Biomedical Signal Processing and Control 39 (2018) 424–430

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Unbiased estimation of permutation entropy in EEG analysis for Alzheimer's disease classification

Lucie Tylová [a,*], Jaromír Kukal [a], Václav Hubata-Vacek [a], Oldřich Vyšata [b,c]

[a] CTU in Prague, Faculty of Nuclear Sciences and Physical Engineering, Department of Software Engineering, Břehová 7, 115 19 Prague 1, Czech Republic
[b] ICT in Prague, Faculty of Chemical Engineering, Department of Computers and Control Engineering, Technická 5, 160 00 Prague 6, Czech Republic
[c] Charles University in Prague, Faculty of Medicine in Hradec Králové, Department of Neurology, Simkova 870, 500 38 Hradec Králové 1, Czech Republic

ARTICLE INFO

Article history:
Received 7 October 2016
Received in revised form 13 June 2017
Accepted 1 August 2017
Available online 30 August 2017

Keywords:
EEG
Alzheimer's disease
Permutation entropy
Unbiased estimation
Hash table
Resampling

ABSTRACT

The EEG signal of healthy patient can be recognized as an output of a chaotic system. There are many measures of chaotic behaviour: Hurst and Lyapunov exponents, various dimensions of attractor, various entropy measures, etc. We prefer permutation entropy of equidistantly sampled data. The novelty of our approach is in bias reduction of permutation entropy estimates, memory decrease, and time complexities of permutation analysis. Therefore, we are not limited by the EEG signal and permutation sample lengths. This general method was used for channel by channel analysis of Alzheimer's diseased (AD) and healthy (CN) patients to point out the differences between AD and CN groups. Our technique also enables to study the influence of EEG sampling frequency in a wide range. The best results were obtained for sampling frequency 200 Hz, using at most window of length 10. In the case of Alzheimer's disease, we observed a statistically significant decrease in permutation entropy at all channels.

1. Introduction

The diagnosis of Alzheimer's disease (AD) is an up-to-date problem which is solved by various techniques. Non-invasive investigation of AD patients is preferred and both magnetic resonance imaging (MRI) and electroencephalography (EEG) are frequently used for image and signal processing related to the diagnosis. Being focused on EEG analysis, spectral analysis is one of the successful tools for AD investigation; however, the analysis of non-linear EEG dynamic offers a more complex signal analysis. Correlation dimension D_2 is a good example of complexity measure offering lower values for AD patients [1]. The same trend is seen for the largest Lyapunov exponent λ. Unfortunately, the calculation of D_2 and λ_1 enforces the analysis of a very long time series [2]. In this study, we prefer entropy estimation from the EEG signal [3,4] because of the relationship between the complexity of non-linear signals [5,6] and entropy estimate. Dauwels et al. [7] and many other authors have shown that Alzheimer's disease increases power in the delta and theta-bands in frequency domain, but the power spectrum is a global characteristics of the EEG signal making it impossible to study and localise events in the signal. There are many possibilities how to organize the entropy evaluation. Our approach is based on permutation entropy [8,9]. There are three reasons why to use permutation entropy for diagnosing AD. Human brain activity can be interpreted as behaviour of a complex chaotic system. Therefore, signals from EEG electrodes carry information about the chaos inside the scalp. The hypothesis of decreasing chaotic behaviour [3] during AD can be tested by using permutation entropy; the second reason for using it is its relative simplicity. This term is exactly defined for the windowed signal and is directly applicable to time series without any methodological difficulties. The time complexity of permutation entropy calculation is acceptable for small windows lengths as discussed in [8,9]. However, the application to windows of size above 12 is impossible. Permutation entropy has not yet been applied to longer windows. We suppose there is a chance to obtain new and significant results if the window length were prolonged. Using hash table as a special data structure, it is possible to calculate permutation entropy for window lengths up to 30, which is the main novelty of our approach, as will be explained in following sections.

2. Permutation entropy

2.1. Shannon entropy and its estimation

Definition. Shannon entropy [3,10] H_S of a discrete random variable X with possible values $x_1, \ldots, x_m$ and probability mass function $p(X)$ is defined as

* Corresponding author.
E-mail address: tylovluc@fjfi.cvut.cz (L. Tylová).

http://dx.doi.org/10.1016/j.bspc.2017.08.012
1746-8094/

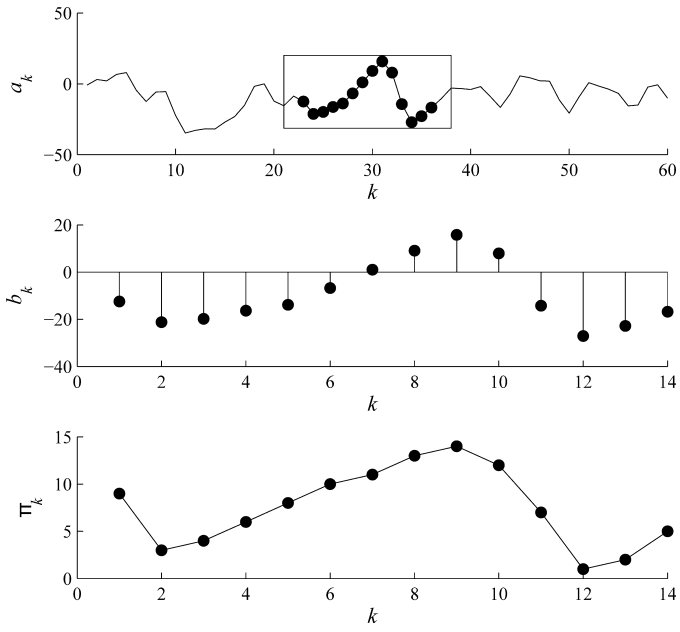

Fig. 1. Permutation analysis: Original EEG (top), windowed $w = 14$ (middle), permutation pattern (bottom).

$$H_S = -\sum_{i=1}^{m} p_i \ln p_i, \tag{1}$$

where $p_i = p(x_i)$.

If the probability function is unknown for an experimental data set, and the number of possible values is finite for random variable X, we estimate probability function p_i by relative frequency $p_{j,N}$ and number of events k_N as

$$p_{j,N} = \frac{n_j}{n}, \tag{2}$$

$$k_N = \sum_{n_j>0} 1 \le k, \tag{3}$$

where n_j is the number of occurrences x_i of random variable X, and n the total number of measurement results. Then we get the *naive estimate* of Shannon entropy as

$$H_N = -\sum_{j=1}^{k_N} p_{j,N} \ln p_{j,N}. \tag{4}$$

This estimate is biased, and therefore it has a systematic error.

Miller [11] modified *naive estimate* H_N using first order Taylor expansion, which produces better estimation

$$H_M = H_N + \frac{k_N - 1}{2n}. \tag{5}$$

2.2. Application to permutation analysis

Entropy estimates can be easily applied to permutation event analysis [8,9]. The methodology from [11] estimates a smaller bias. Let time series be $\{a_k\}\, k = 1_T$ and sliding window $\{b_k\}\, k = 1_w$ of length w, then we can substitute signal values b_k in the window with their orders and then obtain permutation pattern $\{\pi_k\}\, k = 1_w$. The process of pattern conversion is depicted in Fig. 1 for original EEG data 200 Hz-sampled for $w = 14$ and the 8th channel of CN patient.

The universe of random variable X is a set of all permutations of length w. Therefore, the number of possible permutations is

$$m = w!; \tag{6}$$

however, the number of various permutations in a given signal cannot exceed the number of sliding samples as

$$k_n \le n = T - w + 1. \tag{7}$$

The number of occurrences of jth permutation pattern corresponds to n_j, and n is the total number of samples. The differences between typical AD and CN patients are illustrated in Fig. 2. Supposing ordering $n_{(j)} \ge n_{(j+1)}$ for $j = 1, \ldots, m-1$, ten most frequent permutation patterns $(n_{(1)}, \ldots, n_{(10)})$ were plotted to the union diagram for $f_s = 200$ Hz, $w = 14$, $ch = 8$. In the case of AD, we observed a systematic gradual increase or decrease in the EEG signal with a small fraction of exceptions. In the case of CN, however, the patterns rarely increased, no systematic decreasing was observed, and the EEG signal exhibited higher diversity. Therefore, this primary observation is in agreement with hypothesis of diminished EEG signal entropy in the case of AD.

So now, we can directly use (4) and calculate the biased naive estimation H_N as in [10]. Our methodology is based on Miller's approach [11] and direct application of (5) to permutation patterns. The difference between estimates (4) and (5) varies according to the number of distinct patterns and time series length.

3. Permutation analysis for large samples

The main disadvantage of the original procedure of permutation analysis [8] is in its memory and time complexities. The authors implemented permutation memory as a matrix of w columns and $w!$ rows together with counter vector of length $w!$. This enables permutation analysis on a typical computer only for $w < 13$. Traditional applications [8] of permutation entropy use a window of length $w < 8$. The time complexity of single permutation counting is also $w!$, in the worst case. Therefore, for permutation analysis we decided to use more sophisticated data structure. There are many data structures and algorithms for realizing a *look-up table* as a kind of memory with fast access. Our memory has to be optimized only for two operations: FIND and INSERT. We used *hash table* with open addressing and linear probe strategy [12] as a model which is easy to realize. Let $P > n$ be the optional prime number. Then the *loading factor* is defined as a ratio $\alpha = n/P < 1$. The mean number of permutation vector comparisons during the successful FIND operation was determined [12] as

$$ET_{OPT} = \frac{1}{2}\left(1 + \frac{1}{1-\alpha}\right). \tag{8}$$

For unsuccessful FIND operation and INSERT operation, the mean number of permutation vector comparisons is higher [12] than in the previous optimistic case

$$ET_{PES} = \frac{1}{2}\left(1 + \frac{1}{(1-\alpha)^2}\right). \tag{9}$$

Our tiny and fast implementation of permutation memory is a matrix of occurred permutations with w columns and $P > n$ rows together with counter vector of length P. In the best (8) and the worst (9) cases, the time complexity of single permutation counting is constant and dependent only on the loading factor. This enables very fast permutation analysis for higher sample length w and long EEG sequences. The last implementation detail is how to realize hash function $index = h(\boldsymbol{\pi})$ for the given permutation pattern $\boldsymbol{\pi}$. In the first step, by subtracting the vector of units from vector $\boldsymbol{\pi}$, we obtain digital form $\boldsymbol{y} = \boldsymbol{\pi} - 1$. Let $R = w$ be the base of a digital system. In the second step, we calculate the value v of $\boldsymbol{y}$