World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:9, No:2, 2015

Statistical Wavelet Features, PCA, and SVM Based Approach for EEG Signals Classification

R. K. Chaurasiya, N. D. Londhe, S. Ghosh

***Abstract*—**The study of the electrical signals produced by neural activities of human brain is called Electroencephalography. In this paper, we propose an automatic and efficient EEG signal classification approach. The proposed approach is used to classify the EEG signal into two classes: epileptic seizure or not. In the proposed approach, we start with extracting the features by applying Discrete Wavelet Transform (DWT) in order to decompose the EEG signals into sub-bands. These features, extracted from details and approximation coefficients of DWT sub-bands, are used as input to Principal Component Analysis (PCA). The classification is based on reducing the feature dimension using PCA and deriving the support-vectors using Support Vector Machine (SVM). The experimental are performed on real and standard dataset. A very high level of classification accuracy is obtained in the result of classification.

***Keywords*—**Discrete Wavelet Transform, Electroencephalogram, Pattern Recognition, Principal Component Analysis, Support Vector Machine.

Open Science Index, Computer and Information Engineering Vol:9, No:2, 2015 waset.org/Publication/10000595

I. Introduction

ELECTROENCEPHELOGRAPHY is the study of the electrical signals produced by brain. Production of electrical signals as a result of neural activity of the brain starts as early as from the 17th week of prenatal development. Electrical signals generated by the human brain represent the thinking of the mind and the status of the body. The close study of these Electro-EncephaloGram (EEG) signals is useful in many research areas such as detection and classification of event related potentials, seizure detection and prediction, brain–computer interfacing, Study of mental disorders like psychiatric disorders and dementia, and sleep signal analysis. For better understanding of human behavior, the EEG signal waves are further divided in five major sub-bands based on the frequency ranges. These bands from low to high frequencies respectively are called delta (δ)(Range 0.5-4Hz), theta (θ)(Range 4-8 Hz), alpha (α) (Range 8-13 Hz), beta (β)(Range 13-30 Hz), and gamma (γ)(Range 30-45 HZ)[1].

The visual distinction of seizure from common artifacts within an EEG measurement is based on the shape and spikiness of the waveforms. A signal with seizure have a rhythmical and prominent spiky, whereas the most of other artifacts are non-stationary and randomly shaped. But considering the fact that the recorded EEG pattern is a special mapping of signals captured by placement of electrode onto different regions of the scalp, it is extremely difficult for human being to observe and understand the actual behavior of the brain by merely visual inspection. Hence there is an ever increasing demand of easily accessible and fully automatic epileptic seizure detection system using EEG signals.

In this paper we propose a statistical feature based epileptic seizure detection system. Statistical features are extracted from Discrete Wavelet transforms (DWT) of EEG signals. Further, Support Vector Machine (SVM) is used for classification into two classes i.e. is epileptic and normal. In order to reduce the time and space complexity and to avoid redundancy in the observed features, we have applied Principal Component Analysis (PCA) on the normalized feature matrix.

R. K. Chaurasiya is with the Dept. of Electronics and Telecommunication, National Institute of Technology, Raipur, 492010 India (Phone: +91-91 65 971639; e-mail: rkchaurasiya@nitrr.ac.in).

N. D. Londhe is with the Dept. of Electrical Engineering, National Institute of Technology, Raipur, 492010 India (e-mail: nlondhe.ele@nitrr.ac.in).

S. Ghosh was with the NIT, Rourkela. He is now with the Dept. of Electrical Engineering, NIT, Raipur, India (e-mail: sghosh.ele@nitrr.ac.in).

II. Related Work

The electrical signals for brain activity were first recorded by the English scientist Richard Caton in 1875. Hans Berger started the study of EEGs from human brain in 1920 [2]. *Epilepsy* is a Greek word, which means 'to seize or attack'. The very basic concepts of epilepsy can be found in ancient Indian medicine (4500–1500BC) as *apasmara*, which means "loss of consciousness". Babylonian tablet in the British Museum in London also gives the detailed knowledge about the epileptic disease and its cure [1]. Kaufman associated the epileptic attacks with abnormal electrical discharges [3].

Most of the epilepsy analysis methods developed in the 20th century were based on the concept of visual inspection of EEG signals by highly skilled electroencephalographers. However, with the advancement in the field of signal processing and pattern recognition, different automatic techniques of epileptic seizure detection have been developed in last two decades [6], [9].

Spectral analysis based feature extraction method provides poor results for EEG classification as the frequency domain information is provided at the cost of time domain information such as the amplitude distribution and EEG pattern. Hence, both time and frequency domain based feature extraction algorithms such as Discrete Wavelet Transform (DWT) are being used in current research [4]-[6]. The other advantage of DWT over spectral analysis is its suitability for analysis of non-stationary signals like EEG [7], [8]. Kai Fu et al. have recently published their work with Hilbert-Huang Transformed (HHT) based approach [9].

Open Science Index, Computer and Information Engineering Vol:9, No:2, 2015 waset.org/Publication/10000595

III. Data-Set for Experimental Analysis

In last few years, most of the researchers have used publicly available data described in [10] for their research work in the field of epileptic seizure detection. We are also using the same *bench-mark* database in order to compare our results with the results of previous research works. The database is prepared by taking inputs from different subjects and is divided into five sets (A-E) each containing 100 EEG samples recorded through single channel. The mental status of the subjects in each data set (A-E) at the time of data recording was as follows:

A, B: Five healthy volunteers, relaxed in an awake state with eyes open (A) and closed (B).

C, D: Activities measured during seizure free intervals of EEGs from five patients, all of whom had achieved complete seizure control and were correctly diagnosed.

E: Contains seizure activity (recorded from the same patients as for set B and C).

The data set were recorded using a 128-channel amplifier system and standardized 10-20 electrode placement scheme. After recording, the data were sampled and digitized at 173.61 samples per second using 12 bit resolution. As the useful information from the data can be found only in δ, θ, α, β, and γ sub-bands, a band-pass filter with 0.50–40 Hz (12 dB/oct) was applied. In this study, we used the dataset A and E for classification, as only set E contains the samples from confirmed epilepsy (Class I), the data set A consists of sample from persons having no epilepsy (Class II).

IV. Proposed Algorithm

The stepwise details of the proposed algorithm are given in Table I. As described in Table I, we pick one EEG sample at a time and find its DWT coefficients. The features from the DWT coefficients are extracted and appended in the corresponding column of a feature matrix. Note that for a set of 100 EEG samples, the feature matrix will have number of column = 100, and number of rows = the total number of extracted feature. The same procedure is repeated for all the EEG samples. The final feature matrix is normalized and passed for dimension reduction using PCA. Binary SVM classification is performed on the dimension reduced feature matrix for classification.

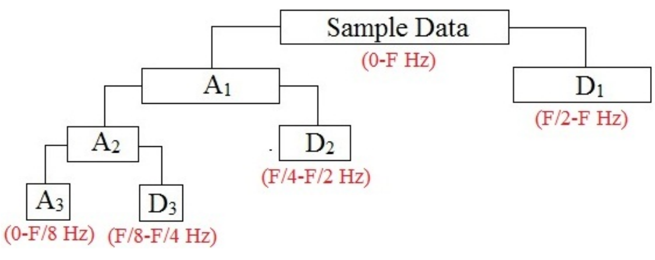


Fig. 1 3-Level wavelet decomposition of the sample data signal having 0-F Hz frequency range. The signal is decomposed into detail coefficients D_1-D_3 and approximation A_3. The frequency range covered in different decompositions and approximation is shown in the bracket

TABLE I
Stepwise Details of the Proposed Algorithm

1.	i = 1
2.	**for** i< = size of the data set, **do**
3.	Decomposition the i^{th} EEG sample using5-level DWT.
4.	Extract the statistical wavelet features from DWT coefficients, and put in i^{th} column of feature matrix *Ftr_Mat*.
5.	**end for.**
6.	Normalization the *Ftr_Mat* (feature wise).
7.	PCA on *Ftr_Mat* for dimension reduction.
8.	Train the SVM and derive the support vectors.
9.	Apply SVM on test data for Classification.
10.	Measure the accuracy obtained by SVM classification.

A. Feature Extraction Using DWT

Fourier transform and other spectral analysis techniques are the popular tools used for analyzing stationary signals. However, for non-stationary signals like EEG, direct application of Fourier transform is not recommended. Hence, time-frequency analysis using wavelet transform have been adapted in the proposed work.

A multi-level wavelet decomposition of the EEG samples provides the information at different resolutions of the samples at different frequency bands [11]. Fig. 1 shows 3-level wave decomposition.

The selection of the level of decomposition and the type of the basic wavelet is a problem specific criterion. For extracting the features from EEG samples, the frequency range of interest is 0-50 Hz. So the level of decomposition is chosen to be 5. Different kinds of wavelets were tried and the accuracy of the SVM classification was measured. It was observed that the Daubechies wavelet suits the EEG signals more and hence it was chosen as filter wavelet. Fig. 2 shows an EEG sample signal from set A, its decomposition D_1-D_5 and approximation A_5. Considering the frequency of our interest, decomposition D_3-D_5 and approximation A_5 are chosen for feature extraction.

Extracted wavelet coefficients provides both time and frequency representation of the EEG samples. Various statistical features are extracted from these coefficients as mentioned below:

(1) Feature 1 to 4 consists of the mean of the absolute values of the approximation (A_5) and details (D_3-D_5).

$$[Ftr(1), Ftr(2), Ftr(3), Ftr(4)] = [mean(abs(A_5)), mean(abs(D_5)), mean(abs(D_4)), mean(abs(D_3))],$$

Here Ftr is the feature vector for one EEG sample.

(2) Feature 5 to 8 consists of the average of the square of the second order norm (equivalent to average power of discrete signals) of the approximation and details.

(3) Feature 9 to 12 consists of the median of the actual values of the approximation and details.

(4) Feature 13 to 16 consists of the standard deviation of the coefficients of the approximation and details.

(5) Feature 17 to 20 consists of the kurtosis; feature 21 to 24 consists of the skewness; and feature 25 to 28 consists of the entropy of the coefficients of the approximation and details.

Open Science Index, Computer and Information Engineering Vol:9, No:2, 2015 waset.org/Publication/10000595

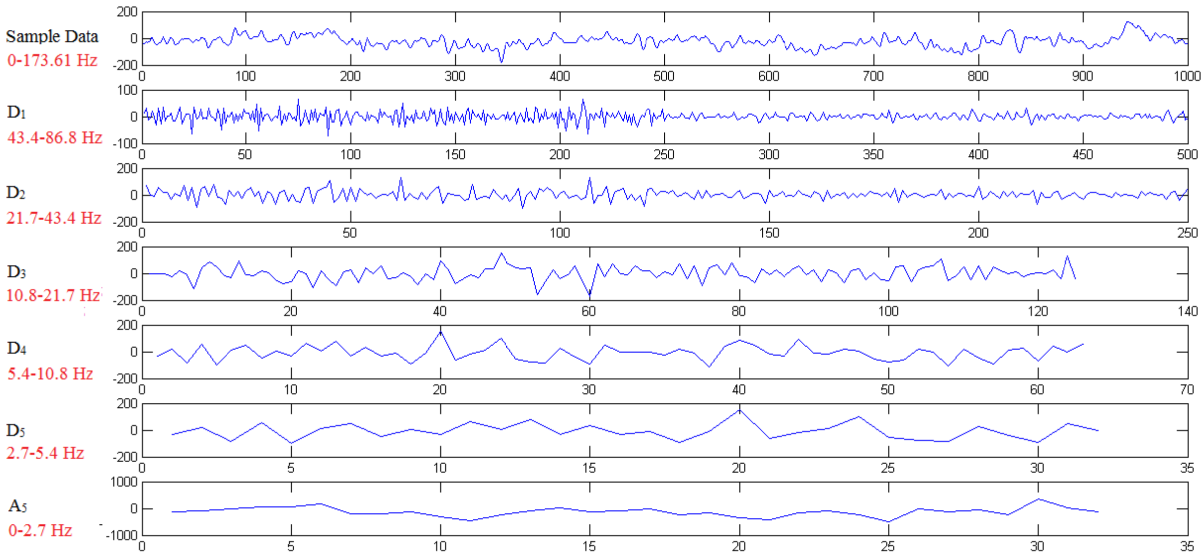


Fig. 2 5-Level wavelet decomposition of sample data (from set A) signal of 0-173.61 Hz. D_1-D_5 are details and A_5 is approximation. (For clear visibility only 1000 initial samples taken from 4097 samples of the sample data and axis are *not* equalon the sub-plots)

(6) Feature 29 to 31 consists of the ratio of the absolute mean values of adjacent sub-bands i.e. approximation and detail. The first three feature sets (1-3) represent the frequency distribution of the signal and the other three (4-6) represents the variation with respect to frequency distribution. It is clear that feature vector for one EEG sample consists of 31 features. As each data set (A to E) consists of 100 EEG samples, the feature matrix for each data set is of the dimension 31×100.

B. Principal Component Analysis (PCA):

PCA is well-established and the most widely used method for dimension reduction. PCA allows to represent a *d*-dimensional data into a lower dimensional space (say *l*, where *l*<*d*). The PCA reduced data set is the *best* representation of the *d*-dimensional data into *l*-dimensional space (best in terms of minimum squared-error-distance).

In the process of minimizing the squared-error distance between the actual data and reduced data, one can derive the method to reduce the *d*-dimension data into *l*-dimensional data through PCA. First the *d*-dimensional mean μ and $d \times d$ dimensional covariance matrix S are computed for original *d*-dimensional data set. Next, *d* eigen values are calculated and are sorted in decreasing order. Say these eigen values (in decreasing order) are $\lambda_1, \lambda_2 \ldots \lambda_d$ and the corresponding eigen vectors are $e_1, e_2 \ldots\ldots e_d$. (all the eigen vectors $e_1, e_2 \ldots\ldots e_d$ are mutually orthogonal). Subsequently, the first *l* eigen vectors $e_1, e_2 \ldots\ldots e_l$, which correspond to largest *l* eigen values $\lambda_1, \lambda_2 \ldots \lambda_l$ are chosen as natural basis for projecting the *d*-dimensional data in *l*-dimensional space. A good value of *l* is decided by the fact that there is a significant comparative difference between l^{th} and $(l+1)^{th}$ eigen value. The more details and mathematical analysis of PCA can be found in [12] [13].

C. Support Vector Machine

The idea of SVM is originated from the idea of controlling the *generalizing capabilities* of machines for automation. The performance of a classifier must be generalized, i.e. it should perform well when it is applied for the data outside the training set. The notion of maximizing the *margin* between the *support vectors* is at the heart of the SVM classifier, in order to perform more accurately on unknown data [13], [14]. Consider the hyper-plane in (1):

$$w^T x + w_0 = 0 \qquad (1)$$

The *margin* is the Euclidian distance $1/_{||w||}$ between the two parallel hyper-planes (support vectors) described in (2):

$$w^T x + w_0 = 1, \text{ and } w^T x + w_0 = -1 \qquad (2)$$

Let x_i are training points, with respective classes $y_i \in \{-1,1\}$, i=1,2,….,N for a 2-class classification problem. The task is to optimize for minimum training error and maximum separating margin between hyper-planes of (2). SVM classifier solves this task by solving the optimization problem of (3):

$$\text{Minimize } L(w, w_0, \xi) = \frac{1}{2}||w||^2 + C\sum_{i=1}^{N}\xi_i \qquad (3)$$

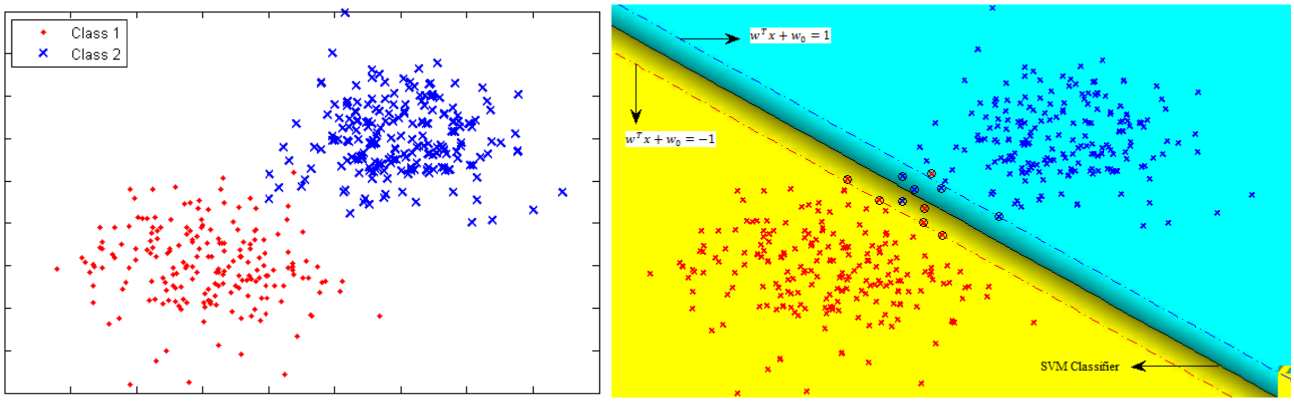
$$\text{Subjected to} \quad \begin{array}{l} w^T x + w_0 \geq 1 - \xi_i, \quad if\ y_i \in 1 \\ w^T x + w_0 \leq -1 + \xi_i, \ if\ y_i \in -1 \end{array}$$

and

$$\xi_i \geq 0 \qquad (4)$$

For the present work, involving two class classification (Epileptic seizure or not), we first learn the classifier equation (similar to (1)) by solving the optimization problem of (3) with constrains of (4), using half of the feature vectors from *Ftr_Mat* as training data. Then the separating hyper-plane is used to classify the remaining feature vectors of *Ftr_Mat*. The pictorial representation of learning the SVM classifier from the training data of different classes is shown in Fig. 3.

Open Science Index, Computer and Information Engineering Vol:9, No:2, 2015 waset.org/Publication/10000595



(a) (b)

Fig. 3 (a) Distribution of 2-dimensional data set of two different classes (b) Support vectors (dotted lines) and SVM classifier (solid line) learnt to optimize for minimum training error and maximum separating margin between hyper-planes

V. Experimental Results

In the presented work, we have used pattern recognition approach for EEG signal classification. EEG samples were decomposed into sub-band using 5-level DWT with Daubechies wavelet. Thirty-one different Statistical features were extracted from the details and approximation sub-bands. This feature extraction process was repeated for all the sample signals of set A and set E. A 31×200 sized *Ftr_Mat* was formed after adding all the feature vectors to it. Here each row corresponds to a particular feature and each column corresponds to a particular EEG sample. The rows of *Ftr_Mat* were then normalized between 0 and 1 using

$$x_i = \frac{x_i - x_{min}}{x_{max} - x_{min}}, for\ i = 1\ to\ 200 \tag{5}$$

where x_i = feature value in the i^{th} particular row, x_{min}= minimum value in that row, and x_{max}= maximum value in that row.

The dimension of the extracted features was reduced to 7 using PCA. After dimension reduction we had a 7×200 *Ftr_Mat*. 100 out of 200 feature points (50 from each class) were used to train the SVM classifier. The remaining feature points were used to test the accuracy of the SVM classifier.

In order to measure the performance of the classifier, sensitivity (True Positive Ratio- TPR) and specificity (True Negative Ratio- TNR) were calculated by using confusion matrix. Equations (6) and (7) describe the formula used for calculating sensitivity and specificity using True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

$$Sensetivity = TPR = \frac{TP}{TP+FN} \times 100\% \tag{6}$$

$$Specificity = TNR = \frac{TN}{TN+FP} \times 100\% \tag{7}$$

TABLE II
THE VALUES OBTAINED FOR PERFORMANCE MEASURING PARAMETERS

Parameter	Numerical Value Obtained
Sensitivity	100.0%
Specificity	99.50%
Accuracy	99.75%

The accuracy is calculated using:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \tag{8}$$

Table II summarizes the values of accuracy, sensitivity and specificity obtained after classification.

VI. Conclusion

In this paper, a pattern recognition approach is proposed to detect the epileptic seizure. The proposed approach is based on extracting the statistical features such as means, power, standard deviation, kurtosis, skewness, entropy, and median from decomposed sub-bands of DWT. In order to avoid the redundancy in the observed features, the normalized values of the extracted wavelet features are passed to PCA for dimension reduction. Reducing the feature dimension using PCA also helps in reducing the time and space complexity. SVM classification is applied to classify the data into one of the two classes: epileptic seizure or not.

A very high level of detection ratio and accuracy is obtained after applying the proposed classification approach on real EEG data set. The higher level of accuracy obtained, makes the system a perfect helping tool for automated classification of EEG signals.

Development of a dedicated hardware set-up and user friendly interface can be considered as the future work for the presented work.

References

[1] Saeid Sanei and J.A. Chambers. *EEG Signal Processing*. John Wiley & Sons, 2007.

[2] A. Massimo, "In Memoriam Pierre Gloor (1923–2003): an appreciation". Epilepsia, vol.-45(7), July 2004, page-882.

[3] M. A. B. Brazier. "A History of the Electrical Activity of the Brain". The First Half-Century, Macmillan, New York, 1961.

[4] M. D. Alessandro, R. Esteller, G. Vachtsevanos, A. Hinson, A. Echauz, and B.Litt. "Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: A report of four patients". IEEE Transactions on Biomedical Engineering-2003. vol.-50 (5), pp.-603–615.

[5] B. P. Marchant. "Time–frequency analysis for biosystem engineering". Biosystems Engineering-2003. vol.-85 (3), pp.-261–281.

[6] A. Subasi. "EEG signal classification using wavelet feature extraction and a mixture of expert model". Expert Systems with Applications-2007, vol-32, pp.-1084–1093.
[7] A. Subasi, and M. Ismail Gursoy. "EEG signal classification using PCA, ICA, LDA and support vector machines". Expert Systems with Applications-2010. vol. 37,pp.-8659–8666.
[8] A. Subasi. "Epileptic seizure detection using dynamic wavelet network". Expert Systems with Applications-2005. vol.-29, pp.-343–355.
[9] K. Fu, J. Qu, Y. Chai, and Y. Dong, "Classification of seizure based on the time-frequency image of EEG signals using HHT and SVM". Biomedical Signal Processing and Control-2014, vol.-13, pp.-15–22.
[10] R.G. Andrzejak, K. Lehnertz, F. Mormann, et al. "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state". Phys. Rev. Ed- 64 (6)-061907.2001.
[11] H. Adeli, Z. Zhou, and N. Dadmehr. "Analysis of EEG records in an epileptic patient using wavelet transform". 2003. Journal of Neuroscience Methods, vol.-123, pp.-69–87.
[12] R. O. Duda, P. E. Hart, and D.G. Stork. *Pattern Classification.*2nd Ed., John Wiley & Sons, 2001.
[13] S. Theodoridis, and K. Koutroumbas. *Pattern Recognition*. 4th Ed., Elsevier - Academic Press, 2009.
[14] P. S. Sastry. "An introduction to Support Vector Machines". Chapter in J.C. Misra (Ed), computing and information sciences: Recent Trends. Narosa Publishing House, New Delhi 2003.

Rahul K. Chaurasiya born in a small village in the state of Madhya-Pradesh, India, received the undergraduate B. Tech. degree from the Maulana Azad National Institute of Technology, Bhopal in 2009. He then joined the Indian Institute of Science, Bangalore, as a graduate student. There, he received M.E. degree in System Science and Automation in 2011.

After receiving his M.E. degree he joined Brocade communications as senior software engineer in Bangalore. In 2013, he moved to Raipur, India and joined National Institute of Technology, Raipur as Assistant Professor. Currently, he is also working towards Ph.D. degree from the same institute. His current research area includes pattern recognition, signal processing and brain-computer interfacing.

Open Science Index, Computer and Information Engineering Vol:9, No:2, 2015 waset.org/Publication/10000595