International Journal of Neural Systems

APPLICATION OF NON-LINEAR AND WAVELET BASED FEATURES FOR THE AUTOMATED IDENTIFICATION OF EPILEPTIC EEG SIGNALS

U RAJENDRA ACHARYA*
Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore 599489
S VINITHA SREE
Global Biomedical Technologies, CA, USA
ANG PENG CHUAN ALVIN
Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore 599489
RATNA YANTI
Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore 599489
JASJIT S. SURI
Fellow AIMBE, CTO, Global Biomedical Technologies, CA, USA and Biomedical Engineering Department, Idaho State University (affl.), ID, USA (jsuri@comcast.net)
*Corresponding author: U R Acharya: Email: aru@np.edu.sg

Epilepsy, a neurological disorder, is characterized by the recurrence of seizures. Electroencephalogram (EEG) signals, which are used to detect the presence of seizures, are non-linear and dynamic in nature. Visual inspection of the EEG signals for detection of normal, interictal, and ictal activities is a strenuous and time-consuming task due to the huge volumes of EEG segments that have to be studied. Therefore, non-linear methods are being widely used to study EEG signals for the automatic monitoring of epileptic activities. The aim of our work is to develop a Computer Aided Diagnostic (CAD) technique with minimal pre-processing steps that can classify all the three classes of EEG segments, namely normal, interictal, and ictal, using a small number of highly discriminating non-linear features in simple classifiers. To evaluate the technique, segments of normal, interictal, and ictal EEG segments (100 segments in each class) were used. Non-linear features based on the Higher Order Spectra (HOS), two entropies, namely the Approximation Entropy (*ApEn*) and the Sample Entropy (*SampEn*), and Fractal Dimension and Hurst Exponent were extracted from the segments. Significant features were selected using the ANOVA test. After evaluating the performance of six classifiers (Decision Tree, Fuzzy Sugeno Classifier, Gaussian Mixture Model, K-Nearest Neighbor, Support Vector Machine, and Radial Basis Probabilistic Neural Network) using a combination of the selected features, we found that using a set of all the selected six features in the Fuzzy classifier resulted in 99.7% classification accuracy. We have demonstrated that our technique is capable of achieving high accuracy using a small number of features that accurately capture the subtle differences in the three different types of EEG (normal, interictal, and ictal) segments. The technique can be easily written as a software application and used by medical professionals without any extensive training and cost. Such software can evolve into an automatic seizure monitoring application in the near future and can aid the doctors in providing better and timely care for the patients suffering from epilepsy.

Keywords: epilepsy, Higher Order Spectra, non-linear analysis, entropies, ictal, interictal, electroencephalograms.

1. Introduction

Epilepsy is a common neurological disorder that causes the patient to have repeated seizures. Globally, between 4 and 10 per 1000 people are estimated to suffer from active epilepsy wherein the patients have continuing seizures and need treatment. In developed countries, between 40 and 70 per 100 000 people are diagnosed with epilepsy every year. However, in the case of developing countries, this figure is often close to twice as high due to the higher risk of experiencing conditions that can lead to permanent brain damage.[63] Early and effective detection of epileptic activity will aid the clinicians in monitoring seizures and in administering appropriate seizure management protocols in order to improve the quality of life of epileptic patients. The Electroencephalogram (EEG) signals are widely used to investigate brain disorders and to study brain electrical activity.[31,47,55,56] The clinicians study the EEG segments for three types of electroencephalographic changes: ictal, which is usually characterized by continuous rhythmical activity that has a sudden onset when the patient is exhibiting a seizure; interictal, which is characterized by small spikes and subclinical seizures that generally occur during the time between seizures in epileptic patients; and normal EEG segments. Epilepsy monitors are commonly used to record long periods of EEG data for presurgical evaluation of epilepsy patients based on the occurrence of normal, interictal, and ictal activities. Detection of these activities by visual

inspection of the EEG signals is a strenuous and time-consuming task due to the huge volumes of EEG segments that have to be studied. Moreover, clinicians doing visual inspection require expert training and good experience in order to make credible predictions and avoid inter-observer variability. These issues can be addressed by developing Computer Aided Diagnostic (CAD) techniques that study the EEG segments and output the type of activity. Such automated techniques generally offer a faster, easier, more effective and affordable way to study epileptic activity.

Many automated CAD techniques extract linear time-domain and frequency-domain based features from the EEG signals and use these features in classification models in order to determine the optimal feature subset and classifier combination that present the highest accuracy. However, since EEG signals are by nature non-linear,[6-7,9,12,14-17,54,55] most of the recent studies[1, 21,36,41,60] extracted non-linear features and used them in classifiers. The use of non-linear features based on Higher Order Spectra (HOS) of the EEG signals has been shown to provide good classification accuracy.[21-23] Moreover, several studies[37,48] have also evaluated the capabilities of entropies in detecting seizures. There are a few studies that extract features from the wavelet transform applied signals. These studies either focused on classifying only the normal and ictal classes[48,53,59] or they used time frames or sub-bands of data.[8,27-30] Details on the techniques employed in such CAD based epilepsy detection studies are provided in the discussion section of this paper. The results of these studies indicate that there is still room for improvement of classification accuracy. The aim of our work is to develop a technique with minimal pre-processing steps that can classify all the three classes of EEG segments, namely normal, interictal, and ictal, using a small number of highly discriminating non-linear features in simple classifiers. Such an automated technique can detect epilepsy because of its ability to classify normal and interictal activities, and can detect seizures in epilepsy monitoring units because of its capability to classify interictal and ictal segments. Thus, the objective of the work is to present to the clinicians an automated, simple, objective, fast, and cost-effective efficient secondary diagnostic tool that can aid in real-time monitoring of EEGs in order to provide additional confidence to their initial diagnosis of the class of the EEG segment. After sufficient clinical trials, such a tool can evolve into a primary diagnostic tool for detection of epileptic EEG signals.

The proposed technique, called (*IntelligentWatch*), is illustrated by a block diagram in Fig. 1. The offline system indicates the steps used for developing the classifiers. In this case, part of the original dataset, called the training dataset, is used for feature extraction. Features include entropies, HOS based features, Hurst exponent (H), and Fractal Dimension (FD). Highly discriminating features are selected using the Analysis of Variance (ANOVA) test and these feature vectors along with the ground truth that dictates the class of the sample are used to train classifiers. The classifiers learn the training parameters that best predict the class of the input sample. In the online real-time system that would be the end-product used in hospitals, no training of the classifiers is required. The significant features found in the training phase are extracted from the test sample, and the training parameters are applied on this feature vector to classify it into one of the three classes. The performances of the classifiers are reported as their accuracy, sensitivity, specificity, and Positive Predictive Value (PPV) registered for the test dataset.

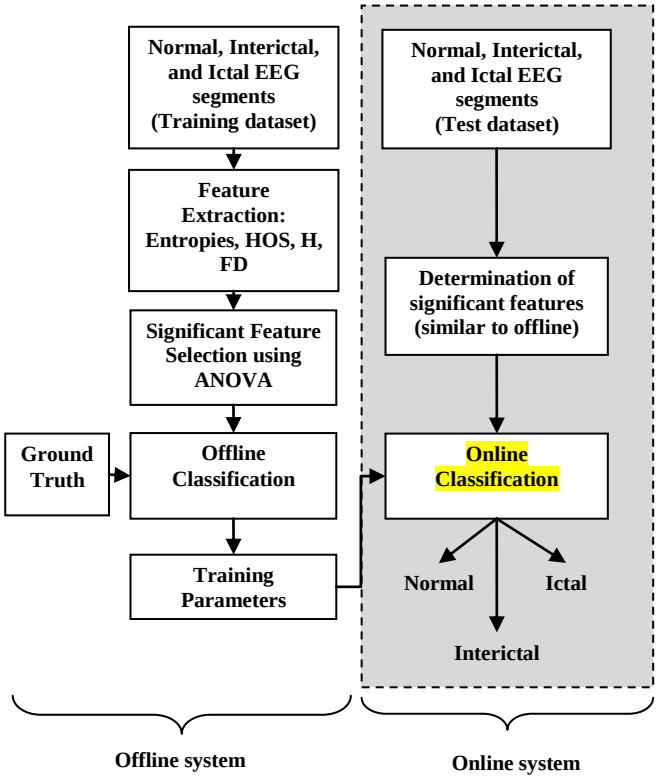

Fig. 1 Block diagram of the proposed system (*IntelligentWatch*) for EEG classification; the blocks outside

the dotted shaded rectangular box represent the flow of offline training system, and the blocks within the dotted box represent the online real-time system.

The key novelty of this paper lies in the determination of highly discriminative combination of non-linear features that improve the classification accuracy to 99.7%. Such a high accuracy has not been reported by any studies in the literature. This paper is organized as follows: Section 2 presents the description of the data used in this work and provides brief description of the features and the feature selection technique. The classifiers and classification methodology are presented in Section 3. The feature selection results and the classification results are presented in Section 4. Section 5 presents a discussion on related studies and compares our results with other published work. The paper is concluded in Section 6.

2. Data and Methods

2.1. *EEG Data*

The EEG dataset used in this work was taken from the artifact free EEG time series data (Sets A, D, and E) available at the University of Bonn [see Andrzejak *et al.*[19] for more details]. EEG segments, each of length 23.6 seconds, were taken from five healthy subjects and five epileptic patients. For each of the three categories: normal (EEG segments taken from healthy subjects), interictal, and ictal/epileptic, 100 segments of the data were selected. Each EEG segment was considered as a separate EEG signal resulting in a total of 300 EEGs. The standard surface electrode placement scheme (the international 10-20 system) was used to obtain the normal segments from the five healthy cases (Set A recorded when patients were in the awake state with eyes open). Both the interictal and ictal segments were obtained from the five epilepsy patients undergoing presurgical evaluations. The interictal segments were recorded during seizure free intervals from the depth electrodes that were implanted into the hippocampal formations (Set D). The ictal segments were recorded from all sites exhibiting ictal activity using depth electrodes and also from strip electrodes that were implanted into the lateral and basal regions of the neocortex (Set E).[19] All the segments were recorded using a 128-channel amplifier system, digitized with a sampling rate of 173.61 Hz and 12-bit A/D resolution, and filtered using a 0.53~40Hz (12 dB/octave) band pass filter. Typical normal, interictal, and epileptic (ictal) EEG signals are illustrated in Fig. 2. In this work, we selected 6 seconds data from each segment to develop an accurate classification framework for shorter segments.

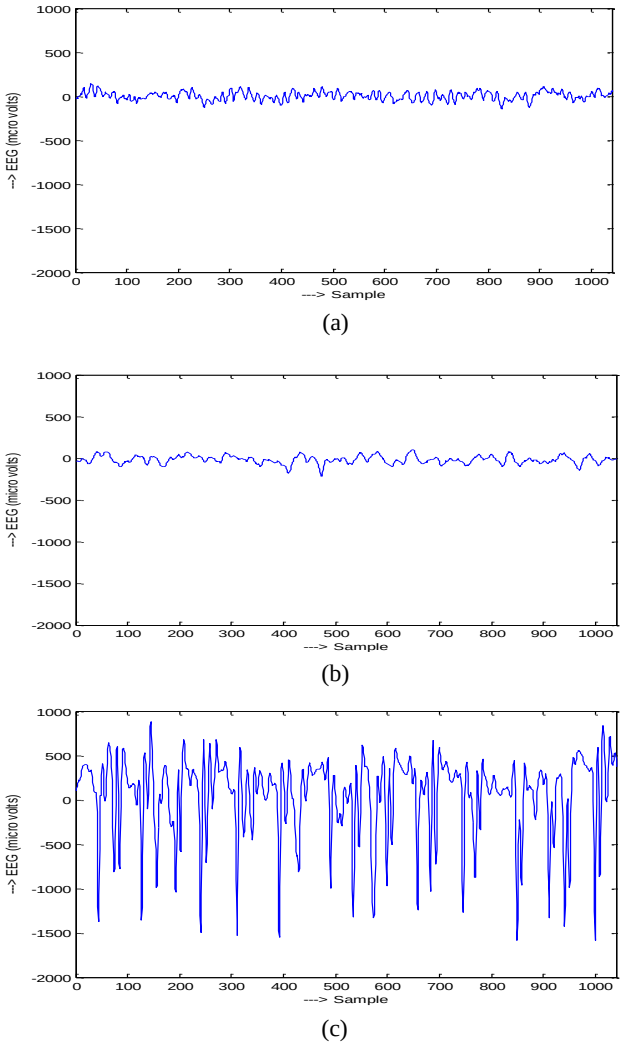

Fig.2. EEG signals[19] (a) normal (b) ictal and (c) interictal

2.2. *Feature Extraction*

Entropies

Approximate Entropy *(ApEn)*

Approximate Entropy is a measure of data regularity as it is defined as the logarithmic likelihood that the trends of the data patterns that are close to each other will remain close in the next comparison with next pattern. Highly regular data result in producing smaller *ApEn* values and vice versa. *ApEn*, proposed by Pincus *et al.*[48,49] detects changes in the underlying episodic

behaviour that is not reflected in the amplitude or peaks of the signal. In order to calculate *ApEn*, consider a time series *x(n), n= 1,2…,N*. A series of patterns of length *e* (called the embedding dimension which is the smallest integer for which the patterns do not intersect with each other) is derived from *x(n)*. *ApEn* is given by

$$ApEn(e,r,N)=\frac{1}{(N-e+1)}\sum_{i=1}^{N-e+1}\log C_i^e(r)-\frac{1}{(N-e)}\sum_{i=1}^{N-e}\log C_i^{e+1}(r) \quad (1)$$

where the index *r* is a fixed parameter which sets the *tolerance* of the comparison and $C_i^e(r)$ is the correlation integral given by

$$C_i^e(r)=\frac{1}{(N-e+1)}\sum_{j=1}^{N-e+1}\Theta\left(r-\left\|x_i-x_j\right\|\right) \quad (2)$$

where Θ is the Heaviside step function, $\Theta(a)=0\forall a\le 0$ and $\Theta(a)=1\forall a>0$. We chose *r* as 0.2 times the standard deviation of the time series, and *e* as 1.[62] Since *ApEn* is very insensitive to low level noise it is very suitable for EEG signal analysis.

Sample Entropy (*SampEn*)

Sample Entropy, proposed by Richman *et al.*,[52] is the negative natural logarithm of an estimate of the conditional probability that patterns of length *e* that match point-wise within a tolerance *r* also match at the next point.[50] Even though *SampEn* is also a measure of data regularity like *ApEn*, *SampEn* is mostly independent of record length and displays relative consistencies under circumstances where *ApEn* does not.[52] To calculate the sample entropy, runs of points matching within the tolerance *r* are carried out until there is no match, while the count of template matches are stored in counters *A(k)* and *B(k)* for all lengths *k* up to *e*. *SampEn* is given by the formula

$$SampEn(k,r,N)=\ln\left(\frac{A(k)}{B(k-1)}\right) \quad (3)$$

where *k = 0, 1,..., e-1* and $B(0)=N$, the length of the input series. The *k* value was empirically chosen to be 5 in this work.[57]

Higher Order Spectra (HOS)-based features

Higher Order Spectra based features quantify the nonlinear behaviour of a process.[44,45] HOS features comprise of higher order moments (order greater than two) and nonlinear combinations of these higher order moments, called the higher order cumulants. The bispectrum, which is the spectrum of the third order cumulants, is one of the most commonly used HOS features. It is a complex valued function of two frequencies given by

$$B(f_1,f_2)=E[X(f_1)\cdot X(f_2)\cdot X^*(f_1+f_2)] \quad (4)$$

where *X(f)* is the Fourier transform of the signal and X^* indicates the complex conjugate. As per the equation, the bispectrum is the product of the three Fourier coefficients. The function exhibits symmetry, and is computed in the non-redundant/ principal domain region Ω as shown in Fig. 3.

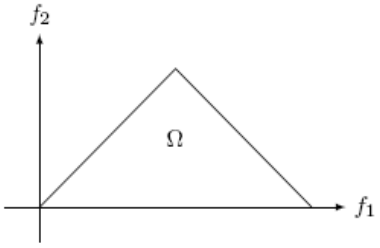

Fig. 3. Principal domain region (Ω) used for the computation of the bispectrum for real signals.

The Bispectral Phase Entropy (BPE),[21-23] obtained from the bispectrum, is defined as

$$BPE=\sum_n p(\psi_n)\log p(\psi_n) \quad (5)$$

where

$$p(\psi_n)=\frac{1}{L}\sum_{\Omega} I(\phi(B(f_1,f_2))\in\psi_n) \quad (6)$$

$$\psi_n=\{\phi\,|\,-\pi+2\pi n/N\le\phi<-\pi+2\pi(n+1)/N\},\quad n=0,1,\ldots,N-1 \quad (7)$$

where *L* is the number of points within the region Ω, ϕ is the phase angle of the bispectrum, and *I(.)* is an indicator function which gives a value of 1 when the phase angle is within the range depicted by ψ_n in equation (7). We also calculated two phase entropies which are similar to the bispectral phase entropy. They

are calculated as follows.[21-22]

Phase Entropy 1: $PhaseEnt1 = -\sum_k p_k log p_k$ (8)

where $p_k = \frac{|B(f_1, f_2)|}{\sum_\Omega |B(f_1, f_2)|}$

Phase Entropy 2: $PhaseEnt2 = -\sum_i q_i log q_i$ (9)

where $q_i = \frac{|B(f_1, f_2)|^2}{\sum_\Omega |B(f_1, f_2)|^2}$

In the above equations, the magnitude and the square of the magnitude are the L1 and L2 norms of the bispectrum and they were normalized by the sum of the norm over Ω such that each norm is now similar to a Probability Distribution Function (PDF) with values estimated from the Ω region. Due to the different non-linear dynamics existing in the three EEG classes, the normal, interictal, and ictal EEG are expected to have different PDF profiles of the bispectrum. We also calculated the Weighted Centre Of Bispectrum (*Wcob*),[65] as follows,

$$f_{1m} = \frac{\sum_\Omega iB(i,j)}{\sum_\Omega B(i,j)} \qquad f_{2m} = \frac{\sum_\Omega jB(i,j)}{\sum_\Omega B(i,j)} \tag{10}$$

where *i*, *j* are the frequency bin index in the non-redundant region, f_{1m} is *Wcobx*, and f_{2m} is *Wcoby*.

Higuchi Fractal Dimension

The concept of fractal dimension (FD) originates from fractal geometry.[42] FD is a powerful tool for transient detection and it can be used to measure the dimensional complexity of biological signals. It can give an indication of how completely the fractal appears to fill space. In this work, the algorithm proposed by Higuchi was used for finding FD of EEG signals.[34] The EEG segment was assumed as a time sequence $x(1), x(2) \ldots, x(n)$. Time series $x(k,m)$ may be constructed as:

$$x(k,m) = \{x(m), x(m+k), x(m+2k), \ldots, x(m + int[(N-m)/k]k)\} \tag{11}$$

where $m = 1, 2, \ldots, k$ and $int[\ ]$ is an integer function. Here m indicates the initial time value and k indicates the discrete time interval between points. The length $L(m,k)$ for each of the k time series or curves $x(k,m)$ was computed as:

$$L(m,k) = \left\{ \frac{\sum_{i=1, int[(n-m)/k]} |x(m+ik) - x[m+(i-1)k]|(n-1)}{int[(n-m)/k]k} \right\} \frac{1}{k} \tag{12}$$

where n is the total length of data sequence x. Mean value of the curve length $L(k)$ was calculated for each k by averaging $L(m,k)$ for all m. Thus, an array of mean values $L(k)$ was obtained. A plot of $log(L(k))$ versus $log(1/k)$ was made and FD was estimated from the slope of least squares linear best fit from the plot. Thus, FD can be defined as

$$FD = log(L(k))/log(1/k) \tag{13}$$

Hurst Exponent (H)

Hurst Exponent (H) is a measure of self-similarity, predictability and the degree of long-range dependence in a time-series. It is also a measure of the smoothness of a fractal time-series based on asymptotic behavior of the rescaled range of the process.[26] According to the Hurst's generalized equation of time series,[35] H is defined as

$$H = \frac{log\ (R/S)}{log\ (T)} \tag{14}$$

where T is the duration of the sample of data and R/S is the corresponding value of rescaled range. R is the difference between the maximum and minimum deviation from the mean while S represents the standard deviation.[24] In other words, Hurst exponent is given by,

$$(R/S)_T = c * T^H \tag{15}$$

Here, c is a constant and H is the Hurst exponent. Hurst exponent is estimated by plotting (R/S) versus T in log-log axes. The slope of the regression line approximates the Hurst exponent.

2.3. *Feature Extraction*

The extracted features are subjected to the one-way Analysis of Variance (ANOVA) test. In this technique, the variation of a feature between classes and the variation within a class are calculated. If the between-class variance is higher than the within-class variance,

the feature is considered to be statistically significant (low value of p).

3. Classification

3.1. *Classifiers used*

In this work, we evaluated six classifiers namely Decision Tree (DT), Fuzzy Sugeno Classifier, Gaussian Mixture Model (GMM), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Radial Basis Probabilistic Neural Network (RBPNN). They are briefly described in this section.

3.1.1. *Decision Tree (DT)*

A DT classifier[40] is a decision support classifier that uses trees built using the input training dataset features. A series of rules are extracted from the tree and these rules are used to recognize the class of the test data. The parent node and leaf node values should be tuned, and in this work, we found that the parent and leaf node values of 6 and 1 respectively resulted in the highest accuracy.

3.1.2. *Fuzzy Sugeno Classifier*

Fuzzy inference is a technique wherein the values in the input vector are interpreted based on a set of rules, and based on the interpretation, the input values are mapped to the output vector. In this work, during training, the subtractive clustering technique was used to generate a Fuzzy Inference System (FIS).[61] This FIS structure contains the inputs, outputs, and a set of fuzzy rules that cover the feature space. These rules are then used to perform fuzzy inference calculations of the test data. In the FIS, the clusters were fixed using radii. It is a vector that specifies a cluster center's range of influence in each of the data dimensions, assuming the data falls within a unit hyperbox. Small radii values generally result in finding a few large clusters. The best values for radii are usually between 0.1 and 0.5. In our work, we have selected the radii of 0.4, the input membership function as 'Gaussian' and the output membership function as 'linear'.

3.1.3. *Gaussian Mixture Model (GMM)*

In the GMM classifier, the PDF of each class is assumed to consist of a mixture of multidimensional Gaussian distributions. The expectation-maximization algorithm[33] is used to determine the parameters of the each Gaussian component and the weights of the mixtures based on the training data. The trained GMMs are used to classify the test data. In this work, we used a GMM containing five cluster centers.

3.1.4. *K-Nearest Neighbor (KNN)*

KNN is a supervised learning algorithm[39] that uses the class labels of the neighbours of the new test data to predict the class of the test data. In the case of new test sample, *K* number of training points closest to the test sample are evaluated and the class that is most common amongst these *K* nearest neighbours is assigned as the class to the new test data. In this work, we varied the *K* nearest neighbors *K*=2 to 6. We obtained maximum accuracy for *K*= 2. The distance was computed using Euclidean distance.

3.1.5. *Support Vector Machine (SVM)*

Support Vector Machine (SVM) is a supervised classifier whose main objective is to find a separating hyperplane that separates the training samples belonging to the three classes which are viewed in an *n*-dimensional space (*n* is the number of features used as inputs) with a maximum margin between the hyperplane and the samples closest to the hyperplane (called the support vectors).[25] If the data are not linearly separable, kernel functions can be used to map the original feature space to a higher dimensional feature space where the features become linearly separable. [43] In this work, we used the radial basis function kernel that has two training parameters: cost (C) which controls over-fitting of the model, and sigma (σ) which controls the degree of nonlinearity of the model. Using a grid search approach, we found that the accuracy was highest when C was 35, and sigma was set to 1.

3.1.6. *Radial Basis Probabilistic Neural Network (RBPNN)*

RBPNN is a two-layer radial basis network[18,38,64] used for classification.[33] The first layer has radial basis activation functions which calculate the distance from the input test sample to the training input vectors and yields a distance vector. The next layer is the competitive layer which adds these distance vectors for each input classes and produces a vector of probabilities as its output.[10, 13] The ‘compete’ transfer function at the

output of the second layer selects the maximum of these probabilities, and assigns a 1 for that selected class and a 0 for the other classes. All biases in the radial basis layer were set to $\sqrt{ln0.5}/s$, where s=spread constant of the RBPNN. In this work, we obtained maximum accuracy when s=0.5.

3.2. *Classification methodology*

We have used a total of 300 EEG time series data segments (100 segments from each class). Seven significant features (as indicated in Table 1 in the Results section) were selected from these data to form the classification dataset. This classification dataset is used to build the classifiers. In order to develop robust classifiers that are adequately generalized to perform well when new samples are input, we chose the ten-fold cross validation data resampling technique for training and testing the classifiers. In this technique, the data is split into ten parts such that each part contains approximately the same proportion of class samples as the classification dataset. Nine parts of the data are used for training the classifier, and the remaining one part for testing. This procedure is repeated nine more times using a different part for testing in each case. Sensitivity, specificity, accuracy, and PPV are then calculated based on the results obtained for the test data in each folds. The average values of these performance measures over the ten folds are taken as the actual estimates of the classifier performance.

4. Results

We calculated *ApEn*, *SampEn*, and in the case of HOS features, we found the *BPE*, *PhaseEnt1*, *PhaseEnt2*, *Wcobx*, *Wcoby*, and we also determined Hurst Exponent (H) and Higuchi Fractal Dimension (FD). ANOVA was used to determine the significant features. Table 1 shows the range of selected six features for normal, interictal, and ictal classes. *ApEn* and *SampEn* were both found to be significant as indicated by the low p-values. Among the HOS features, *PhaseEnt2 and Wcoby* were found to be significant. Both H and FD were also significant features. The entropy features (*SampEn* and *ApEn*) have lower values for the abnormal (ictal and interictal) classes compared to the normal class. Such varied ranges for each feature indicate that various combinations of these features should be evaluated in each of the classifiers to determine the feature set and the classifier combination that presents the highest classification accuracy.

Table 2 shows the results of accuracy, sensitivity, specificity, and PPV recorded by the six classifiers using all the six features. Our results show that the Fuzzy classifier performs better than the other classifiers by registering 99.7% for accuracy and 100% for sensitivity, specificity, and PPV. The results of this classifier for various other combinations of features are listed in Table 3. It is evident that on using all the six significant features, the maximum accuracy of 99.7% was registered indicating that all the features have a significant part in classifying the segments accurately.

Table 1 Range (Mean ± Standard Deviation) of features for normal, interictal, and ictal EEG classes.

Features	Normal	Interictal	Ictal	p-value
Entropies				
ApEn	2.202 ±3.546E-02	1.8178 ± 0.170	1.8343 ± 0.209	<0.0001
SampEn	1.3180 ± 0.114	1.0301 ± 0.133	0.91719 ± 0.122	<0.0001
HOS				
PhaseEnt2	0.45268 ±8.433E-02	0.28522 ± 8.369E-02	0.33037 ± 0.179	<0.0001
Wcoby	10.365 ± 10.7	3.3568 ± 1.70	8.3844± 2.55	<0.0001
FD	-1.4695 ± 9.093E-02	-1.2919 ± 8.128E-02	-1.3286 ± 0.124	<0.0001
H	0.66828 ± 8.706E-02	0.79709± 5.445E-02	0.52875 ± 0.122	<0.0001

Table 2 The sensitivity, specificity, accuracy and PPV values presented by the six classifiers using all six features for training and testing (Mean ± Standard Deviation)

Classifiers	Acc (%)	PPV (%)	Sn (%)	Sp (%)
DT	97.3±2.63	97.2±3.25	99±2.11	94±6.99
PNN	97.7±2.74	98.1±2.49	99±2.11	96±5.16
KNN	97.7±2.74	98.1±2.49	99±2.11	96±5.16
FUZZY	99.7±1.05	100±0	100±0	100±0
GMM	99.3±1.41	99±2.01	100±0	98±4.22
SVM	99±1.61	99±2.01	99.5±1.58	98±4.22

Table 3 The sensitivity, specificity, accuracy and positive predictive value registered by the Fuzzy classifier for various feature combinations (Mean ± Standard Deviation)

Feature set	Acc (%)	PPV (%)	Sn (%)	Sp (%)

Entropies	81.3±5.49	98.9±2.28	89.5±5.99	98±4.22
HOS	82±5.26	95.2±5.37	95±4.08	90±11.55
FD + H	82±9.05	86.5±10.82	88.5±6.69	69±27.67
Entropies + HOS	92±4.22	99±2.01	98.5±2.42	98±4.22
HOS + FD + H	93.3±4.16	94.7±4.74	96±3.16	89±9.94
Entropies + FD + H	98.3±1.76	99±2.01	98.5±2.42	98±4.22
All features	99.7±1.05	100±0	100±0	100±0

5. Discussion

Several studies report the use of linear and non-linear features for the automatic detection of the three EEG classes. A summary of the methodology employed in these studies and the classification accuracies obtained is presented in this section. All these studies used the Bonn University dataset.

Entropy-based studies: Several studies have demonstrated that the entropies extracted from the ictal segments were generally lower than those from the normal class.[20,37] A similar trend was observed in this work (Table 1). Ocak *et al.*[46] used *ApEn* and DWT to classify epileptic and normal EEG signals. The accuracy was 96% when *ApEn* was extracted from the DWT applied EEG signal, and it was 73% when it was calculated from raw EEG data. Kannathal *et al.*[36] used non-linear features like Correlation Dimension (CD), Largest Lyapunov Exponent (LLE), Hurst exponent (H), and entropy to detect normal and epileptic EEG signals with an accuracy of 92.2%. Using various entropies in an Adaptive Neuro Fuzzy Inference System (ANFIS), the same group[37] obtained an accuracy of more than 90% for differentiating normal and epileptic EEG signals.

Wavelet-based studies: Adeli *et al.*[5,11] used discrete Daubechies and harmonic wavelets to analyze and characterize epileptiform discharges in patients in the absence of seizure. They captured and localized transient features in both time and frequency domain using their method. A novel wavelet-chaos-neural network methodology was used for classification of healthy, ictal, and interictal EEG classes.[27] Wavelets were used to decompose the EEG signals into delta, theta, alpha, beta, and gamma sub-bands. Three features were then extracted from these sub-bands: standard deviation, CD, and LLE. They developed a Spiking Neural Network (SNN) using three training algorithms (SpikeProp, QuickProp, and RProp). On using RProp as the training algorithm, a classification accuracy of 92.5% was obtained using the mixed-band feature space consisting of nine parameters (CD and LLE from wavelet coefficients). The same group obtained a higher accuracy of 96.7% using the Levenberg-Marquardt back propagation neural network and a mixed-band feature space consisting of nine parameters[28]. Again, the same group[30] developed a Multi-Spiking Neural Network (MuSpiNN) and obtained an accuracy range of 90.7%-94.8% using the mixed-band feature space. On applying PCA to the nine-parameter mixed-band feature space, they obtained an accuracy of 96.6% using a cosine radial basis function neural network (RBFNN).[29] In another study,[7] CD and LLE features were extracted from wavelet based EEG sub-bands to detect the three classes. They did not observe significant differences in the values of the features obtained from the original EEG signals. However, when these features were obtained from the specific EEG sub-bands, differences were evident. They found that for the higher frequency beta and gamma sub-bands, the CD feature differentiates between the three groups, whereas for the lower frequency alpha sub-band, the LLE feature differentiates between the three groups.

HOS-based studies: HOS features were found to be significant enough for differentiating normal, interictal and epileptic EEG signals.[21] On using the HOS features in GMM and SVM classifier, accuracies of 93.11% and 92.67%, respectively, were reported.[22] The same group[23] calculated entropies from the power spectra of the segments and used them in GMM classifier and obtained a lower accuracy of 88.78%. The HOS cumulants and entropies were also found to be significant features. In our recently published work[4], we used the HOS cumulants extracted from Wavelet Packet Decomposition (WPD) coefficients as features in SVM classifier and obtained an accuracy of 98.5%. In this work, we improved the accuracy to 99.7% by using entropy based features, Hurst Exponent, and Higuchi FD in addition to the HOS based features.

Other non-linear features: Guler *et al.*[32] used Lyapunov features in a Recurrent Neural Network (RNN) for classifying the three classes with an efficiency of more than 96%. Recently, chaotic features like CD, Hurst exponent, LLE, fractal dimension, and *ApEn* were used to extract the important hidden features from the normal, interictal and ictal EEG signals.[2] On using the significant features in the GMM classifier, a

classification efficiency of 95% was obtained. In a study by Raiesdana *et al.*,[51] Recurrence Quantification Analysis (RQA) measures were used to study the effect of seizures on the recorded EEG signals. They observed strong rhythmicity in the recurrence plots and RQA measures when seizures occurred. Our group recently[3] studied on the use of RQA parameters for classification of the three classes of EEG segments. We were able to use ten significant RQA parameters in SVM classifier to obtain a classification accuracy of 95.6%. Table 4 presents a summary of these studies.

Table 4 Summary of the non-linear studies on automated identification of normal, interictal, and ictal EEG segments

Authors	**Features (No. of features)**	**Classifier**	**Accuracy (%)**
Ghosh-Dastidar *et al.*[27]	Mixed-band feature space (9)	Spiking Neural Network	92.5
Ghosh-Dastidar *et al.*[28]	Mixed-band feature space (9)	Back Propagation Neural Network	96.7
Ghosh-Dastidar *et al.*[29]	Mixed-band feature space (9)	Radial Basis Function Neural Network	96.6
Ghosh-Dastidar *et al.*[30]	Mixed-band feature space (9)	Multi-Spiking Neural Network	90.7-94.8%
Guler *et al.*[32]	Lyapunov exponents (4)	Recurrent Neural Network	96.79
Chua *et al.*[22, 23]	HOS based features (3)	SVM and GMM	93.11
Acharya *et al.*[2]	Non-linear features (5)	SVM and GMM	95
Acharya *et al.*[3]	RQA parameters (10)	SVM	95.6
Acharya *et al.*[4]	HOS cumulants from WPD coefficients (4)	SVM	98.5
This work	Entropies+HOS+Higuchi FD+Hurst (6)	Fuzzy	99.7

It is evident from the results of these studies that the accuracy obtained is lower than the 99.7% accuracy obtained in this work. The following are the key observations and conclusions from this work:

a. We attribute the high accuracy to the highly discriminative features. During seizure, there is a sudden increase in neural discharge causing an increase in the variability of the EEG signal. This variability is better captured by the non-linear HOS features, entropies, Hurst exponent and FD. The Fuzzy classifier has superior approximation capabilities compared to crisp classifiers, and hence, it resulted in the highest accuracy in this work.
b. The number of features used to obtain such a high accuracy is only six. Moreover, there are no specific pre-processing steps required prior to feature extraction. Thus, the entire process is computationally simple and fast to carry out.
c. An accuracy that is close to our accuracy was obtained by Srinivasan *et al.*[58] In their work, they performed an automated classification of normal and epileptic states using Elman network and obtained an accuracy of 99.6%. However, we were able to achieve 99.7% accuracy in classifying all three classes of EEG segments, mainly attributed by the combination of various highly discriminative non-linear features.
d. Since cross validation was used as the resampling technique, the resultant classifiers are more robust compared to those classifiers that are trained using hold-out technique.
e. The technique is fully automated, can be easily written as a software application and installed, and is cost-effective. The results obtained are objective.

This preliminary study has shown promising results for EEG activity classification. However, in order to establish the effectiveness of this technique, we are planning to evaluate the technique using a larger multi-ethnic database. Over time, we expect more bench marking and diversification of data collection and its validation.

6. Conclusion

CAD based techniques are gaining more popularity due to their objective diagnostic capabilities, speed, efficiency, easy deployment and training, and cost-effectiveness. Such techniques give a valuable second opinion to the doctors on important diagnoses. In this work, we have presented a CAD technique that we developed for classification of normal, interictal, and ictal EEG segments. This technique uses non-linear

features extracted from the EEG segments in a Fuzzy classifier to classify the segments. The built Fuzzy classifier has demonstrated that it is capable of classifying the three classes with 99.7% accuracy. Such a high accuracy has not been reported by previously published studies that used linear and non-linear features for EEG epileptic activities classification. More studies will be carried out with larger databases to evaluate our technique and also to establish the significance of the features we discovered in this work.

References

1. U. R. Acharya, O. Faust, N. Kannathal, T. J. Chua and S. Laxminarayan, Non-linear analysis of EEG signals at various sleep stages, *Computer Methods and Programs in Biomedicine* **80**(1) (2005) 37–45.
2. U. R. Acharya, K. C. Chua, T. C. Lim, Dorithy and J. S. Suri, Automatic identification of epileptic EEG signals using nonlinear parameters, *Journal of Mechanics in Medicine and Biology* **9**(4) (2009) 539–553.
3. U. R. Acharya, S. Vinitha Sree, S. Chattopadhyay, W. Yu, and A. P. C. Alvin, Application of Recurrence Quantification Analysis for the automated identification of epileptic EEG signals, *International Journal of Neural Systems* **21**(3) (2011a) 199-211.
4. U. R. Acharya, S. Vinitha Sree, J.S. Suri, Automatic detection of epileptic EEG signals using higher order cumulant features, *International Journal of Neural Systems* **21**(5) (2011b) 1-12.
5. H. Adeli, Z. Zhou and N. Dadmehr, Analysis of EEG records in an epileptic patient using wavelet transform, *Journal of Neuroscience Methods* **123**(1) (2003) 69–87.
6. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, Alzheimer's disease and models of computation: Imaging, classification, and neural models, *Journal of Alzheimer's Disease* **7**(3) (2005a) 187–199.
7. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, Alzheimer's disease: Models of computation and analysis of EEGs, *Clinical EEG and Neuroscience* **36**(3) (2005b) 131–140.
8. 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 Transactions on Biomedical Engineering* **54**(2) (2007) 205–211.
9. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A Spatio-temporal wavelet-chaos methodology for EEG-based diagnosis of Alzheimer's disease, *Neuroscience Letters* **444**(2) (2008) 190–194.
10. H. Adeli and A. Panakkat, A probabilistic neural network for earthquake magnitude prediction, *Neural Networks* **22** (2009) 1018–1024.
11. H. Adeli, S. Ghosh and Dastidar, Automated EEG based diagnosis of neurological disorders: Inventing the future of neurology, CRC Press (2010).
12. M. Ahmadlou and H. Adeli, Wavelet synchronization methodology: A new approach for EEG-based diagnosis of ADHD, *Clinical EEG and Neuroscience* **41**(1) (2010a) 1–10.
13. M. Ahmadlou and H. Adeli, Enhanced probabilistic neural network with local decision circles: A robust classifier, *Integrated Computer-Aided Engineering* **17**(3) (2010b) 197–210.
14. M. Ahmadlou, H. Adeli and A. Adeli, New diagnostic EEG markers of the Alzheimer's disease using visibility graph, *Journal of Neural Transmission* **117**(9) (2010c) 1099–1109.
15. M. Ahmadlou, H. Adeli and A. Adeli, Fractality and a wavelet-chaos-neural network methodology for EEG-based diagnosis of autistic spectrum disorder, *Journal of Clinical Neurophysiology* **27**(5) (2010d) 328–333.
16. A. Ahmadlou, H. Adeli and A. Adeli, Fractality and a wavelet-chao methodology for EEG-based diagnosis of Alzheimer's disease, *Alzheimer Disease and Associated Disorders* **25**(1) (2011a) 85–92.
17. M. Ahmadlou and H. Adeli, Fuzzy synchronization likelihood with application to attention deficit hyperactivity disorder, *Clinical EEG and Neuroscience* **42**(1) (2011b) 6–13.
18. P. Anand, B. V. N. Siva Prasad and C. Venkateswarlu, Modeling and optimization of a pharmaceutical formulation system using radial basis function network, *International Journal of Neural Systems* **19**(2) (2009) 127–136.
19. R. G. Andrzejak, K. C. Lehnertz, F. Rieke, P. Mormann, 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* **E64** 2001 061907. (Website: http://www.meb.uni-bonn.de/epileptologie/science/physik/eegdataold.html, Last accessed on 11th Dec 2011)
20. D. Bai, T. Qiu and X. Li, The sample entropy and its application in EEG based epilepsy detection, *Journal of Biomedical Engineering* (2007) 200–205.
21. K. C. Chua, V. Chandran, U. R. Acharya and C. M. Lim, Analysis of epileptic EEG signals using higher order spectra, *Journal of Medical & Engineering Technology* **33**(1) (2009a) 42–50.
22. K. C. Chua, V. Chandran, U. R. Acharya and C. M. Lim, Automatic identification of epileptic EEG signals using

higher order spectra, *International Journal of Engineering in Medicine* **223**(4) (2009b) 485–495.

23. K. C. Chua, V. Chandran, U. Rajendra Acharya and C. M. Lim, Application of higher order spectra to identify epileptic EEG, *Journal of Medical Systems*, DOI: 10.1007/s10916-010-9433-z.
24. S. Dangel, P.F. Meier, H.R. Moser and S. H. Plibersek, Time series analysis of sleep EEG, *Computer Assisted Physics* **14** (1999), 93-95.
25. V. David and A. Sanchez, Advanced support vector machines and kernel methods, *Neurocomputing* **55**(1-2) (2003) 5-20.
26. J. Feder, Fractals, New York: Plenum Press, (1988).
27. S. Ghosh-Dastidar, H. Adeli, Improved spiking neural networks for EEG classification and epilepsy and seizure detection, *Integrated Computer-Aided Engineering* **14**(3) (2007a) 187–212.
28. 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* **54**(9) (2007b) 1545–1551.
29. S. Ghosh-Dastidar, H. Adeli and N. Dadmehr, Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection, *IEEE Transactions on Biomedical Engineering* **55**(2) (2008) 512–518.
30. S. Ghosh-Dastidar and H. Adeli, A new supervised learning algorithm for multiple spiking neural networks with application in epilepsy and seizure detection, *Neural Networks* **22** (2009) 1419–1431.
31. L. B. Good, S. Sabesan, S. T. Marsh, K. S. Tsakalis and L. D. Iasemidis, Control of synchronization of brain dynamics leads to control of epileptic seizures in rodents, *International Journal of Neural Systems* **19**(3) (2009) 173–196.
32. N. F. Guler, E. D Ubey and I. Guler, Recurrent neural network employing Lyapunov exponents for EEG signals classification, *Expert Systems with Applications* **29**(3) (2005) 506–514.
33. J. Han, M. Kamber and J. Pei, *Data Mining: Concepts and Techniques*, 2nd edn. (Morgan Kaufmann, 2005).
34. T. Higuchi, Approach to an irregular time series on the basis of the fractal theory, *Physics D* **31**(2), (1988) 277-283.
35. H. E. Hurst, Long-term storage of reservoirs: an experimental study, *Transactions of the American society of civil engineers* **116** (1951) 770-799.
36. N. Kannathal, U. R. Acharya, C. M. Lim, Q. Weiming, M. Hidayat and P. K. Sadasivan, Characterization of EEG: A comparative study, *Computer Methods and Programs in Biomedicine* **80**(1) (2005a) 17–23.
37. N. Kannathal, C. M. Lim, U. R. Acharya and P. K. Sadasivan, Entropies for detection of Epilepsy in EEG, *Computer Methods and Programs in Biomedicine* **80**(3) (2005b) 187–194.
38. A. Karim and H. Adeli, Radial basis function neural network for work zone capacity and queue estimation, *Journal of Transportation Engineering* **129**(5) (2003) 494–503.
39. D. T. Larose, Discovering Knowledge in Data: An introduction to data mining, New Jersey, USA: Wiley Interscience, ch.5 (2004a).
40. D. T. Larose, Discovering Knowledge in Data: An introduction to data mining. New Jersey, USA: Wiley Interscience, ch.6 (2004b).
41. K. Lehnertz, Epilepsy and nonlinear dynamics, *Journal of Biological Physics* **33**(3–4) (2008) 253–266.
42. B.B. Mandelbrot, The Fractal Geometry of Nature, W.H. Freeman and Company (1982).
43. K. R. Muller, S. Mika, G. Ratsch, K. Tsuda and B. Scholkopf, An introduction to Kernel Based Learning Algorithms, *IEEE Transactions on Neural Networks* **12**(2) (2001) 181-201.
44. C. L. Nikias and A. P. Petropulu, Higher-order spectra analysis: A nonlinear signal processing framework, Englewood Cliffs, HJ: Prentice Hall (1993).
45. C. L. Nikias and J. M. Mendel, Signal processing with Higher-Order Spectra, *IEEE Signal Proc. Mag.* **10**(3) (1993) 10-37.
46. H. Ocak, Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy, *Expert Systems with Applications* 3**6**(2) (2009) 2027–2036.
47. J. P. Pijn, D. N. Velis, M. J. van der Heyden, J. DeGoede, C. W. van Veelen and F. H. Lopes da Silva, Nonlinear dynamics of epileptic seizures on basis of intracranial EEG recordings, *Brain Topography* **9**(4) (1997) 249–270.
48. S.M. Pincus, Approximate entropy as a measure of system complexity, in *Proc. of the National Academy of Sciences* **88** (1991), pp.2297-2301.
49. S.M. Pincus and D.L. Keefe, Quantification of hormone pulsatility via an approximate entropy algorithm, *American Journal of Physiology* **262**(5) (1992) E741–E754.
50. PhysioToolkit. http://physionet.incor.usp.br/physiotools/sampen/. (Last accessed on 11th Dec 2012)
51. S. Raiesdana, S. M. Golpayegani, S. M. Firoozabadi and J. Mehvari Habibabadi, On the discrimination of patho-physiological states in epilepsy by means of dynamical measures, *Computers in Biology and Medicine* **39**(12) (2009) 1073–1082.
52. J.S. Richman and J.R. Moorman, Physiological time-series analysis using approximate entropy and sample

entropy, *American Journal of Physiology* **278**(6) (2000) H2039–H2049

53. N. Sadati, H. R. Mohseni and A. Magshoudi, Epileptic seizure detection using neural fuzzy networks, in *Proc. of the IEEE International Conference on Fuzzy Syst.* (Canada, 16–21 Jul 2006), pp. 596–600.
54. Z. Sankari and H. Adeli, Probabilistic neural networks for EEG-based diagnosis of alzheimer's disease using conventional and wavelet coherence, *Journal of Neuroscience Methods* in press (2011a).
55. Z. Sankari, H. Adeli and A. Adeli, Intrahemispheric, interhemispheric and distal EEG coherence in alzheimer's disease, *Clinical Neurophysiology*, (2011b) doi:10.1016/j.clinph.2010.09.008.
56. A. Shoeb, J. Guttag, T. Pang and S. Schachter, Non-invasive computerized system for automatically initiating vagus nerve stimulation following patient specific detection of seizures or epileptiform discharges, *International Journal of Neural Systems* **19**(3) (2009) 157–172.
57. Y. Song and P. Liò, A new approach for epileptic seizure detection: sample entropy based feature extraction and extreme learning machine, *Journal of Biomedical Science and Engineering* **3**(6) (2010) 556-567.
58. V. Srinivasan, C. Eswaran and N. Sriraam, Artificial neural network based epileptic detection using time domain and frequency domain features, *Journal of Medical Systems* **29**(6) (2005) 647–660.
59. A. Subasi, EEG Signal classification using wavelet feature extraction and a mixture of expert model, *Expert Systems with Applications* **32**(4) (2007) 1084–1093.
60. D. P. Subha, K. P. Joseph, U. R. Acharya and C. M. Lim, EEG signal processing: A survey, *Journal of Medical Systems* **34**(2) (2010), DOI: 10.1007/s10916-008-9231-z.
61. M. Sugeno, Industrial applications of fuzzy control, Elsevier Science Pub. Co. (1985).
62. N.V. Thakor and S. Tong, Advances in quantitative electroencephalogram analysis methods, *Annual Reviews* **6** (2004) 453-495.
63. World Health Organization. Website: http://www.who.int/mediacentre/factsheets/fs999/en/index.html (Last accessed on 11th Dec 2011).
64. D. Wu, K. Warwick, Z. Ma, Gasson, M. N. Burgess, J. G. S. Pan and T. Z. Aziz, Prediction of parkinson's disease tremor onset using radial basis function neural network based on particle swarm optimization, *International Journal of Neural Systems* **20**(2) (2010) 109–116.
65. J. Zhang, C. Zheng and A. Xie, Bispectrum analysis of focal ischemic cerebral EEG signal, *IEEE Transactions on Biomedical Engineering* **47**(3) (2000) 352-359.