Biomedical Signal Processing and Control 23 (2016) 28–41

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Improved multiscale permutation entropy for biomedical signal analysis: Interpretation and application to electroencephalogram recordings

Hamed Azami*, Javier Escudero

Institute for Digital Communications, School of Engineering, University of Edinburgh, King's Buildings, Edinburgh EH9 3JL, UK

ARTICLE INFO

Article history:
Received 22 January 2015
Received in revised form 26 June 2015
Accepted 5 August 2015
Available online 25 August 2015

Keywords:
Multiscale permutation entropy
Biomedical signal analysis
Electroencephalogram
Non-linear analysis
Signal regularity
Complexity

ABSTRACT

Permutation entropy (PE) is a well-known and fast method extensively used in many physiological signal processing applications to measure the irregularity of time series. Multiscale PE (MPE) is based on assessing the PE for a number of coarse-grained sequences representing temporal scales. However, the stability of the conventional MPE may be compromised for short time series. Here, we propose an improved MPE (IMPE) to reduce the variability of entropy measures over long temporal scales, leading to more reliable and stable results. We gain insight into the dependency of MPE and IMPE on several straightforward signal processing concepts which appear in biomedical activity via a set of synthetic signals. We also apply these techniques to real biomedical signals via publicly available electroencephalogram (EEG) recordings acquired with eyes open and closed and to ictal and non-ictal intracranial EEGs. We conclude that IMPE improves the reliability of the entropy estimations in comparison with the traditional MPE and that it is a promising technique to characterize physiological changes affecting several temporal scales. We provide the source codes of IMPE and the synthetic data in the public domain.

1. Introduction

There are several main types of measures, such as, entropies and fractal dimensions, to compute the complexity of a system or signal. These are used to compare signals and distinguish or detect regular and random epochs [1]. As mentioned in [2], healthy subjects and people with disease can often be distinguished by the complexity of their physiological activity [3].

Entropy is one of the most popular and powerful concepts to evaluate the dynamical characteristics of a signal. This metric measures the uncertainty and irregularity of a time series. Higher entropy generally demonstrates higher uncertainty, whereas lower entropy shows more regularity and certainty of a system [1,4]. There are a number of entropy approaches commonly applied to physiological recordings, such as approximate entropy (ApEn) [5], sample entropy (SaEn) (Richman and Moorman 2000), fuzzy entropy (FuEn) (Chen, Wang et al., 2007), permutation entropy (PE) [1] and wavelet entropy (Rosso, Blanco et al., 2001), each of which has its own advantages and disadvantages [6].

PE is based on the order relations among values of a signal, the permutation patterns. It is analogous to the Lyapunov exponents for some well-known chaotic dynamical systems, such as the noise-free logistic map, although PE yields more meaningful results in the presence of observational and dynamical noise [1,7]. Compared with the other entropies, such as ApEn and FuEn, PE is theoretically simple and it has fewer parameters, it is relatively robust to artifacts and noise, and is computationally fast. Furthermore, the PE can be used for both the non-stationary and nonlinear signals. With respect to the signal length, PE is more robust than the zero-crossing rate (ZCR) [1]. Because of the aforementioned advantages, PE has been extensively employed in the numerous real world physiological signal and image processing applications [7,8]. For example, Li et al. investigated the behavior of PE to predict absence seizures in rats using EEG signals [9]. They showed that PE can track the dynamical changes of EEG recordings and that PE can predict absence seizures better than SaEn [9]. Ferlazzo et al. employed PE to reveal abnormalities of cerebral activity in patients with typical absences [10]. They concluded that PE is a valuable tool to detect abnormalities of cerebral electrical activity which are not revealed by conventional approaches for EEG signals [10].

However, PE is limited to assessing the values of entropy for only one temporal scale, the one associated with the original sampling of the signals. This may limit the ability of PE to inspect dynamics

* Corresponding author.
E-mail addresses: hamed.azami@ed.ac.uk (H. Azami), javier.escudero@ed.ac.uk (J. Escudero).

http://dx.doi.org/10.1016/j.bspc.2015.08.004
1746-8094/

residing at longer temporal scales. In this sense, multiscale entropy (MSE), proposed by Costa et al. [3,11], calculates entropy over a range of scales to evaluate the complexity of a time series. In the original definition of MSE, SaEn was the metric used to assess the entropy over the temporal scales [3,11]. Nonetheless, the concept of multiscale evaluation of entropy can be extended to other entropy metrics. Morabito et al. used multiscale PE (MPE) to assess the complexity of electroencephalogram (EEG) recordings in Alzheimer's disease [8].

The multiscale evaluation of entropy has notorious advantages. First of all, it allows us to inspect dynamics along more than one temporal scale. This is very significant for biological systems which need to operate across multiple spatial and temporal scales, and therefore their complexity is also multiscaled [3,12]. Secondly, unlike SaEn, MSE is consistent with the Fogedby study [13] illustrating that the complexity of $1/f$ noise is higher than white Gaussian noise (WGN).

The coarse-graining of MPE and MSE methods are based on Costa's algorithm [3,8,11]. The first step of MPE, the coarse-graining process, considerably reduces the time series length because, to inspect the deeper temporal scales, MPE uses a procedure similar to sub-sampling. This may yield an imprecise estimation of entropy when the time series is too short. Hence, the basic MPE method may not provide a reliable analysis for short time series. To overcome this problem, in this paper, an improved MPE (IMPE) is proposed. This is in contrast with the alternative MPE algorithm by [14] called modified MPE (MMPE). In the MMPE, a coarse-grained sequence is built subsampling the original signal by taking one out of τ samples, where τ denotes the temporal scale. However, no filtering is used. Therefore, this procedure will necessarily lead to aliasing, thus changing important properties of the signal. For example, for $\tau = 3$ and $i = 2$, $y^{(3)}(2,j)=\{x_2,x_5,x_8,...\}$, some important information of the original time series $\{x_1,x_2,...,x_N\}$ may be omitted in the MMPE.

Because of the relevance and the possible usefulness of MPE and IMPE in a number of biomedical signal analyses, it is important to understand and exemplify the behavior of the measure for different kinds of classical signal concepts such as frequency, amplitude, noise power, and signal bandwidth. This study addressed this issue to help to illustrate the dependency of both MPE and IMPE on these concepts and to compare both techniques. Moreover, we will illustrate the application of MPE and IMPE to five different datasets of real EEG signals.

In the following section, the concepts of PE and MPE are described. Our proposed method is explained in Section 3. In Section 4, the synthetic signals and real EEG datasets employed in this paper are introduced. Then, the results and discussions of the proposed method (IMPE) and the conventional MPE are explained in Sections 5 and 6, respectively. The conclusions of the paper are drawn in the last section.

2. Background on PE and MPE

In this section, we briefly describe PE and MPE.

2.1. Permutation entropy

Assume we have a given time series of length N, and let the time series be $y = \{y_1,y_2,...,y_N\}$. At each time t of $\mathbf{y}$, a vector including the dth subsequent values is constructed as: $Y_t^{d,l}=\{y_t,y_{t+1},...y_{t+(d-2)l},y_{t+(d-1)l}\}$ for $t = 1,2,\ldots,N-(d-1)l$, where d, which is named the embedding dimension, determines how much information is contained in each vector and l is the time delay. To calculate the PE, the d values y_i are associated with numbers from 1 to d and arranged in increasing order as $\{y_{t+(j_1-1)l}, y_{t+(j_2-1)l}, \ldots, y_{t+(j_{d-1}-1)l}, y_{t+(j_d-1)l}\}$. For different samples, there will be $d!$ potential ordinal patterns, π, which are named "motifs". For each π_t, $p(\pi_t)$ demonstrates the relative frequency as follows:

$$p(\pi_i^{d,l}) = \frac{\#\{t \mid t \le N-d, \text{type}(Y_t^{d,l}) = \pi_i^{d,l}\}}{N-d+1} \quad (1)$$

where $\#\{\}$ denotes the cardinality of the set (the number of elements) [1,7]. The PE is computed as follows:

$$H(y,d,l) = -\sum_{\pi_k=1}^{\pi_k=d!} p(\pi_k)\ln p(\pi_k) \quad (2)$$

When all motifs have equal probability, the largest value of PE is obtained, which has a value of $\ln(d!)$. In contrast, if there is only one $p(\pi_k)$ different from zero, which illustrates a completely regular signal, the smallest value of PE is obtained as much as 0 [1,7].

2.2. Multiscale permutation entropy

MPE, like MSE, includes two main steps. First, a "coarse-graining" process is applied to a time series. Consider a real-valued time series $\{x_1,x_2,...,x_N\}$ of length N. Multiple successive coarse-grained versions are made by averaging the time data points within non-overlapping windows of increasing length τ, which is called scale factor. A schematic illustration of the coarse-grained procedure is shown in Fig. 1. According to the following equation, each element of the coarse-grained time series $y_j^{(\tau)}$ is defined as:

$$y_j^{(\tau)} = \frac{1}{\tau}\sum_{i=(j-1)\tau+1}^{j\tau} x_i \quad 1 \le j \le \left\lfloor \frac{N}{\tau} \right\rfloor \quad (3)$$

where $\lfloor a \rfloor$ denotes the largest integer not greater than a. The length of each coarse-grained time series is $\left\lfloor \frac{N}{\tau} \right\rfloor$. Second step is calculating the PE for each coarse-grained time series. The attained values can be plotted as a function of the scale factor τ [3,8,11].

3. Improved multiscale permutation entropy

The conventional MPE has two main drawbacks. Firstly, the MPE is not symmetric. For example in scale 3, we could rationally expect the metric to behave the same for x_3 and x_4, in comparison with x_2 and x_3. However, at scale 3, x_1, x_2 and x_3 are separated from x_4, x_5 and x_6. The second drawback is the relative variability of the MPE results for long temporal scales. When the MPE is computed, in the coarse-graining process, the number of samples of the resulting coarse-grained sequence is $\left\lfloor \frac{N}{\tau} \right\rfloor$. When the scale factor τ is high, the number of samples in the coarse-grained sequence decreases. This may yield an unstable measure of entropy.

To overcome these problems, the IMPE is proposed based on the idea originally reported by Wu for MSE [15]. Here, because of some advantages of PE over SaEn, we use PE instead. Hence, the IMPE is calculated in two main steps:

1) In the first step, $z_i^{(\tau)} = \{y_{i,1}{}^{(\tau)}, y_{i,2}{}^{(\tau)}, \ldots\}$ are generated where $y_{i,j}{}^{(\tau)} = \frac{\sum_{f=0}^{\tau-1} x_{f+i+\tau(j-1)}}{\tau}$ As can be observed in Fig. 2, in the IMPE algorithm, for each τ, we have τ different time series $z_i^{(\tau)} | (i = 1, \ldots, \tau)$, while in the MPE method, only $Z_1(\tau)$is considered.
2) For a defined scale factor τ and embedding dimension d, PE of each of $z_i^{(\tau)} | (i = 1, \ldots, \tau)$ is separately calculated. Then, the average of PE values is computed as follows:

$$\text{IMPE}(x,\tau,d) = \frac{1}{\tau}\sum_{i=1}^{\tau} \text{PE}(z_i^{(\tau)}) \quad (4)$$