Biomedical Signal Processing and Control 40 (2018) 312–317

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Short communication

An accurate system to distinguish between normal and abnormal electroencephalogram records with epileptic seizure free intervals

Salim Lahmiri

Montreal Neurological Institute, McGill University, Montreal, Canada

ARTICLE INFO

Article history:
Received 3 March 2017
Received in revised form 7 September 2017
Accepted 7 October 2017

Keywords:
Electroencephalographic
Epilepsy
Seizure
Generalized hurst exponent
Multifractal
Support vector machine
Classification

ABSTRACT

The number of people affected by epilepsy is growing. Therefore, the design of accurate automated systems for detection and classification of electroencephalogram (EEG) signals of epileptic patients is a great aid in the diagnosis process. The purpose of this study is to present an accurate and fast automated diagnosis system to distinguish between normal and abnormal EEG records with seizure free intervals. The system is based on generalized Hurst exponent estimates at different scales used to characterize EEG records, and subsequently on a support vector machine classifier with different kernels to be employed for classification purpose. Statistical tests such as *t*-test, *F*-test, Kruskal-Wallis test and Kolmogorov-Smirnov test all show that multifractal based features are significantly different across normal EEG records and those with seizure free intervals. Finally, classification experiments following ten-fold and leave one out method cross-validation techniques yielded to 100% accuracy with low time processing cost.

1. Introduction

Epilepsy is a chronic disorder that happens in human brain and is associated with an increased vulnerability to seizures. The latter are in fact due to a disturbance in the electrical activity of the brain. When brain electrical signals are disrupted, memory can be affected and the patient may hurt himself. The analysis of the electroencephalogram (EEG) records used to measure the electrical activity in the brain is a common approach to help the physician in the diagnosis of seizures. In this regard, in recent years, various and interesting automated systems have been proposed in the literature to assist physicians and to improve the diagnosis outcome, while reducing duration of consultation [1–8].

For instance, the authors in [1] employed a genetic algorithm to automatically obtain Fourier transform based characteristics and used *k*-nearest neighbour (*k*-NN) to classify segments with seizure activity against non-seizure segments either belonging to an epileptic subject or to a healthy one. The proposed system achieved 98.53% accuracy. In a similar problem, the authors in [2] employed an approach based on bandwidths of intrinsic mode functions fed to least squares support vector machine having Morlet wavelet as kernel function. The maximum (minimum) classification accuracy among ten classification results for seizure against non-seizure EEG signals was 100% (95.50%) using second intrinsic mode function based bandwidths. In a subsequent study [3], fuzzy Sugeno classifier trained with four different entropy measures reached 98.1% accuracy when used to classify healthy, pre-ictal, and ictal EEG records. An automated system that combines smoothed Hilbert-Huang transform and root mean square feature was proposed in [4] to detect seizure in epileptic brain. The system achieved 90.72% and 8.23% in terms of sensitivity and false discovery rate respectively. In an interesting study [5], fractional linear prediction was employed to model error energy in ictal and seizure-free EEG signals, and support vector machine classifier with radial basis function kernel for classification. The system achieved 96% sensitivity, 95% specificity, and 95.33% accuracy. The authors in [6] used statistical features extracted from time-frequency image in Hilbert-Huang domain and support vector machine to detect epileptic seizures in EEG records. The proposed system yielded to an average accuracy of 99.125% (97.5–100%). In [7], the proposed system for automatic seizure detection in EEG signals used spectral entropies and energy all obtained from Hilbert marginal spectrum to train support vector machine classifier. In the case of classification of healthy against seizure segments, the system achieved 99.85% accuracy and 98.80% in the case of classification of segments with seizure activity against non-seizure signals either belonging to an epileptic subject or to a healthy one. More recently [8], Gabor filter was applied to EEG signals to obtain four filter responses from which local binary pattern based traits are extracted. Finally, the latter were fed to the nearest neighbour algorithm to decide whether the original EEG signal belongs to seizure or seizure-free

E-mail address: salim.lahmiri@mail.mcgill.ca

https://doi.org/10.1016/j.bspc.2017.10.004
1746-8094/

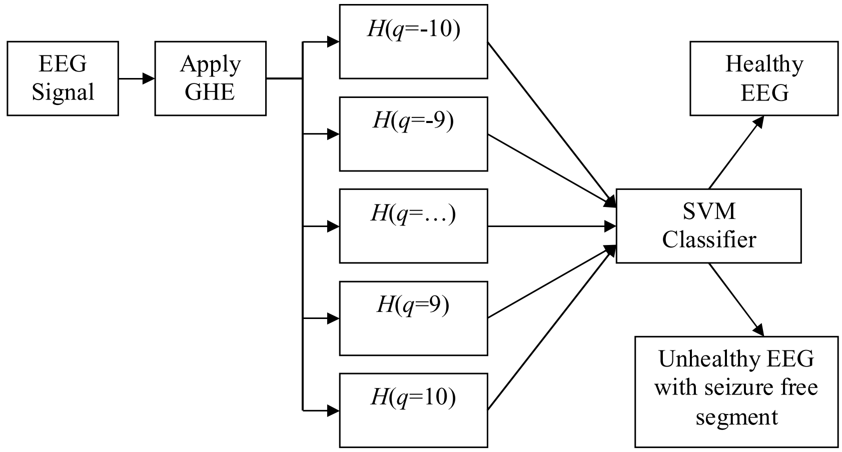

Fig. 1. Flowchart of the proposed system.

category. Following 10-fold cross validation protocol, the presented system achieved an average accuracy of 98.33%.

In this study, we propose a new automated system for classification of EEG signals. In particular, we aim to design an accurate, simple, and fast automated system to effectively distinguish between healthy and seizure free segments belonging to an epileptic patient. In fact, there are three reasons that motivate our study. First, it can be really difficult to diagnosis epilepsy since there is no evident sign a person has epilepsy, unless seizure occurs. Second, the investigation of this issue is definitely interesting in clinical applications. Third, prior works [1–8] have not focused on this problem.

In order to tackle this problem, we propose an automated system composed of two stages. In the first stage, multifractal of the EEG signal are computed at different time scales based on generalized Hurst exponent (GHE) [9]. Then, in the second stage, support vector machine (SVM) [10] classifier is trained with computed generalized Hurst exponent estimates to distinguish between healthy EEG signals and EEG signals with seizure free segments belonging to an epileptic patient. In this work, we rely on fractality assessment of the EEG signal at different time scales to better capture its dynamics; particularly, self-similarity. For instance, measuring the long-range dependence in EEG signal at different scales could help revealing its general power-law correlations in its short and long term variations separately. Therefore, the underlying stochastic behaviour of the EEG signal could be better characterized by the level of persistence in its short and long term variations. Obviously, such intrinsic characteristics are expected to be distinct across healthy and seizure free segments belonging to an epileptic patient. In addition, the GHE provides reliable statistics [11] and was found to be effective in characterization and classification of brain magnetic resonance images [12–16]. Finally, the SVM classifier is chosen thanks to its ability to execute the principle of structural risk minimization; in order to avoid local minima; and capability to generalize the results [10]. Further, it was found to be effective in classification of EEG signal in the context of epilepsy detection [2,5,6] and also in other biomedical engineering applications [17,18].

The rest of our work follows. Section 2 presents methods, Section 3 provides results, and Section 4 discusses our work. Finally, our study is concluded in Section 5.

2. Methods

The goal of employing the GHE [9] in this study is to measure self-similarity at different scales in healthy EEG signals and seizure free segments from EEG signals belonging to an epileptic patient. Then, the computed GHE estimates (for instance; $H(q)$ for $q = -10,\ldots,10$) are fed to the SVM [10] to distinguish between healthy and seizure free signals. Finally, the performance of the proposed system is evaluated based on accuracy, sensitivity, and specificity. In order to obtain robust and general results, the ten-fold cross validation protocol is adopted. Consequently, average and standard deviation of each performance measure is calculated. The flowchart of the proposed EEG classification system is shown in Fig. 1. The GHE, SVM, and performance metrics are all described next.

2.1. Generalized hurst exponent

The GHE [9] based approach for signal analysis is adopted in our study to compute Hurst exponents used to evaluate long memory in a given time series at different time scales. Let a signal $S(t)$ defined at discrete time intervals $t = \nu, 2\nu, \ldots T$ and T is an integer multiple of ν. The qth-order moments of the distribution that describe the statistical behaviour of $S(t)$ is expressed by [9]:

$$K_q(d) = \frac{\langle \| S(t+d) - S(t) \|^q \rangle}{\langle | S(t) |^q \rangle} \tag{1}$$

where $d \in [\nu, d_{max}]$ is a time interval and $d_{\ max}$ is its preset upper limit. The generalized Hurst exponent $H(q)$ is determined from the scaling behaviour of $K_q(d)$ according to the following empirical relation:

$$K_q(d) \propto= \left(\frac{d}{\nu}\right)^{qH(q)} \tag{2}$$

Therefore, when $K_q(d)$ and d satisfy a linear association for a given order q in log-log scale, the generalized Hurst exponent $H(q)$ can be calculated by running a linear regression of $\log(K_q(d))$ versus $\log(d)$. Consequently, $H(q)$ evaluates long-memory dependence or persistence in the original signal $S(t)$. In this regard, the multiscaling structure of signal $S(t)$ is linked to diverse orders q of the exponent $H(q)$. In this paper, the range of q is fixed to the interval from -10 to 10 (excluding -1 and 1) so as to cover a large width, and ν and $d_{\ max}$ are respectively set to 5 and 19. According to the interval to which q belongs, eighteen features –values of $H(q)$- are fed to the support vector machine classifier to distinguish between healthy and epileptic EEG signals free of seizure intervals.

2.2. Support vector machine and performance evaluation

With the purpose of splitting data points, the support vector machine classifier [10] executes a nonlinear mapping of the data in