

# Statistical analysis of epileptic activities based on histogram and wavelet-spectral entropy

Ahmad Mirzaei, Ahmad Ayatollahi, Hamed Vavadi

School of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran.  
Email: ahmad.mirzaei.ce@gmail.com; ayatollahi@iust.ac.ir; vavadi@elec.iust.ac.ir

Received 15 October 2010; revised 4 December 2010; accepted 7 December 2010.

## ABSTRACT

**Epilepsy is a chronic neurological disorder which is identified by successive unexpected seizures. Electroencephalogram (EEG) is the electrical signal of brain which contains valuable information about its normal or epileptic activity. In this work EEG and its frequency sub-bands have been analysed to detect epileptic seizures. A discrete wavelet transform (DWT) has been applied to decompose the EEG into its sub-bands. Applying histogram and Spectral entropy approaches to the EEG sub-bands, normal and abnormal states of brain can be distinguished with more than 99% probability.**

**Keywords:** Electroencephalogram (EEG); EEG Sub-Bands; Epileptic Seizures; Discrete Wavelet Transform (DWT); Histogram; Spectral Entropy (SEN)

## 1. INTRODUCTION

Physiological rhythms are important indicators of health. Diseases which cause some irregularities in these rhythms may lead to death. Epilepsy is a common disorder of this type. About 1% of the Americans are affected by epilepsy [1].

Epilepsy is characterized by abnormal irregular firing of neurons due to synchronous or excessive neuronal activity in the brain. Due to high complexity of brain we should apply several linear and non-linear signal processing methods to analyse Electroencephalogram (EEG) signal truly. EEG signals extracted from the brain show a non-stationary behaviour.

Duke and Pritchard studied about the chaos of brain. They proved that, because of non-stationary behaviour and complex dynamic of the brain, chaotic methods are appropriate for EEG signal analysis [2]. Since the neuronal activities in ictal, interictal, and healthy states significantly differ from each other, chaotic methods such as entropies can be applied to distinguish between these three states [3].

It is helpful to use discrete wavelet transform (DWT) because of its advantages such as time-frequency localization, multi-rate filtering, and scale-space analysis [3]. EEG sub-bands have more accurate information about neuronal activity compared to the original full spectrum EEG. Also, DWT is a powerful transform to analyse non-stationary signals, because it has a good localization in both time and frequency domains [4,5]. Fast events and changes in neuronal activities like spikes that are not obvious in full spectrum EEG, can be recognized in sub-bands. Thus to detect epileptic seizures accurately, each sub-band should be analysed separately [3].

Recently, chaotic methods like correlation dimension, largest Lyapunov exponent and entropies have been applied to EEG sub-bands acquired from preprocessing analysis based on DWT [3]. Iasemidis and Sackellares were the first researchers who studied the nonlinear behaviour of EEG and epilepsy [6]. They applied largest Lyapunov exponent as a chaotic parameter to show the reduction of chaos in preictal phase [7,8]. Elger and Lehnertz similarly concluded that complexity of neuronal activities is reduced in preictal phase. They used correlation dimension and moving-window dimension analysis for patients with temporal lobe epilepsy [9,10]. Tayaranian *et al.* using three features; Detrended Fluctuation Analysis (DFA), Bis-spectral Analysis (BIS), and Standard Deviation (SD), for EEG analysis [11]. They also apply a fuzzy classifier to separate the three groups.

Kumar *et al.* applied entropy method based on Shannon and spectral entropy and concluded that the Shannon entropy value in ictal seizure activity is low compared to healthy and interictal states. Also they reported that the spectral entropy (SEN) value in normal and interictal states is low compared to ictal state [12]. We previously developed a method for distinguishing ictal state from healthy one by calculating the SEN values for wavelet coefficients of EEG and its sub-bands [13].

In this study the EEG signal has been decomposed into five sub-bands by discrete wavelet transform (DWT). Then histogram and Wavelet-spectral entropy have been

presented to analyse EEG sub-bands for epileptic seizures detection. To validate the results statistical analysis has been applied at the end of each process. The block diagram of the procedure is depicted in **Figure 1**.

We applied our method to three different groups [14]:

- 1) Healthy subjects that contain state one and two, corresponding eyes open and closed, respectively.
- 2) Interictal subjects (seizure-free intervals) that contain two different sets: focal interictal activity and non-focal interictal activity.
- 3) Ictal subjects (epileptic activity).

## 2. MATERIALS AND METHODS

### 2.1. Description of the EEG Database

Our EEG data, acquired from university of Bonn, contains three different cases: 1) healthy, 2) epileptic subjects during seizure-free interval (interictal), 3) epileptic subjects during seizure interval (ictal) [14].

Each case has five datasets named: O, Z, F, N, and S. Sets O and Z are obtained from healthy subjects under condition of eyes open and closed; respectively by external surface electrodes. Sets F and N are attained from interictal subjects.

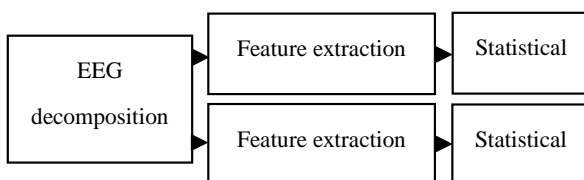
Set F taken from epileptogenic zone of the brain shows focal interictal activity; set N obtained from hippocampal formation of the opposite hemisphere of the brain indicates non-focal interictal activity, and set S is got from an ictal subject.

Each set contains 100 single channel EEG segments of 23.6 sec duration. Sampling frequency is 173.61 Hz, so each segment contains  $N = 4096$  samples [14]. All these EEG segments are recorded with the same 128-channel amplifier that converts by 12 A/D convertor with bit rate of 12, and then were sampled on 173.61 Hz [14].

Since the sampling frequency of EEG records is 173.61 Hz, according to Nyquist sampling theorem, maximum frequency of EEG applied should be in the range 0-86.81 Hz. based on the physiological researches, frequencies above 60 Hz in EEG signal are considered as noise and can be neglected [3]. Thus we have eliminated these frequencies using a low-pass FIR filter.

### 2.2. Wavelet Decomposition

DWT is a proper transform to analyse signals in both time and frequency domains. Using the multi-scale



**Figure 1.** Block diagram of proposed method.

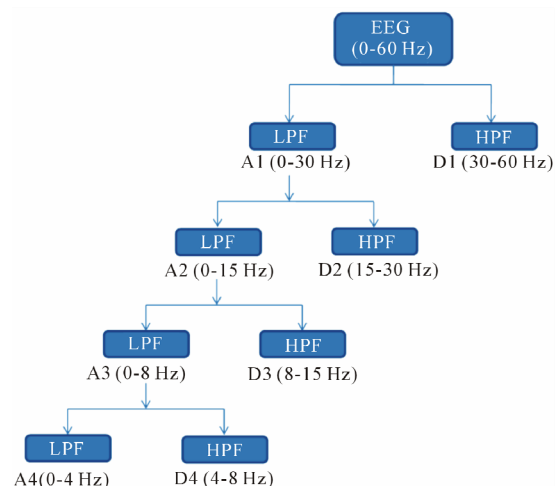
form of DWT, signal can be decomposed into separate frequency bands which each band represents a specific roughness. Wavelet transform uses a variable window size over the length of the signal, which allows the wavelet to be stretched or compressed depending on the signal frequency [3].

In this study fifth-order Daubechies (DB5) DWT has been applied to the band-limited EEG (0-60 Hz). After the first level of decomposition, the band-limited EEG has been decomposed into its high resolution frequency band, D1 (30-60 Hz), and low resolution frequency band, A1 (0-30 Hz), which should be decomposed in next level. In the second level of decomposition, A1 has been decomposed into its high, D2 (15-30), and low, A2 (0-15 Hz) resolution bands. This process has been repeated four times. After full decomposition five sub-bands have been attained: high frequency sub-bands (details) of levels 1 to 4 (D1 (30-60 Hz), D2 (15-30 Hz), D3 (8-15 Hz), D4 (4-8 Hz)) as well as the low frequency sub-band (approximate) of the last level (A4 (0-4 Hz)). **Figure 2** illustrates this multi-level decomposition process schematically.

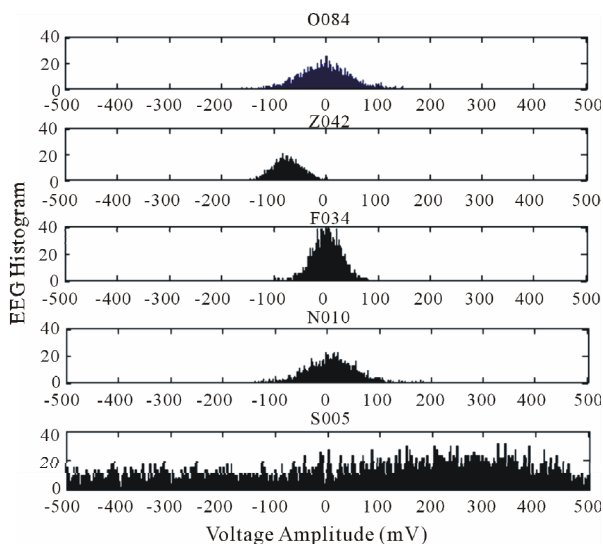
These five frequency sub-bands are almost corresponding to five physiological EEG bands, delta (0-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30), and gamma (30-60 Hz).

### 2.3. Statistical Analysis of Histograms

Histogram is a visual feature for indicating the dispersion of EEG data. Each rectangle in the histograms shows the population of samples falls into several bins corresponding to different amplitudes in EEG sequence. Plots of **Figure 3** show outstanding differences in dispersion of data in histograms for epileptic patients in conniption time from healthy and epileptic patients in seizure-free interval. This led us to consider the histogram



**Figure 2.** Schematic of multi-level decomposition.



**Figure 3.** EEG histograms for five typical subjects.

as a criterion to assess epileptic seizures. To quantify the discrepancies of the data distributions, mean and standard deviation (STD) have been calculated for each histogram. This has been repeated for all EEG sub-bands and the results are reported in **Table 1**.

### 2.4. Statistical Analysis of Entropies

In this section we have applied the spectral entropy method to differentiate between the mentioned cases. After development of information theory scientists introduced the concept of entropy [15]. Entropy expressed first by Shannon in 1940s [15]. To calculate the spectral entropy of a signal it is necessary to have the power spectrum values. The square of the Fourier transform  $|F(\omega)|^2$  is called power spectrum, which indicates the distribution of signal energy in frequency domain [16]. It should be noted that, the power spectrum is defined for

**Table 1.** Mean and standard deviations (in parenthesis) of histograms.

	O n = 100	Z n = 100	F n = 100	N n = 100	S n = 100
EEG (0-60 Hz)	-12.476 (61.012)	-6.223 (40.641)	-6.175 (65.503)	-8.8518 (50.746)	-4.8647 (305.91)
A4 (delta) (0-4 Hz)	-12.481 (29.63)	-6.2259 (28.904)	-6.1735 (56.911)	-8.8561 (44.738)	-4.8409 (184.61)
D4 (theta) (4-8 Hz)	0.0044 (34.555)	0.0042 (18.188)	-0.0001 (26.058)	0.0045 (19.136)	-0.0309 (170.16)
D3 (alpha) (8-15 Hz)	0.0014 (36.658)	-0.0008 (18.813)	-0.0011 (15.333)	-0.0001 (11.667)	0.0016 (144.53)
D2 (beta) (15-30 Hz)	-0.0007 (13.055)	-0.0005 (9.909)	-0.0003 (6.4223)	-0.0002 (5.4961)	0.0056 (57.906)
D1(gamma) (30-60 Hz)	0.0001 (3.3764)	-0.0001 (2.3939)	0 (1.643)	0 (1.5903)	0 (9.4568)

stochastic and stationary processes, while the EEG signal is a stochastic and non-stationary sequence. Hence, to assess the stationarity condition entire series have been divided into 32 sub-segment (epoch). Each of these sub-segments contains 128 samples with 0.74 second duration. The value of spectral entropy for each epoch has been calculated in the following steps:

*Step 1:* the power spectrum,  $p(f)$  of the signal,  $f(t)$  has been calculated as:

$$p(f) = \left( \int f(t) \exp(j\omega t) dt \right)^2 \quad (1)$$

where, equation  $\int f(t) \exp(-j\omega t) dt$  is continuous-time Fourier transform.

Then, it has been normalized to its summation:

$$Q(f) = \frac{p(f)}{\sum_f p(f)} \quad (0 \leq Q(f) \leq 1) \quad (2)$$

*Step 2:* The Shannon function,  $f(x) = x \log(1/x)$  has been applied to the normalized power spectrum components  $Q(f)$ :

$$H(f) = Q(f) \log(1/Q(f)) \quad (3)$$

*Step 3:* Normalized entropy, SEN, has been calculated as:

$$SEN = \frac{\sum_f H(f)}{\log(N(f))} \quad (0 \leq SEN \leq 1) \quad (4)$$

where,  $N(f)$  is the total number of frequency components. Normalized entropy is a measure of regularity. SEN value equal to one shows the maximum irregularity, and equal to zero shows the complete regularity. For each 128 samples of power spectrum an entropy value has been calculated. We have assigned the average value of 32 entropies to each data-segment. This has been repeated for all 100 data-segments. Then, mean and STD of these SEN values have been estimated for each band limited EEG and its sub-bands. The results are collected in **Table 2**.

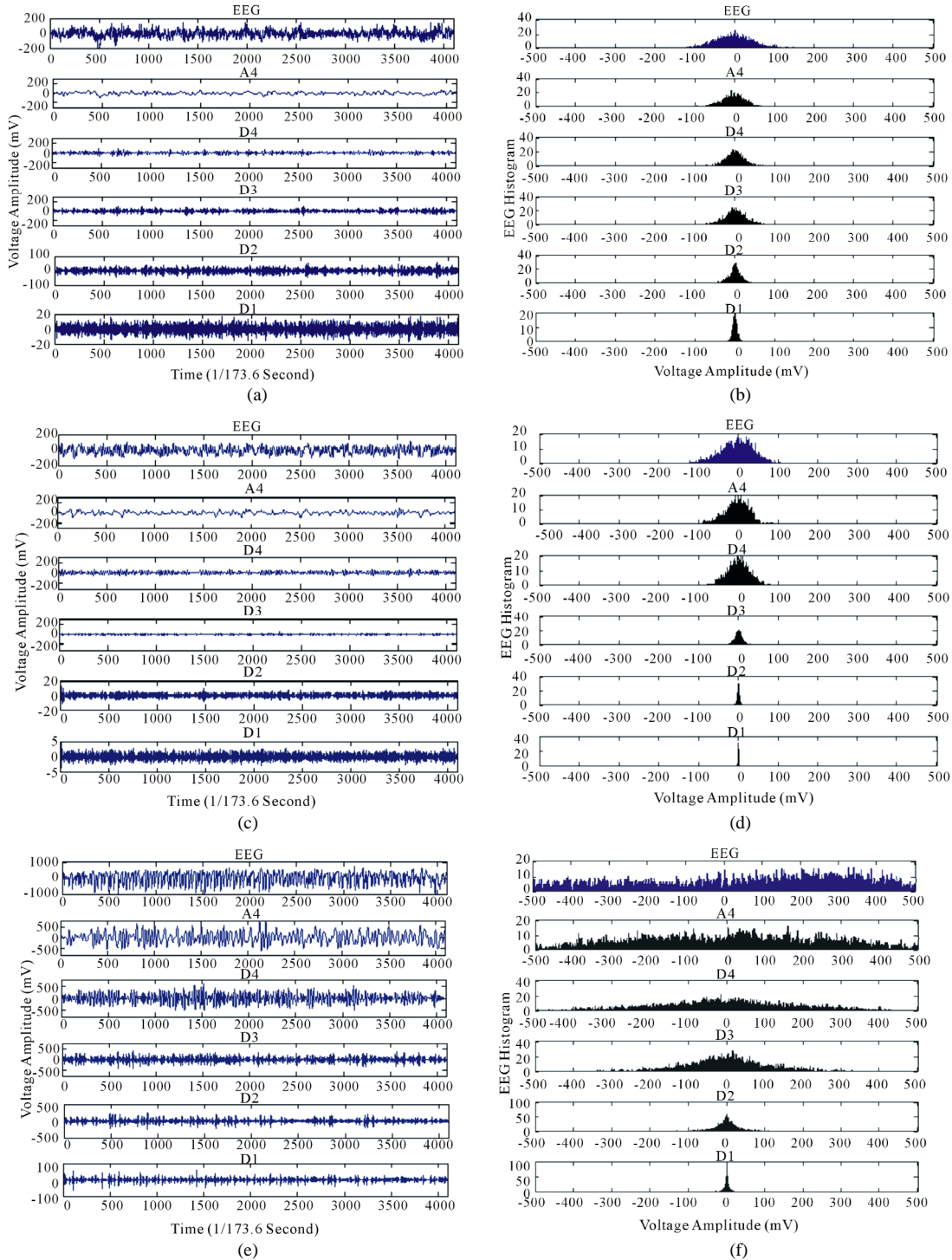
**Table 2.** Mean and standard deviations (in parenthesis) of SEN values.

	O n = 100	Z n = 100	F n = 100	N n = 100	S n = 100
EEG (0-60 Hz)	0.629 (0.073)	0.634 (0.085)	0.538 (0.093)	0.539 (0.091)	0.653 (0.069)
A4 (delta) (0-4 Hz)	0.375 (0.09)	0.393 (0.097)	0.419 (0.096)	0.425 (0.097)	0.484 (0.095)
D4 (theta) (4-8 Hz)	0.563 (0.074)	0.58 (0.063)	0.567 (0.071)	0.565 (0.067)	0.553 (0.074)
D3 (alpha) (8-15 Hz)	0.641 (0.086)	0.69 (0.061)	0.684 (0.062)	0.685 (0.057)	0.667 (0.064)
D2 (beta) (15-30 Hz)	0.762 (0.055)	0.761 (0.055)	0.79 (0.053)	0.794 (0.05)	0.759 (0.054)
D1(gamma) (30-60 Hz)	0.786 (0.051)	0.81 (0.05)	0.821 (0.054)	0.822 (0.045)	0.871 (0.041)

### 3. Results

We have pointed to widespread shape of pure epileptic histogram versus histograms of seizure-free and normal intervals in section 2.3. **Figure 4** depict this notion in

detail. **Figures 4(a), (c), (e)** show the EEG data recorded from a typical healthy, interictal, and ictal subject, respectively, and **Figures 4(b), (d), (f)** display the corresponding histograms. Comparison between these histograms shows that, in ictal state EEG samples spread



**Figure 4.** Band limited EEG data recorded from a typical healthy (a), interictal (c), and ictal (e) subject, respectively, these corresponding histograms in (b), (d), and (f) respectively.

more than EEG samples obtained from healthy and interictal cases. This uniformity in histogram of seizure interval demonstrates the existence of high amplitude samples as well as low amplitude ones. **Table 1** shows the means and standard deviations (STD) for all cases. As we expected from histograms, the STD of the band-limited EEG and its sub-bands in ictal state are high compared to healthy and interictal states. This shows the low frequency and high amplitude activity of neurons in conniption time, while, in other cases the brain has only a low amplitude activity. As we mentioned in section 2.4, SEN is an appropriate parameter to separate normal and epileptic EEGs.

It is clear from **Table 2** that for the band-limited EEG, in ictal state the SEN value is high compared to normal and interictal states. This result is similar to the result of Kumar *et al.*'s work [12] and contravenes the physiological aspects. Because according to physiological aspects the regularity should be increased on conniption times. As mentioned above when regularity is increased the entropy values tending to zero. Thus we conclude that the SEN values in ictal state should be less compared to interictal and healthy states. This result is not satisfied in Kumar *et al.*'s work and in our study in band-limited EEG. Thus we applied DWT to decompose the band-limited EEG and evaluate its sub-bands but Kumar *et al.* evaluate just band-limited EEG. This result is repeated for delta (low frequencies) and gamma (high frequencies) sub-bands. But, for middle sub-bands, theta, alpha, and beta, SEN values of ictal state are low compared to healthy and interictal states. These SEN values of middle sub-bands satisfy the physiological aspect. It can be seen that three middle sub-bands have the same results as the results obtained from histogram analysis. This implies that the regularity of EEG signal in its middle sub-bands increases in conniption time.

Comparison of three middle sub-bands shows that unlike the beta sub-band in alpha and theta sub-bands the disparity of the SEN values of the healthy and interictal subjects is negligible. Hence if the beta sub-band is considered to be a representation of brain dynamics by itself, the regularity becomes low in epileptic patients during seizure-free interval.

**Figure 5** shows the values of normalized SEN for the typical band-limited EEGs and their beta sub-bands.

By means of T-Student distribution test<sup>1</sup> we have emphasized that in beta sub-band, healthy subjects can be distinguished from epileptic patients during seizure-free interval with more than 99% probability [17]. The probabilities of distinguishing between interictal state and two other states in beta sub-band are shown in **Table 3**.

<sup>1</sup>For more information refer to Appendix.

In fact in this study we focus on comparison between interictal state and healthy and ictal states, because it is clear that for prevention treatments we should evaluate the interictal activities to help prevent the seizure. **Table 4** shows the comparison of our result and Kumar *et al.* result based on Spectral entropy. Their results are for band limited EEG that do not satisfy physiological aspect totally but our results are for beta sub-band that satisfy physiological aspect and achieve better results.

As we see from **Table 4** that we compare the focal and non-focal interictal with ictal, separately.

#### 4. Conclusion

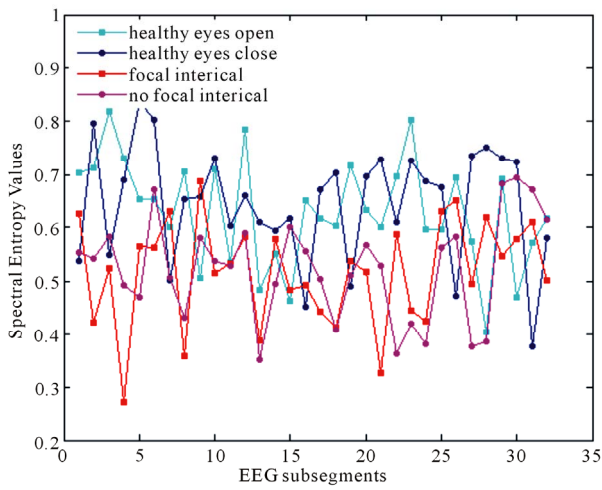
Two methodologies based on statistics and chaos theory in time-frequency domain have been presented for analysis of EEGs and its delta, theta, alpha, beta, and gamma sub-bands for detection of epilepsy. Since the EEG is an overall representation of brain dynamics, the observed changes in the band-limited EEG are actually the result of the total activity of neurons which performs a substantial role in forming the shape of the EEG signal. One method of studying these underlying activities are to study the component physiological sub-bands of the EEG which can be assumed to represent the neurological activities at a finer level. It has been shown that the statistical analysis of histograms of EEGs and their sub-bands can reveal the differences between frequencies dependent amplitudes of EEG signal in ictal state from healthy and interictal states. It has been observed that the spectral entropy of the band-limited EEG can distinguish between the three groups of subjects. The decomposition of the original EEG into its five constituent sub-bands alters the results obtained from original signal. However, when the statistical analysis is performed on the EEG sub-bands, it can be seen that the SEN used within certain physiological sub-bands may also be employed to distinguish between all three groups. Therefore, it has been emphasized that changes in the dynamics are not spread out equally across the spectrum of the EEG, but instead, are limited to certain frequency bands.

**Table 3.** Probability of differentiation between a typical interictal case and two other cases in healthy and ictal state.

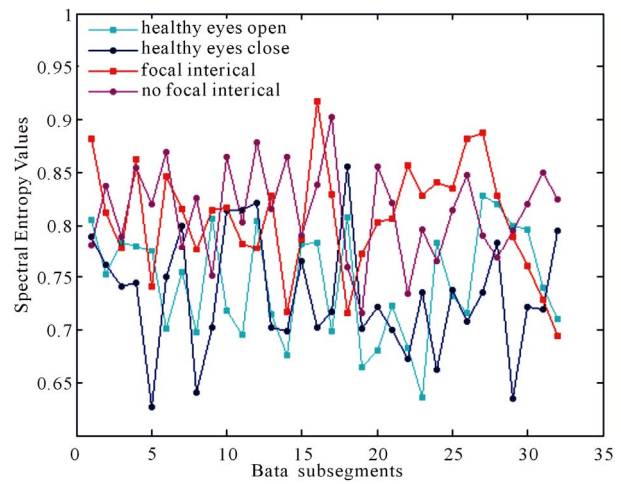
	F-O	F-Z	F-S	N-O	N-Z	N-S
beta	99.9%	99.9%	99.93%	99.94%	99.94%	99.95%

**Table 4.** Comparison of our work and Kumar *et al.*

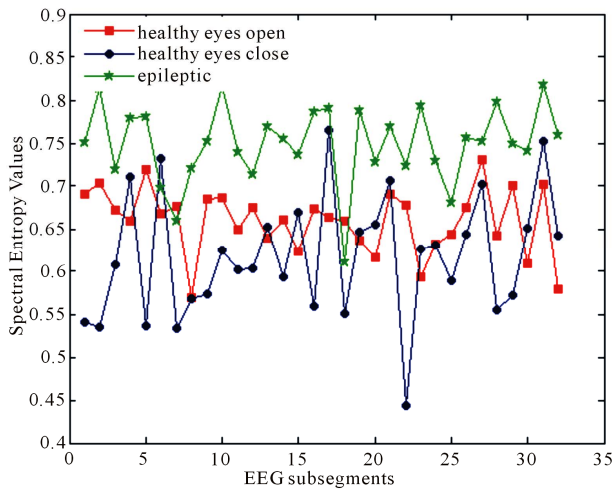
Kumar <i>et al.</i>	Comparison between F, N, and S states	
	94.6%	
Our work	Comparison between F and S	Comparison between N and S
	99.93%	99.95%



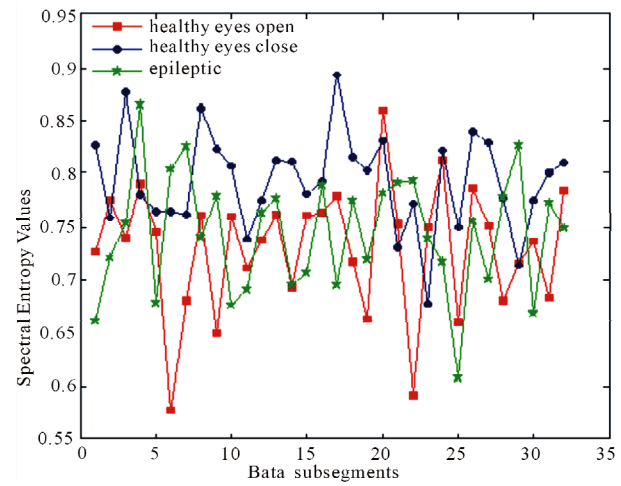
(a)



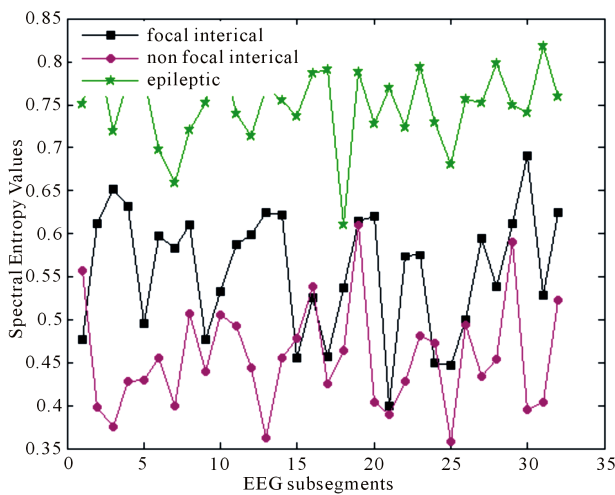
(b)



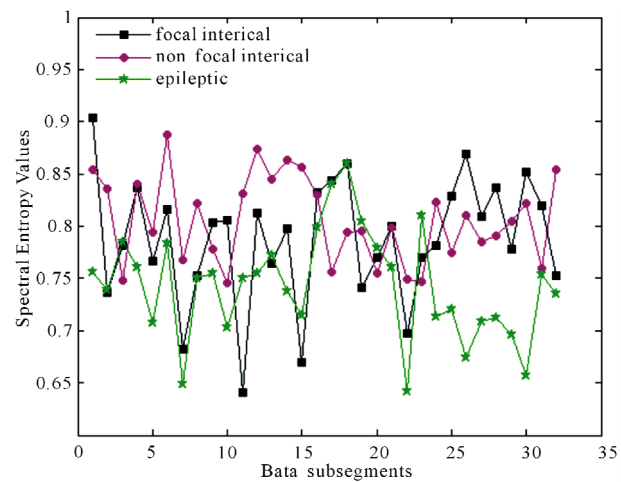
(c)



(d)



(e)



(f)

**Figure 5.** Comparison of normalized SEN values of typical EEGs and their beta sub-bands between healthy and interictal (a), (b), healthy and ictal (c), (d), and ictal and interictal (e), (f) states.

## REFERENCES

- [1] Bethesda, M.D. (2004) Seizures and epilepsy: Hope through research National Institute of Neurological Disorders and Stroke (NINDS). [http://www.ninds.nih.gov/health\\_and\\_medical/pubs/seizures\\_and\\_epilepsy\\_htr.htm](http://www.ninds.nih.gov/health_and_medical/pubs/seizures_and_epilepsy_htr.htm)
- [2] Duke, D., and Pritchard, W. (1991) Measuring chaos in the human brain. World Scientific, Singapore.
- [3] Adeli, H., Ghosh-Dastidar, S., and Dadmehr, N. (2007) A wavelet-chaos methodology for analysis of EEGs and EEG sub-bands to detect seizure and epilepsy. *IEEE Transaction on Biomedical Engineering*, **54**, 205-211. [doi:10.1109/TBME.2006.886855](https://doi.org/10.1109/TBME.2006.886855)
- [4] Gao, W. and Li, B.L. (1993) Wavelet analysis of coherent structures at the atmosphere-forest interface. *Journal of Applied Meteorology*, **32**, 1717-1725. [doi:10.1175/1520-0450\(1993\)032<1717:WAOCSA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1993)032<1717:WAOCSA>2.0.CO;2)
- [5] Meyer, Y. (1993) Wavelets: Algorithms and applications. SIAM, Philadelphia, PA.
- [6] Iasemidis L.D. and Sackellares J.C. (1991) The temporal evolution of the largest Lyapunov exponent on the human epileptic cortex. World Scientific, 49-82.
- [7] Iasemidis, L.D., Shiau, D.S., Sackellares, J.C., and Pardalos, P.M. (2000) Transition to epileptic seizures: Optimization. In: *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*. American Mathematical Society, Providence, RI, **55**, 55-74.
- [8] Iasemidis, L.D., Shiau, D.S., Chaovalitwongse, W., Sackellares, J.C., Pardalos, P.M., Principe, J.C., Carney, P.R., Prasad, A., Veeramini, B., and Tsakalis, K. (2003) Adaptive epileptic seizure prediction system. *IEEE Transaction on Biomedical Engineering*, **50**, 616-627. [doi:10.1109/TBME.2003.810689](https://doi.org/10.1109/TBME.2003.810689)
- [9] Elger, C.E. and Lehnertz, K. (1994) Ictogenesis and chaos. In: Wolf, P. Ed., *Epileptic Seizures and Syndromes*. Libbey, London, 547-552.
- [10] Elger, C.E. and Lehnertz, K. (1998) Seizure prediction by non-linear time series analysis of brain electrical activity. *European Journal of Neuroscience*, **10**, 786-789. [doi:10.1046/j.1460-9568.1998.00090.x](https://doi.org/10.1046/j.1460-9568.1998.00090.x)
- [11] Tayaranian, P., Shalbaf, R., and Nasrabadi, A.M. (2010) Extracting a seizure intensity index from one-channel EEG signal using bispectral and detrended fluctuation analysis. *Journal of Biomedical Science and Engineering*, **3**, 253-261. [doi:10.4236/jbise.2010.33034](https://doi.org/10.4236/jbise.2010.33034)
- [12] Kumar, S.P., Sriraam, N. and Benakop, P.G. (2008) Automated detection of epileptic seizures using wavelet entropy feature with recurrent neural network classifier. *TENCON IEEE*, 1-5. [doi:10.1109/CICSyN.2010.84](https://doi.org/10.1109/CICSyN.2010.84)
- [13] Mirzaei, A., Ayatollahi, A., Gifani, P., and Salehi, L. (2010) Spectral entropy for epileptic seizures detection. *Second International Conference on Computational Intelligence, Communication Systems and Networks*, Liverpool.
- [14] Andrzejak, R.G., Lehnertz, K., Rieke, C., Mormann, F., David, P., and Elger, C.E. (2001) Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. <http://www.meb.unibonn.de/epileptologie/science/physik/eegdata.html>.
- [15] Shannon, C.E. (1948) A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379-423.
- [16] Qian, S., and Chen, D. (1996) Joint time-frequency analysis: methods and applications. Prentice Hall, Upper Saddle River.
- [17] Mendenhall, W., Beaver, R.J., and Beaver B.M. (2006) Introduction to probability and statistics. 12th Edition, Cengage Learning, Stamford.

## APPENDIX

### 1. Two- Sample t Test

T-Student is applied to two independent samples to estimate the probability (p-value) of distinguishing between two populations means.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (\text{A.1})$$

$t$  is the T-student test and  $s_1^2$  and  $s_2^2$  are samples variances.  $\bar{X}_1$  and  $\bar{X}_2$  are means of samples.  $n_1$  and  $n_2$  are sample intervals that are equal  $n_1 = n_2 = 100$ . To

obtain the p-value, T-Student degrees of freedom,  $df$  should be calculated as:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} \quad (\text{A.2})$$

We can read p-value from the T-Student table (A.1) that is arranged according to  $t$  and  $df$  parameters calculated by above equations.