heart rate and blood pressure in dilated cardiomyopathy*, *PACE* **25** (2002) 1307–1314.

[20] M. Baumert, V. Baier, J. Haueisen, N. Wessel, U. Meyerfeldt, A. Schirdewan and A. Voss, *Forecasting of life threatening arrhythmias using the compression entropy of heart rate*, *Methods Inf. Med.* **43** (2004) 202–206.