Computers in Biology and Medicine 87 (2017) 141–151

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Noisy EEG signals classification based on entropy metrics. Performance assessment using first and second generation statistics

David Cuesta–Frau [a,*], Pau Miró–Martínez [b], Jorge Jordán Núñez [b], Sandra Oltra–Crespo [a], Antonio Molina Picó [a]

[a] Technological Institute of Informatics, Polytechnic University of Valencia, Alcoi Campus, Plaza Ferrandiz y Carbonell 2, Alcoi, Spain
[b] Department of Statistics, Polytechnic University of Valencia, Alcoi Campus, Alcoi, Spain

ARTICLE INFO

Keywords:
Electroencephalograms
Signal classification
Approximate Entropy
Sample Entropy
Fuzzy Entropy
EEG artifacts

ABSTRACT

This paper evaluates the performance of first generation entropy metrics, featured by the well known and widely used Approximate Entropy (ApEn) and Sample Entropy (SampEn) metrics, and what can be considered an evolution from these, Fuzzy Entropy (FuzzyEn), in the Electroencephalogram (EEG) signal classification context. The study uses the commonest artifacts found in real EEGs, such as white noise, and muscular, cardiac, and ocular artifacts. Using two different sets of publicly available EEG records, and a realistic range of amplitudes for interfering artifacts, this work optimises and assesses the robustness of these metrics against artifacts in class segmentation terms probability. The results show that the qualitative behaviour of the two datasets is similar, with SampEn and FuzzyEn performing the best, and the noise and muscular artifacts are the most confounding factors. On the contrary, there is a wide variability as regards initialization parameters. The poor performance achieved by ApEn suggests that this metric should not be used in these contexts.

1. Introduction

Electroencephalography is a very important medical monitoring technique based on recording and analysing the brain's electrical activity. These recordings are termed electroencephalograms (EEGs), and are usually obtained non invasive by placing electrodes on the surface of scalps. The resulting time series can then be used to study the electrical activity of different brain regions and their correlation with clinical variables [1]. This analysis, performed by skilled operators using classical signal processing algorithms, was successfully used to assess a multitude of brain disorders, damage or processes.

For example, the authors in Ref. [2] propose a method based on the EEG power spectrum to estimate users' level of alertness while they performed critical tasks. Similarly [3], report a method to classify states of fatigue and alertness while driving. Another field of extensive research is the assessment of sleep or anesthesia depth. In Rodriguez et al. [4], the authors describe an unsupervised sleep stages classification method based on pattern recognition techniques and a feature optimisation algorithm. EEG has also been used to evaluate the brain function after a stroke. The study [5] proposes a dense–array EEG to capture stroke effects, with a high correlation with the NIH stroke scale by partial least squares modelling. EEG and different types of dementia form another very active field of research. In Ref. [6], the authors carried out a meta–analysis based on 4157 papers to assess the correlation between abnormal EEGs and early–onset dementia (EOD). A clear relationship was found and demonstrated the capability of EEG to become a reliable tool for EOD diagnosis and prognosis. EEG analysis and processing can also contribute significantly to diagnosing and managing epilepsy [7] with a number of specific applications, such as seizure type determination or identification of epileptogenic regions, among many more.

However, not all the information provided by EEGs can be directly extracted because some information may be buried far down in the dynamics of the time series itself. In order to place this information within reach of the understanding of physicians, it is necessary to implement advanced mathematical methods and algorithms that extract additional subclinical information efficiently and expeditiously [8]. In line with this, one of the most successful groups of tools is the time series entropy estimation methods.

A diverse varied collection of these methods has been proposed in the last few decades, including Approximate Entropy, Sample Entropy, Fuzzy Entropy, Lempel–Ziv complexity, Permutation Entropy, Distribution Entropy, Renyi Entropy, Detrended Fluctuation Analysis, and some others, with a broad range of capabilities and applications in mainly economy and medicine. Specifically, in the field of EEG processing, two

* Corresponding author.
E-mail address: dcuesta@disca.upv.es (D. Cuesta–Frau).

http://dx.doi.org/10.1016/j.compbiomed.2017.05.028
Received 14 February 2017; Received in revised form 5 May 2017; Accepted 28 May 2017

of the most widely used and successful entropy estimators are Approximate Entropy (ApEn) [9] and Sample Entropy (SampEn) [10], with hundreds of studies in the scientific literature.

ApEn quantifies the similarity probability of patterns of length m and $m+1$. Unlike other previous non linear methods, ApEn has demonstrated its robustness against noise and its capability to detect complexity changes using finite size datasets, and has provided at least 1000 data values whenever available [9]. By using similarity threshold r, defined as a fraction of the standard deviation of the input data, ApEn is also scale-independent. ApEn has been used to find EEG differences in schizophrenia patients [11], with lower entropy values obtained for these patients, or in comatose patients [12]. A significant number of studies has assessed anaesthesia depth where ApEn was the chosen tool, e.g., the work described in Ref. [13]. In that study, the ApEn metrics was able to track EEG changes in different anesthesia stages. Other research works have focused on measuring the effects of specific treatments or therapies on a range of neurological conditions through quantifiable changes in EEG. For example, the authors in Ref. [14] investigated the effect of current stimulation on aphasic patients. EEG changes due to aging or sleep have also been assessed using ApEn, as in Ref. [15], where ApEn was able to distinguish consciousness levels, and to find differences between age groups.

SampEn is a similar statistic. It also measures the probability of subsequences being close at two lengths m and $m+1$ within tolerance r. However, SampEn does not include self–comparisons and exhibits greater consistency than ApEn [16]. The algorithm to compute SampEn is also faster than that of ApEn, but its execution time is still $O(N^2)$, with N being the length of the time series [17]. SampEn has not yet been used as extensively as ApEn as this was proposed later, but it is quickly catching up given its better performance. The scope of application is very similar to that of ApEn. So, there are works that have studied EEG differences between control subjects and individuals with traumatic brain injury [18]. Sleep stages have also been classified using SampEn, as in Refs. [19,20]. Alzheimer screening using EEG and SampEn is another promising area of research with already significant results [21].

ApEn and SampEn are very successful data entropy estimators, but they also have their weaknesses. As stated above, ApEn is biased since it includes self–matches in the count, and SampEn requires a relatively large r to find similar subsequences and to avoid the $\log(0)$ problem (Table 1). They are also very sensitive to input parameters m, r, and N. More recently, an evolution of these metrics, Fuzzy entropy (FuzzyEn), has been proposed to mitigate these problems [22]. FuzzyEn is based on a continuous function to compute the dissimilarity between two zero–mean subsequences and, consequently, it is more stable in noise and parameter initialisation terms. This metrics is still scarcely used in EEG studies, but it is expected to replace ApEn and SampEn because of its excellent stability, mainly when applied to noisy or short records. At present, very few studies have already demonstrated its capability to detect epileptic seizures [23], EEG abnormalities in Alzheimer's disease [24], or in recognizing wake or sleep stages [25,26].

ApEn and SampEn have played, or are playing, a very important role in unveiling hidden information in EEGs, and will still be used for some time unless a more efficient metrics, such as FuzzyEn, completely replaces these older methods. To distinguish between these two generations of metrics, those initially proposed, even decades ago, and those proposed less than 5 years ago as an evolution or improvement of the initial ones, we coined the terms first- and second-generation metrics, which will be used throughout this paper.

Signal classification efficiency is often assessed in relation to more robustness against difficult processing conditions: class separability, initialisation dependence, data size or noise. This paper focuses specifically on the effect on entropy metrics of EEG signals noise. Biomedical records are often corrupted with artifacts and noise, and EEGs are no exception. In general, biomedical record interferences can be of a physiological (EEGs are corrupted with data from other biosignals) or technical (EEGs are corrupted with noise generated by acquisition or other nearby systems) origin, with a myriad of methods to remove, or at least, reduce these artifacts proposed in the scientific literature [27–30]. However, this is not always possible: signal and artifacts overlap in time and/or frequency domains (they cannot be removed without degrading the underlying valid signal), there is a high computational cost or complexity of the required algorithms, and the parameter optimization needs of filtering or cancelling methods cannot be addressed due to lack of time or resources.

As a result, a certain level of interference should be expected in any EEG signal, and the methods applied must therefore be robust against it. The present study addresses this issue by assessing of the performance of the above cited methods, ApEn, SampEn, and FuzzyEn, in the noisy EEG signal classification context. Specifically, we analyse the influence of the commonest physiological artifacts in EEG records: ocular artifacts [31], cardiac artifacts [32] and muscular artifacts [33]. The study also includes technical artifacts, such as noise and spikes [34]. The objective of the study is to improve the understanding of the metrics' behaviour under real conditions, and to provide practical advice about optimal performance.

The methodology employed is based on quantitative research. The analysis involves the collection of labelled EEG data, considered as the ground truth, since they do not contain artifacts (intra-cranial visually inspected EEGs), and apply a correlational research to find differences among the three entropy metrics studied (ApEn, SampEn, and FuzzyEn), based on a statistical treatment. The ultimate goal is to support or refute the robustness against artifacts hypothesis of each one of the metrics.

Table 1
Mathematical definition of ApEn, SampEn, and FuzzyEn ($\mu(d, r)$: Fuzzy membership function).

	ApEn(m, r, N)	SampEn(m, r, N)	FuzzyEn(m, r, N)
1) Create a set of subsequences of length m	$\mathbf{x}_i = \{x_i, x_{i+1}, \ldots, x_{i+m-1}\} i = 1, \ldots, N-m+1$	$\mathbf{x}_i = \{x_i, x_{i+1}, \ldots, x_{i+m-1}\} i = 1, \ldots, N-m+1$	$\mathbf{y}_i = \{x_i, x_{i+1}, \ldots, x_{i+m-1}\} \overline{y}_i = \text{mean}(\mathbf{y}_i)$ $\mathbf{x}_i = \{x_i - \overline{y}_i, x_{i+1} - \overline{y}_i, \ldots, x_{i+m-1} - \overline{y}_i\}$ $i = 1, \ldots, N-m+1$
2) Dissimilarity computation	$d_{ij} = \max(\lvert x_{i+k} - x_{j+k} \rvert), 0 \le k \le m-1$	$d_{ij} = \max(\lvert x_{i+k} - x_{j+k} \rvert), 0 \le k \le m-1, j \ne i$	$d_{ij} = \max(\lvert x_{i+k} - x_{j+k} \rvert)$, $D_{ij} = \mu(d_{ij}, r), 0 \le k \le m-1, j \ne i$
3) Count matches	$B_i(r)$ no. of j so that $d[X_m(i), X_m(j)] \le r$ $A_i(r)$ no. of j so that $d[X_{m+1}(i), X_{m+1}(j)] \le r (1 \le j \le N-m+1)$	$B_i(r)$ no. of j so that $d[X_m(i), X_m(j)] \le r A_i(r)$ no. of j so that $d[X_{m+1}(i), X_{m+1}(j)] \le r (1 \le j \le N-m, j \ne i)$	$\phi_i^m(r) = \frac{1}{N-} m-1 \sum_{j=1, j \ne i}^{N-m} D_{ij}^m$
4) Statistics	$B_i^m(r) = \frac{1}{N-} m+1 B_i(r) A_i^m(r) = \frac{1}{N-} m A_i(r)$ $\phi^m(r) = \frac{1}{N-} m+1 \sum_{i=1}^{N-m+1} \log B_i^m(r)$ $\phi^{m+1}(r) = \frac{1}{N-} m \sum_{i=1}^{N-m} \log A_i^m(r)$	$B_i^m(r) = \frac{1}{N-} m-1 B_i(r) B^m(r) = \frac{1}{N-} m \sum_{i=1}^{N-m} B_i^m(r)$ $A_i^m(r) = \frac{1}{N-} m-1 A_i(r) A^m(r) = \frac{1}{N-} m \sum_{i=1}^{N-m} A_i^m(r)$	$\varphi^m(r) = \frac{1}{N-} m \sum_{i=1}^{N-m} \phi_i^m(r)$ $\varphi^{m+1}(r) = \frac{1}{N-} m \sum_{i=1}^{N-m} \phi_i^{m+1}(r)$
5) Result	$\text{ApEn}(m, r) = \lim_{N \to \infty} [\phi^m(r) - \phi^{m+1}(r)]$ $\text{ApEn}(m, r, N) = [\phi^m(r) - \phi^{m+1}(r)]$	$\text{SampEn}(m, r) = \lim_{N \to \infty} \left(-\log \left[\frac{A^m(r)}{B^m(r)} \right] \right)$ $\text{SampEn}(m, r, N) = -\log \left[\frac{A^m(r)}{B^m(r)} \right]$	$\text{FuzzyEn}(m, r) = \lim_{N \to \infty} [\log\varphi^m(r) - \log\varphi^{m+1}(r)]$ $\text{FuzzyEn}(m, r, N) = [\log\varphi^m(r) - \log\varphi^{m+1}(r)]$