Ahmad MIRZAEI[1], Ahmad AYATOLLAHI[2], Ali Motie NASRABADI[3]

Iran University of Science and Technology (1 and 2), University of Shahed (3)

Automated detection of epileptic seizures using mixed-methodology: Wavelet-Chaos-KNN Classifier-Mutual Information

***Abstract**. Electroencephalogram (EEG) is the brain signal that contains the valuable information about different states of the brain. In this study EEG signals are analyzed for evaluating epileptic seizures in these signals and their sub-bands and comparing epileptic states with other states. A discrete wavelet transform is applied for decompose the EEGs into its sub-bands. The chaotic behavior of EEGs is evaluated by means of normalized Shannon and spectral entropies. Entropy method is presented for detection of epileptic seizures through the analysis of EEGs and their sub-bands. At the end the mixture K-nearest neighbor and mutual information method is applied as a classifier to classify the different states in EEGs and their sub-bands. This method is applied to three different groups of EEG signals: 1) healthy states, 2) epileptic states during a seizure-free interval (interictal EEG), 3) epileptic states during a seizure (ictal EEG). The proposed method could classify different states with 99% accuracy.*

***Streszczenie.** Elektroencefalografia EEG jest analizą sygnału mózgu. W artykule przedstawiono metody analizy sygnału EEG stosowane w celu wykrycia epilepsji. Zastosowano dyskretną transformatę falkową do dekompozycji sygnału EEG. Wykorzystano metodę entropii do detekcji sygnału związanego z epilepsją. Metody zastosowano do trzech grup pacjentów: zdrowych, chorych na epilepsję i chorych w czasie ataku epilepsji. (**Automatyczna detekcja ataku epilepsji przy wykorzystaniu połączenia różnych metod: falkowej – klasyfikacji KNN i wzajemnej informacji**).*

Keywords: discrete wavelet transform (DWT); normalized Shannon entropy (NShEn) normalized spectral entropy (NSEn); k-nearest neighbor (KNN); mutual information, encephalography, epileptic seizures.
Słowa kluczowe: dyskretna transformata falkowa, znormalizowane spektrum entropii, analiza sygnału epilepsji, encefalografia.

Introduction

Epilepsy is known as the most prevalent neurological disorder that affects nearly 1% of people worldwide [1]. Epileptic seizures occur when the set of neurons firing synchronic. Traditional seizure detection methods, like visual scanning of EEG, are very time consuming, expensive and inaccurate methods. Thus automatic seizures detection methods are necessary. Recently many automatic methods were applied with researchers. Most of these methods are based on chaotic methodology, time-frequency analysis and neural networks or mixed methods.

Iasemidis was among the first to study the chaotic behavior of EEG signal and conclude that the chaos in EEG was decreased in ictal state [2], [3], [4]. Adeli *et al.* applied correlation dimension (CD) and Largest lyapunov exponent (LLE) as chaotic parameters and concluded that complexity was reduced in conniption time [5], [6]. They also applied mixed band methodology, using chaos-neural network, and classify different states from each other [7]. Kumar *et al.* was applied Shannon and Spectral entropies to EEG signals and were concluded that entropy was reduced in ictal state [8]. Ocak was applied approximate entropy (APEn) as a chaotic measure and distinguished the normal and epileptic states from each other [9].

We apply the new method for evaluation the EEG behavior in different states and compare and classify these states from each other. In this study we apply the Discrete Wavelet Transform (DWT) for preprocessing and Normalized Shannon and spectral entropy (NShEn and NSEn) as chaotic measures for chaotic analysis. Also we classify our signals by means of K-Nearest Neighbor (KNN) classifier. The results were improved when we applied mutual information. This new method is explained in details in following sections. Figure 1 shows the block diagram of the proposed process. The proposed process is applied to three different groups of EEG signals: 1) healthy states; 2) epileptic states during a seizure free interval (interictal EEG); 3) epileptic states during a seizure (ictal EEG).

Materials and Methods

A. Description of the EEG data used in this study

The data used in this study is acquired from university of Bonn [10]. This datasets contain three different states: 1) healthy states, 2) interictal states, and 3) ictal state. These three states involve five different datasets; O, Z, F, N, and S. Sets O and Z expressing two different normal states with eyes open and close respectively. Sets F and N expressing two different interictal states. Set F is acquired from epileptogenic area of the brain that shows focal interictal activity; set N is acquired from hippocampal formation of the opposite hemisphere of the brain that shows non-focal interictal activity. Set F is obtained from ictal state. Each dataset comprises 100 single channel EEG segments with 23.6 sec time duration and 173.61 Hz sampling frequency, so each segment contains 4097 samples [10]. Figure 2 illustrates the five typical EEG data sets in 5 sec time duration that contain 868 samples.

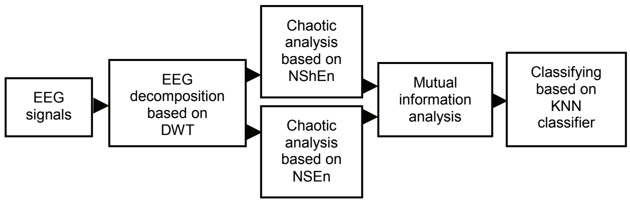

Fig.1. Block diagram of proposed process

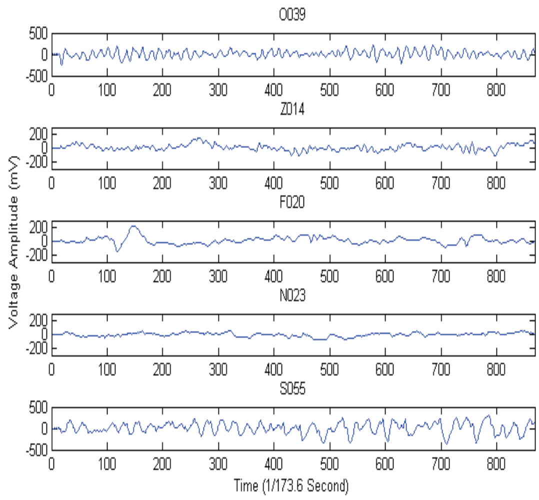

Fig.2. Typically EEG (0-5 sec) for (from top to bottom) healthy (O39), healthy (Z14), interictal (F20), interictal (N23), and ictal (S55)

B. EEG preprocessing based on DWT

DWT is a useful multilevel decomposition tool. DWT is useful for decomposing the non-stationary signals like EEG that traditional methods like Fourier transform are incapable. By means of DWT the EEG signals decomposed into low frequencies (approximates (A)) and high frequencies (details (D)) in each level that obtained using low and high pass filters respectively. In this study a fourth-order Daubechies (DB4) DWT has been used to decompose the band-limited EEG (0-60 Hz) into its sub-bands. The sampling frequency is 173.61 and according to Nyquist sampling theorem; maximum useful frequency is half of this measure or 86.8. According to physiological aspects frequencies greater than 60 Hz can be recognized as noise and could be filtered [5]. Thus in this study the band limited EEG (0-60 Hz) is decomposed by means of DWT. In each level we keep the high frequency band and low frequency band is decomposed in next level. This process is applied for four level and five frequency bands are yielded. Figure 3 shows decomposition process.

At the end the standard physiological sub-bands; delta, theta, alpha, beta, and gamma are yielded.

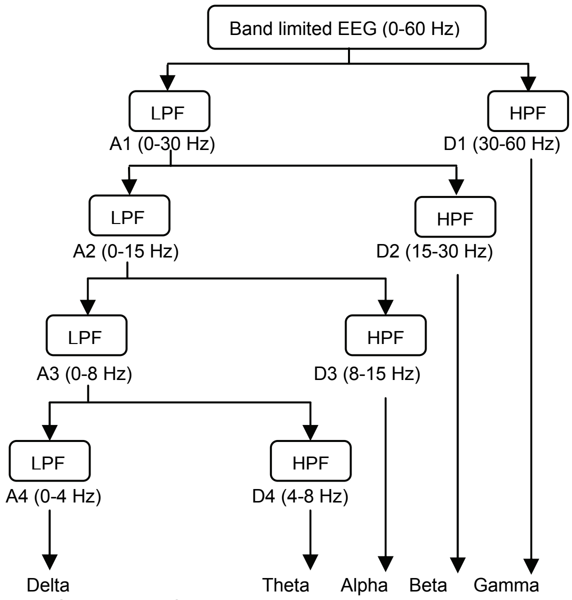

Fig.3. Schematic of multi-level decomposition

Figure 4 shows a typically full time duration interictal signal and its sub-bands that yielded from DWT multi-level decomposition.

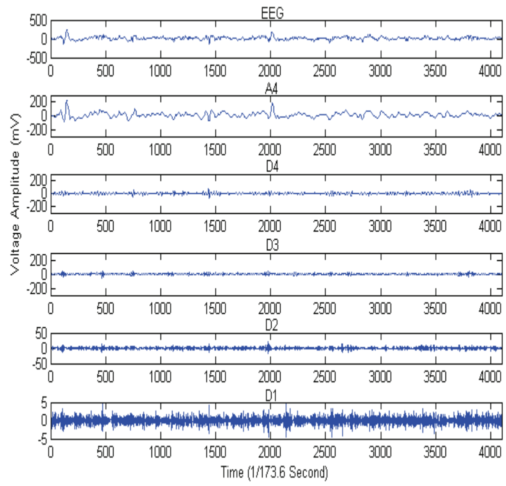

Fig.4. A typically interictal EEG and its sub-bands.

C) Chaotic analysis of EEG based on entropy criterion

1) Normalized Shannon entropy

In this study two entropy measures are applied; normalized Shannon entropy and normalized spectral entropy. The basic concept of entropy is developed first by Shannon [11] that expresses by following equation that called Shannon entropy:

$$ShEn = -\sum_i P_i \log P_i \tag{1}$$

where p_i is the probability measure of occurrence of amplitude a_i anywhere in the signal. When the ShEn is divided by $\log k$ the normalized Shannon entropy is yielded that k is the number of bins that obtained from histogram analysis [12]:

$$NShEn = \frac{ShEn}{\log k} \tag{2}$$

2) Normalized spectral entropy

If the P_i in (1) is replaced by the power spectrum P_n of the signal the spectral entropy is obtained:

$$SEn = -\sum_i P_{ni} \log P_{ni} \tag{3}$$

if the SEn that obtained from above equation divide by $\log N_f$ the normalized spectral entropy, NSEn, is yielded:

$$NSEn = \frac{SEn}{\log N_f} \tag{4}$$

where N_f is the frequency components. The NShEn and NSEn values are between 0 to 1; values near to zero show maximum regularity and near to one show maximum irregularity.

3) Entropy analysis

As mentioned before five data sets are analyzed in this study. To protect the stationary condition, each whole segment is divided to 32 sub-segments with 0.74 sec time duration samples. The NShEn and NSEn values are calculated for each sub-segment. At the end the average and standard deviation of these NShEn and NSEn values are calculated. Table 1 and 2 show the NShEn and NSEn values respectively.

Classifying based on KNN Classifier

In this study, the k-Nearest Neighbor (KNN) classifier is used for detection of epileptic subjects. KNN is a supervised learning algorithm. The classification of this algorithm is based on using a reference data set containing both the input and the target members and then it compares the unknown input members to that reference data set. Assignment of the class is done by determining the distance of the unknown to the k nearest neighbors by obtaining a majority vote from them. In the first step, 32 Shannon (or spectral) entropy values for each segment were extracted as features. For improving classification performance, the features were reduced to 20 features by applying minimal-redundancy-maximal relevance criterion (MRMR) [13]. This method is expressed in the next section.

Table 1. Average and standard deviation (in parenthesis) of NShEn

	O N=100	Z N=100	F N=100	N N=100	S N=100
EEG	0.8673 (0.025)	0.8676 (0.028)	0.8618 (0.043)	0.8678 (0.037)	0.8495 (0.032)
Delta	0.896 (0.037)	0.8903 (0.04)	0.878 (0.046)	0.878 (0.044)	0.8801 (0.035)
Theta	0.8504 (0.04)	0.8515 (0.039)	0.8451 (0.046)	0.8532 (0.04)	0.8495 (0.04)
Alpha	0.8468 (0.034)	0.8413 (0.034)	0.8172 (0.05)	0.8386 (0.037)	0.8052 (0.044)
Beta	0.8295 (0.031)	0.8301 (0.030)	0.7983 (0.052)	0.8256 (0.034)	0.7494 (0.051)
Gamma	0.8378 (0.031)	0.8369 (0.025)	0.8069 (0.047)	0.8295 (0.03)	0.7231 (0.052)

Table 2. Average and standard deviation (in parenthesis) of NSEn

	O N=100	Z N=100	F N=100	N N=100	S N=100
EEG	0.629 (0.073)	0.634 (0.085)	0.538 (0.093)	0. 539 (0.091)	0.653 (0.069)
Delta	0.375 (0.09)	0.393 (0.097)	0.419 (0.096)	0.425 (0.097)	0.484 (0.095)
Theta	0.563 (0.074)	0.58 (0.063)	0.567 (0.071)	0.565 (0.067)	0.553 (0.074)
Alpha	0.641 (0.086)	0.69 (0.061)	0.684 (0.062)	0.685 (0.057)	0.667 (0.064)
Beta	0.762 (0.055)	0.761 (0.055)	0.79 (0.053)	0.74 (0.05)	0.759 (0.054)
Gamma	0.786 (0.051)	0.81 (0.05)	0.821 (0.054)	0.822 (0.045)	0.871 (0.041)

The optimized features extracted from dataset of 120 segments (including 60 epileptic (S) and 60 healthy or interictal (O-Z-F-N)) were used for training the classifier. Then the validation of classification was done on features from 80 segments of dataset (including 40 epileptic and 40 other states).

Minimum Redundancy - Maximum Relevance feature selection

In 2005 a new feature-selection method based on mutual information was proposed by Peng *et.al* [13]. The principle idea of this method is considering relevant and redundant features simultaneously. Based on Shannon's information theory, the mutual information between two random variables x and y is defined as follow:

$$I(X;Y)=\sum_{x\in X}\sum_{y\in Y}P(x,y)\log\frac{P(x,y)}{P(x)P(y)} \tag{5}$$

which p(x) and p(y) are probability distributions and p(x, y) is joint probability distribution. The relevance criterion of a typically set E for the class C is defined by the mean value of all mutual information values (6). In this case the mutual information is calculated between f_i and C. f_i is a matrix which its dimension is 100 * 32 (Its rows consist of segments that each segments contains 32 features). C is a 100*1 matrix which expresses the corresponding class of each segment.

$$D(S,c)=\frac{1}{|S|}\sum_{f_i\in S}I(f_i;c) \tag{6}$$

The redundancy of all features in the typically set E is the mean value of all mutual information values between the features f_i and f_j:

$$R(S)=\frac{1}{|S|^2}\sum_{f_i,f_j\in S}I(f_i;f_j) \tag{7}$$

The MRMR criterion is obtained from subtracting redundancy from relevancy measures by following equation:

$$\max_S \phi_S(D,R)=\frac{1}{|S|}\sum_{f_i\in S}I(f_i;c)-\frac{1}{|S|^2}\sum_{f_i,f_j\in S}I(f_i;f_j) \tag{8}$$

The optimized features are selected by finding the subset E which maximizes the MRMR criterion. The MRMR maximization is obtained by maximizing $D(E)$ and minimizing $R(E)$.

Results of the KNN classifier based on mutual information for NShEn and NSEn values are shown in table 3 and 4 respectively. In this process classification is done for ictal state (S) and other states (healthy (O and Z) and interictal (F and N)):

Table 3. Result of KNN classifying of EEG and its sub-bands based on NShEn values.

	O-S	O-Z	F-S	N-S
EEG	72	75	78.5	75
Delta	63.25	64	73.75	66.5
Theta	70.25	80.75	78.5	78
Alpha	85.75	82	79.5	89.5
Beta	94.5	94.25	92.5	95.5
Gamma	99.75	99.25	94.5	99.85

Table 4. Result of KNN classifying of EEG and its sub-bands based on NSEn values.

	O-S	Z-S	F-S	N-S
EEG	72.75	73.5	96.85	97.5
Delta	84.75	81.8	83.75	84.5
Theta	69.75	82	84.75	80
Alpha	74.75	79.7	86.25	84
Beta	70	73.6	84.75	83.8
Gamma	98.5	96.5	97.5	97.7

Conclusion

In this study we applied a mixed methodology based on chaotic - KNN classifying - mutual information analysis. By means of mutual information analysis the results of KNN classifier are improved approximately 3-4 percent. The classifying results that are collected in tables 3 and 4 show improvement than results that were done with other researchers in recent years. In recent years some results are reported based on chaotic analysis like; approximate, Shannon, spectral entropy and some other methods. In 2008 Kumar et.al could classify the ictal state from normal state with 99.75% and ictal state from interictal state with 96.3% accuracy based on Shannon entropy. In 2009 Ocak could classify the ictal subjects in 43.4-86.8 Hz sub-band with 96% accuracy based on approximate entropy. In this study we compare ictal state with both healthy and interictal states based on mutual information analysis. According to results the best classification based on Shannon entropy is occurred in gamma sub-band that compare ictal and non-focal interictal with 99.85% accuracy, also in gamma sub-band between ictal and healthy with eyes open state with 99.75% accuracy. Also based on spectral entropy the best result is occurred in gamma sub-band for comparison of ictal state and healthy with eyes open with 98.5% accuracy.

REFERENCES

[1] H. Witte, L. D. Iasemidis, and B. Litt, "Special isse on epileptic seizures prediction," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 5, pp. 537-539, May 2003.

[2] L. D. Iasemidis and J. C. Sackellares, "The temporal evolution of the largest Lyapunov exponent on the human epileptic cortex," in Measuring Chaos in the Human Brain, D. W. Duke and W. S. Pritchard, Eds. Singapore: World Scientific, 1991, pp. 49-82.

[3] L. D. Iasemidis, D. S. Shiau, J. C. Sackellares, and P. M. Pardalos, "Transition to epileptic seizures: Optimization," in *DIMACS Series in Discrete Mathematics and Theoretical Computer Science.* Providence, RI: American Mathematical Society, 2000, vol. 55, pp. 55-74.

[4] L. D. Iasemidis, D. S. Shiau, W. Chaovalitwongse, J. C. Sackellares, P. M. Pardalos, J. C. Principle, P. R. Carney, A. Prasad, B. Veeramani, and K. Tsakalis, "Adaptive epileptic seizure prediction system," IEEE Trans. Biomed. Eng., vol. 50, no. 5, pp. 616-627, May 2003.

[5] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A wavelet-chaos methodology for analysis of EEGs and EEG sub-bands to detect seizure and epilepsy," *IEEE Trans. Biomed. Eng, vol. 54, no. 2,* pp. 205–211, Feb. 2007.

[6] H. Adeli, Z. Zhou, and N. Dadmehr, "Analysis of EEG records in an epileptic patient using wavelet transform," *J. Neurosci. Meth, vol. 123,* pp. 69–87, 2003.

[7] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Mixed-Band Wavelet-Chaos-Neural Network Methodology for Epilepsy and Epileptic Seizure Detection," *IEEE Transactions on Biomedical Engineering, vol. 54,* No. 9, September 2007.

[8] S. P. Kumar, N. Sriraam and P. G. Benakop, "Automated Detection of Epileptic Seizures Using Wavelet Entropy Feature with Recurrent Neural Network Classifier," *TENCON IEEE*, pp. 1-5. 2008.

[9] H. Ocak, " Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy," Elsevier Ltd, vol. 36. Pp. 2027-2036. 2009.

[10] R. G. Andrzejak, K. Lehnertz, C. Rieke, F. Mormann, P. David, and C. E. Elger, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state, *Physical Review E, vol. 64, 061907,* Nov. 2001. http://www.meb.uni-bonn.de/epileptologie/science/physik/eegdata.html.

[11] C. E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J. 27* (1948) 379-423.

[12] R. Ferenets, T. Lipping, A. Anier, V. Jäntti, S.Melto, and S. Hovilehto, Comparison of Entropy and Complexity Measures for the Assessment of Depth of Sedation, *IEEE Transactions on Biomedical Engineering, vol. 53,* No. 6, June 2006.

[13] H. Peng, F. Long, and C. Ding. "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, No. 8, pp, 1226-1238, 2005.

***Authors**: Ahmad Mirzaei, Iran University of Science and Technology, E-mail: amce_500@yahoo.com: dr. Ahmad Ayatollahi, Iran University of Science and Technology, E-mail: ayatollahi@iust.ac.ir: dr. Ali Motie Nasrabadi, Shahed university, E-mail: Nasrabadi@shahed.ac.ir*