International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 3 (2015) pp. 6405-6417

http://www.ripublication.com

Performance Analysis of SVD, ICA and Hidden Markov Model in Classification of Epilepsy Risk Level from EEG Signal

Dr.R. Hari Kumar[1] and Dr.M.Balasubramani[2]

[1]Professor, Department of Electronics and Communication Engineering
Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu State, India.
[2]IAssistatnt Professor, Department of Electronics and Communication Engineering
Info Institute of Engineering, Kovilpalayam, Coimbatore 641107, Tamilnadu State, India.
Email: harikumarrajaguru@gmail.com / rishiveda@yahoo.com

Abstract

Various methods for dimensionality reduction and classifier combination have been designed to classify the risk levels of epilepsy. Providing large number of features in training data can affect the classification performance of machine learning algorithms. Here the standard 16 channel bipolar EEG signal recordings data have been used in order to diagnose epilepsy by classifying it into normal and abnormal subjects. The advantage of the proposed approach is that it reduces the size of feature space by way of using feature extraction methods. The decisions obtained from these methods have been coalesced to form a fused data. The idea behind this work is to discover a reduced feature space so that a classifier built using this tiny data set would perform no worse than a classifier built from the original data set. This paper investigates the possibility of a hybrid classifier for epilepsy risk level classification from EEG signals. The dimensionality reduction of EEG signal space is done by Singular Value Decomposition (SVD) and Independent Component Analysis (ICA).The final risk level classification is performed by Hidden Markov Model (HMM). A group of twenty patients with known epilepsy findings are analyzed. Here, the classification performance of Hidden Markov Model has been evaluated using measures such as Mean Square Error (MSE) and Kappa function.

Keywords: Epilepsy, Epileptic Seizure, Singular Value Decomposition (SVD), Independent Component Analysis (ICA), Hidden Markov Model (HMM), Kappa Function, Mean Square Error (MSE).

1. Introduction

Knowledge of neuronal functions and neuro physiological properties of the brain together with the mechanisms underlying the generation of signals and their recordings is vital for analysing these signals for detection, diagnosis, and treatment of brain disorders and the related diseases. Epilepsy is a neurological disorder with preponderance of about 1-2% of the world's population (Harikumar et.al). It is due to excessive synchronization of cortical neuronal networks and is characterized by sudden recurrent and transient disturbances of perception or behaviour (Pellouchoud et.al). It is a condition in which an individual experiences abnormal bursts of electrical discharges. Epilepsy is marked by the term "epileptic seizures" (Gotman et.al). Epileptic seizures are classified as partial or focal, generalized, unilateral and unclassified seizures. Focal epileptic seizures involve only part of cerebral hemisphere and in corresponding parts of the body. Generalized epileptic seizures involve the entire brain and produce bilateral motor symptoms usually with loss of consciousness. Both types of epileptic seizures can occur at all ages (Gotman et.al). One of the most important tools for diagnosing Epilepsy includes monitoring the brain activity through the electroencephalogram (EEG). An EEG signal is a measurement of currents that flow during synaptic excitations of the dendrites of many pyramidal neurons in the cerebral cortex. When brain cells (neurons) are activated, the synaptic currents are produced within the dendrites. The EEG signature is a transient waveform of isolated spikes, spike trains, sharp waves or spike-wave complexes. It also assists in classifying the underlying epileptic syndrome. Various methods for dimensionality reduction and classifier combination have been designed to classify the risk levels of epilepsy. However use of large number of features in training data can affect the classification performance of machine learning algorithms (Ocal et.al). The objective of this paper is to present a novel risk level classifier for Epileptic EEG from different set of patients. Here the standard 16 channel bipolar EEG signal recordings data have been used in order to diagnose epilepsy. The paper is organized as follows, section 1 introduces the physiology Epilepsy detection and the related works are discussed in section 2. Dimensionality reduction method like SVD, ICA and HMM classifier are discussed in section 3. Results are discussed in section 4 and paper is concluded in section 5. This paper investigates the possibility of a hybrid classifier for epilepsy risk level classification from EEG signals.

2. Related work

Since 1970, seizure predictability based on scalp-EEG and depth-EEG signals has been studied. In 1982 Gotman presented a computerized system for detecting a variety of seizures (Gotman et.al).An increase of interest in seizure detection was motivated by advanced linear and non-linear approaches in the 1990s. A variety of techniques such as self-organizing map neural networks (Gabor et.al), genetic programming (Marchesi et.al), seizure intensity, and fractal dimension (Esteller et.al) were utilized on seizure detection. More recent methods on seizure detection include autoregressive modeling, rule-based analysis of amplitude and frequency changes of scalp EEG (Ocal et.al), multi-dimensional probability evolution (McSharry et.al), singular-spectrum techniques (Celka et.al) for scalp EEG. In 1995 (Shim-Yih Tseng *et al*) worked on autoregressive (AR) and autoregressive-moving average (ARMA) modeling for EEG signal analysis. They concluded that the AR model is preferred for estimating EEG signal. In 1997, (Qu and Gotman et.al) proposed the use of the nearest-neighbour classifier on EEG features extracted in both the time and frequency domains to detect the onset of the epileptic seizures. In 1998 (Penny et al). discovered that HMMs are capable of detecting non-stationary changes and are suitable for EEG analysis. However they concluded that operating HMMs on AR coefficients is flawed and the state and state transitions in HMM model are estimated incorrectly due to the windowing procedure used in AR models. In 2005 Phillips *et al.* described that spatio-temporal data reduction and soft computing techniques like the Bayesian statistics can also be used in order to reduce the computational loads (Gabor et.al).Works on EEG classification usually apply HMMs to the time changing feature vectors extracted by an AR model or by some other digital signal processing techniques. (Huang et.al) use the mean frequency features, calculated from FFT spectrum, for detecting the arousal state changes. But application of FFT for signals that has high probabilistic variation may not produce good results. Obermaier, Guger, and Pfurtscheller (Obermaier et.al) compare LDA (linear discrimination analysis) and HMMs on band pass-filtered feature vectors and experiment with the structure parameters of HMMs**.** Later Graja and Boucher have investigated the use of hidden Markov tree models for segmenting ECG signals encoded with the discrete wavelet transform (Graja et.al). More recently (Easwarmoorthy *et al*).focused on the method which indicated the state of illness of epileptic patient from EEG recording. Analysis based wavelet decomposition through DWT was performed. Though DWT is good tool for signal analysis it does not produce effective results as that of HMM.

In this paper, Epileptic EEG is acquired and dimensionality reduction is performed by Singular Value Decomposition (SVD) and Independent Component Analysis (ICA). A further reduction in the input is achieved by Vector Quantization. The final risk level classification is done by Hidden Markov Model (HMM).

3. Proposed Dimensionality reduction techniques and HMM classifier

Epileptic EEG signal for the purpose of analysis was acquired from twenty epileptic patients from neurology department of Sri Ramakrishna Hospital, Coimbatore, India (Harikumar et.al). The signal is acquired using a non-invasive method of electrode placing where sixteen channel EEG signals is acquired. The continuous EEG signal is divided into a signal of smaller duration, with duration of two seconds (Clin et.al). The international electrode system introduces artifacts such as muscle noise, eye movements, cardiac signals and line noise due to the low conductivity of skull and synchronization of electrical activity (Pellouchoud et.al) With an EEG signal free of artifacts, a reasonably accurate detection of epilepsy is possible. An artifact free EEG is obtained with the help of neurologist. The Epileptic EEG is diagnosed by acquiring the Epileptic EEG, extracting and processing the SVD and ICA components. The various steps involved in processing the features are elaborated.

Twenty set of patients are considered in this study. EEG signal from each patient is divided into four epochs with duration of two seconds each and sampled at frequency of 200Hz. The dimensionality reduction is done by Singular Value Decomposition (SVD) and Independent Component Analysis (ICA) (Ganesh et.al). Cluster heads were defined using Vector Quantization. The final risk level is obtained the Hidden Markov Model (HMM). Performance of HMM is evaluated by Mean Square Error (MSE), Kappa Function and extended Kappa Function. The system involves two phases namely the Training and the Testing phase. The system is trained using 20 set of patients. Similarly the testing of the system is done and the performance is evaluated. The entire block diagram of hybrid epilepsy risk level classification is shown in Figure.1

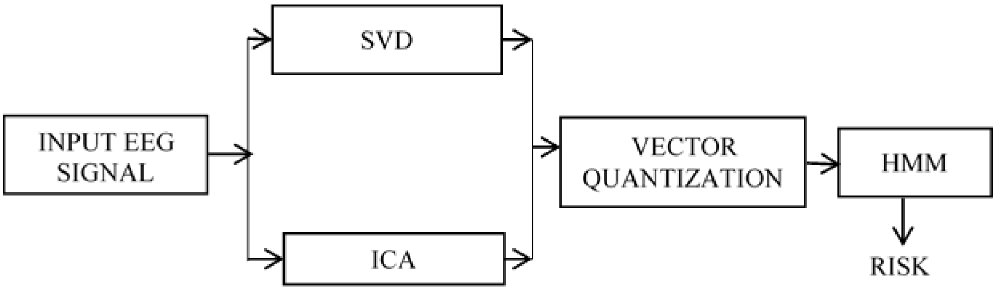

Figure 1. Block Diagram of Hybrid Epilepsy Risk Level Classification System

Each channel has 400 samples of EEG signals per epoch and four such epochs of data forms a bin. There are sixteen such, a bins are available per patient. The data volume for a patient is around 25,600 samples. Hence this large amount of data necessitate the dimensionality reduction technique level for processing the EEG signal. This number is enough to give reliable statistics in mean and variance calculation.

3.1. Dimensionality Reduction using Singular Value Decomposition (SVD)

Dimensionality reduction is the process of integrating multiple sources of information such that their combination yields better results than if the data sources are used individually. Feature extraction aims at finding the smallest feature set having the most beneficial impact on machine learning. SVD is widely used in multi rate signal processing where the initial dataset is separated into complementary subspaces called signal and noise subspaces in order to enhance the signal to noise ratio (Stewart et.al). The Singular Value Decomposition (SVD) is a useful tool that is used to separate the signal and noise. Consider a real M x N matrix X of observations which may be decomposed as follows:

$$X = USV^T \quad \textbf{(1)}$$

Where S is an M x N non-square matrix with zero entries everywhere, except on the diagonal with elements s_i arranged in descending order of the magnitude. Each s_i is equal to $\sqrt{\lambda_i}$, the square root of eigen values of $C=X^TX$. The smallest eigenvalues are considered due to noise (Stewart et.al), (Abdi et.al). The columns of V are eigen vectors of C. The M x M matrix U is the matrix of projections of X onto eigenvectors of C. A routine for performing SVD is as follows:

Step 1. Find the N non-zero eigenvalues, λ_i, of matrix $C=X^TX$ and form a non-square diagonal matrix S by replacing the square roots $s_i= \sqrt{\lambda_i}$, of the N eigenvalues in descending order of magnitude on leading diagonal and setting all other elements of S to Zero.

Step 2. Find the orthogonal eigenvectors of the matrix X^TX corresponding to the obtained eigen values, and arrange them in the same order. This ordered collection of column vectors, forms the matrix V.

Step 3. Find the first N column vectors of matrix U: $u_i = s_i^{-1}$.

Step 4. Add the rest of M-N vectors to the matrix using the Gram Schmidt Orthogonalization process.

3.2. Dimensionality Reduction using Independent Component Analysis (ICA)

ICA describes a model for multivariate data describing large database of samples. The variables in the model are assumed non-Gaussian and mutually independent and they are called the independent components of the observed data. These are also called sources or factors (McSharry and Clifford).The components extracted in ICA are orthogonal in nature and hence the redundant data sets are eliminated (Penny et.al), (Lee et.al). Independent Component Analysis is generally applied where multi-channel data extraction is preferred (Xu et.al), (Oja). We assume that we observe *n* linear mixtures of $x_{1,\dots,}x_n$ of independent components as in (2):

$$X_j=a_{j1}s_1+a_{j1}s_2+\dots\dots+a_{jn}s_n,\ j=1,n \qquad \textbf{(2)}$$

In this equation the variable x_j and the independent component s_i are random variables, $x_j(t)$and $s_i(t)$ are samples of random variables. The variable and the independent component are also assumed to have zero mean reducing the problem to the model zero-mean (Choudery et.al), denoted by (3):

$$X=X_i-E(X) \qquad \textbf{(3)}$$

Let *x* and *s be* the random vector whose elements are x1,….,x_n and $s_1,\dots..s_n$ respectively. Let **A** be the matrix containing the elements a_{ij} which is expressed as in (4).

$$x=As \quad or \quad x=\textstyle\sum_{i=1}^{n} a_i s_i \qquad \textbf{(4)}$$

The above equation is called independent component analysis or ICA where only the measures variable *x* is available and the objective is to determine both the matrix **A** and the independent components. The model is assumed to have independent and non-Gaussian components (McSharry et.al), (Penny et.al).

3.3 Vector Quantization

A vector quantizer is based on the principle of Block Coding and maps k-dimensional vectors in the vector space R^k into a finite set of vectors $Y = \{y_i: i = 1, 2, ..., N\}$. Each vector in Y *is* called a code vector or a code word .The codebook comprises of set of code words (Penny et.al). Voronoi region is the region that is associated with each code word, and it is defined by (9).

$$Vi = \{x\epsilon R^K: \left\| |x - y_i| \right\| \leq \left\| |x - | \right\|, for\ all\ j \neq i\} \qquad \textbf{(5)}$$

Step 1. The following steps are involved in construction of a code book.
Step 2. The number of code words N is determined. The initial code book is the N code words chosen randomly.
Step 3. The vectors around the code words are clustered using the Euclidean distance. This is done by taking each input vector and finding the Euclidean distance between it and each code word. The input vectors that are at a minimum distance from the code word are clustered.
Step 4. The new set of code words is determined by averaging the cluster as shown in (10).

$$y_i = \frac{1}{m}\sum_{j=1}^{m} x_{ij} \qquad \textbf{(6)}$$

Where i is the component of each vector (x, y, z directions), m is the number of vectors in the cluster. Repeat steps 2 and 3 until the either the code words don't change or the change in the code words is small. A vector quantizer is comprises of Encoder and Decoder. The encoder outputs the index of the code word that offers the lowest distortion taking the input vector (Makhoul et.al),(Jeng-shyang pan et.al)]. Once the closest code word is found, the index of that code word is sent through a channel When the encoder receives the index of the code word, it replaces the index with the associated code word (Kekre et.al), (Mario et.al). Figure 2. Shows a block diagram of the encoder and decoder.

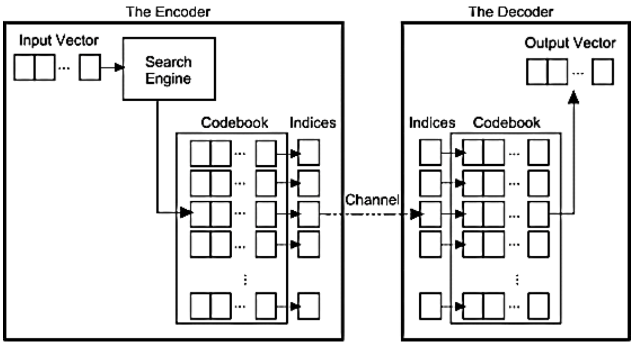

Figure 2. The Encoder and Decoder in a Vector Quantizer.

3.4 Base classifier used (Hidden Markov Model –HMM)

A hidden Markov model (HMM) is a statistical Markov model especially designed for system with probabilistic variations. The state is directly visible to the observer in simpler markov model, and therefore the state transition probabilities are the only parameters (Choudrey et.al). In a Hidden markov model, the state is not directly visible, but output is dependent on the state that is visible. A probability distribution is established for output tokens (Penny et.al) (Rabiner).Therefore the information about the sequence of states is generated by the sequence of tokens. A hidden Markov model (HMM) is a triple (π, A, B).

Where,

$\pi = \pi_i$ the vector of the initial state probabilities

$A=(a_{ij})$ *the sate transition matrix,* $p_r(x_i|x_{j\ t-1})$

$B = (b_{ij})$ *the emission matrix,* $p_r(y_i|x_j)$ **(7)**

Each probability in the state transition matrix and in the emission matrix is time independent that is, the matrices do not change in time as the system evolves. In practice, this is one of the most unrealistic assumptions of Markov models about real processes.

3.5 Baum-Welch (EM) algorithm

The Baum-Welch (EM) algorithm is a general method of finding the maximum likelihood (ML) estimate of the parameters of an underlying distribution. On each iteration of the EM (Expectation - Maximation) algorithm there are two steps-called the expectation step or the E-step and the maximization step or the M-step. There are two main applications of the EM algorithm. The first one occurs when there is problem or limitation of observation process and hence the data has missing values. The second one is done by assuming the existence of additional but missing values or (or hidden) parameters for solving the likelihood function. Now defining the Maximum likelihood, we have a density function that is governed by the set of parameters. We also have a data set of size N, drawn from this distribution, i.e., $X=\{X_1, X_2, \ldots\ldots X_N\}$. That is, we assume that these data vectors are Independent and Identically Distributed (IID) with distribution p. Therefore, the resulting density for the samples is

$$p(\chi|\theta) = \prod_{i=1}^{N} p\,(x_i|\theta) = p(\theta|\chi) \qquad \textbf{(8)}$$

The function $\mathcal{L}(\theta|\chi)$is called the likelihood function. The likelihood is thought of as a function of the parameters θ where the data χ is fixed. In the maximum likelihood our goal is to find the θ that maximizes $\mathcal{L}$. Therefore, we wish to find $\theta^{*}$where

$$\theta^{*}= ar|gm|ax(\theta|\chi) \qquad \textbf{(9)}$$

Here as it is analytically easier we maximize $\mathcal{L}(\theta|\chi)$. Now we will describe the two basic steps of EM methods. As before, we assume that data χ is observed and is generated by some distribution. We call χ as incomplete data. We assume that a complete data set exists$\mathbb{Z}$ = (X,Y) and also assume a joint density function p$(\mathbb{Z}|\theta)$ = p$(x,y|\theta)$ = p$(y|x,\theta)p(x|\theta)$.The new likelihood function is defined by this density function, $\mathcal{L}(\theta|\mathbb{Z})$ = p$(X,Y|\theta)$ called the complete-data likelihood. Here the missing information Y is unknown, random, and presumably governed by an underlying distribution, hence this function can be called a random variable. The original likelihood $\mathcal{L}(\theta|\chi)$is referred to as the incomplete-data likelihood function. Given the observed data X and the current parameter the EM algorithm first finds the expected value of the complete-data log-likelihood $log\ p\mathcal{L}(\theta|\chi)$with respect to the unknown data Y. That is, we define the following $\tilde{Q}$ function.

$$\theta^{(i)} = argmax\tilde{Q}\left(\theta,\theta^{(i-1)}\right) \qquad \textbf{(10)}$$

These two steps are repeated. Each iteration is guaranteed to increase the log likelihood and the algorithm is guaranteed to converge to a local maximum of the likelihood Function. Now this EM algorithm is applied to HMM case, where the EM algorithm is usually known as the Baum-Welch algorithm that established convergence properties for this algorithm. We consider O = $\{O_1,O_2 \ldots O_T\}$ to be the observed data and the underlying state sequence Q = $\{q_1, \ldots q_T\}$ to be hidden or unobserved. The incomplete-data likelihood function is given by P(O|λ) whereas the complete-data likelihood function is P(O,Q| λ). The function $\tilde{Q}$ therefore is

$$\tilde{Q}(\lambda,\lambda')\sum_{q\in Q} P(q|O,\lambda)logP(q|\lambda',O) \qquad \textbf{(11)}$$

Where λ' are our initial (or guesses, previous) estimates of the parameters and where Q is the space of all state sequences of length T.

4. Results and Discussion

Performance of Hidden Markov model classifier have been evaluated using various measures. Kappa coefficient is one of the performance measures for classifier. Given an n x n contingency table of desired versus actual results, a typical performance measure is ratio of individuals lying in the associated region with corresponding classes to the total number of individuals as (Rabiner):

$$p_o = \frac{\sum_i n_{ii}}{N} \qquad (i=1,\ldots,k) \qquad \textbf{(12)}$$

N in total number of vectors in the test set, k is the number of classes, n_{ii} are the diagonal elements of confusion matrix (Lu et.al). However p_o does not take into account the agreement that might be due to chance shown in (Lee et.al).

$$p_c = \frac{\Sigma_i(\Sigma_j n_{ij} \Sigma_j n_{ji})}{N^2} (i, j = 1, \ldots, k) \tag{13}$$

Where

$$k = \frac{p_o - p_c}{1 - p_c} \tag{14}$$

Hence, the *k* coefficient is a corrected measure of agreement between the desired and actual group assignments (Cohen), (Everitt). If the agreement is due to strictly to chance then κ=0. If the agreement is greater than chance then κ>0. There is complete agreement when κ=1. (κ<0 if the agreement is less than chance. The Mean Square Error is calculated by subtracting the Observed Value from original Target Value as shown in (Penny and Robert).

$$\text{MSE}=(T_j - O_j)^2 (i, j = 1, \ldots, k) \tag{15}$$

Table.1 Shows the overall performance of MSE in the HMM classifiers for twenty patients. Average training error is 0.000014 and average testing error is 0.003059. Kappa Function is also calculated as a parameter for estimation. Table.2 shows the performance of Kappa Function in HMM classifiers for twenty patients. It can be inferred that the performance of Kappa Function is better than the MSE. The Computation time for all the dimensionality reduction techniques and the final risk level classifier is shown in Table.3. The computation time for ICA is low comparing SVD.

Table.1 Overall Performance of Mean Square Error in HMM Classifier

Target Values	SVD		ICA	
	Observed Value	Mean Square Error	Observed Value	Mean Square Error
0.85	0.79	0.0036	0.78	0.0049
0.85	0.81	0.0016	0.8	0.0025
0.65	0.6	0.0025	0.59	0.0036
0.45	0.39	0.0036	0.4	0.0025
0.25	0.19	0.0036	0.18	0.0049
0.35	0.32	0.29	0.3	0.0025
0.65	0.61	0.6	0.58	0.0049
0.25	0.2	0.2	0.22	0.0009
0.05	0.02	0.025	0.026	0.0005
0.1	0.07	0.06	0.07	0.0009
0.85	0.78	0.8	0.79	0.0036
0.35	0.28	0.28	0.29	0.0036

0.45	0.41	0.4	0.41	0.0016
0.65	0.61	0.59	0.59	0.0036
0.2	0.14	0.14	0.17	0.0009
0.15	0.1	0.11	0.11	0.0016
0.2	0.13	0.14	0.12	0.0064
0.65	0.57	0.58	0.55	0.01
0.25	0.21	0.19	0.21	0.0016

Table.3 Performance of Kappa Function in HMM Classifier

Kappa Function (SVD)	Kappa Function (ICA)
0.0023	0.0027
0.0010	0.0010
0.0014	0.0018
0.0021	0.0011
0.0026	0.0025
0.0001	0.0025
0.0011	0.0023
0.0012	0.0002
0.0002	0.0003
0.0031	0.0004
0.0030	0.0019
0.0006	0.0016
0.0008	0.0006
0.0008	0.0020
0.0005	0.0005
0.0020	0.0004
0.0018	0.0031
0.0021	0.0002
0.0019	0.0003
0.0038	0.0049

Table.3 Computation Time of SVD, ICA and HMM classifier.

MACHINE (16 Channel, 3 Epochs) (Sec)	MODULE (Sec)	HMM (Sec)	Total Time (Sec)
20-Patients	2 (SVD)	53.128	75.128
20-Patients	0.014988 (ICA)	51.526	70.487

5. Conclusion

In this paper, we have proposed a HMM model for classification of epileptic EEG.HMM outperforms other methods for risk level classification because it is inherited large transitions and probabilistic variation. The system is trained using

Independent Component Analysis (ICA). Finally the system is tested by a separate set of data matrix obtained from individuals who were not in the training phase. The performance evaluation of the system is analysed using Mean Square Error and Kappa function. The system is better evaluated using the Kappa Function ICA based dimensionality reduction with HMM classifier is quiet suitable for classification of epilepsy risk level from EEG Signal.

References

1. Harikumar.R, GaneshBabu.C & Vijayakumar.T(2012). Performance Analysis of Elman Neural Networks as Post Classifiers for Wavelet Transforms Based Feature Extraction using Hard and Soft Thresholding Methods in the Classification of Epilepsy Risk Levels. *European Journal of Scientific Research*,71, 221-232.
2. Ganesh Babu.C & Vanathi.P.T.(2011).Performance Analysis of Voice Activity Detection Algorithm for Robust Speech Recognition System under Different Noisy Environment. *Journal of Scientific and Industrial Research*, 69, 7, 132-141.
3. Gotman.J.(1982).Automatic recognition of epileptic seizures in the EEG. *Electroencephalogr. Clin. Neurophysiol.*, 54, 530–540.
4. ShinyinTsens, Sciencedirect, (1995). Evaluation of parametric methods in EEG signal analysis .*Expert system*, 17, 34-46.
5. Jolliffe IT. (1988).Principal Component Analysis New York: *Springer*-Verlag.
6. Choudrey RA, Roberts SJ. (2003) Bayesian, Ica with hidden markov model sources. *International Conference on Independent Component Analysis*. 809-814.
7. McSharry PE &Clifford GD.(2004) A comparison of nonlinear noise reduction and independent component analysis using a realistic dynamical model of the electrocardiogram. *Proc of SPIE International Symposium on Fluctuations and Noise* .5467(09):78.88.
8. Penny W., Everson R & Roberts S.(2000). Hidden Markov Independent Components Analysis", 2000. URL citeseer.ist.psu.edu397426.html.
9. Makhoul J. Roucos s & Gish,Vector H.(1985). Quantization in Speech Coding. *proc.IEEE*.73,1551-1588.
10. Rabiner R.(1989).Hidden Markov Models and Selected Application in Speech Processing.*proceedings of IEEE*, 77,2.
11. Lee T-W, Girolami M., Bell A.J.& Sejnowski T.J.A unifying Information-theoretic framework for Independent Component Analysis. *International Journal on Mathematical and Computer Modeling, in press*.
12. Qu H & Gotman J. (1998).A patient-specific algorithm for the detection of seizure onset in EEG monitoring: Possible use as a warning device. *IEEETrans. Biomed. Eng.*,44,2, 115–122.
13. Penny W & Roberts S.(1998).Gaussian observation hidden markov models for EEG analysis.*Tech. Rep. TR*-98-12.

14. Jeng-Shyang Pan, Zhe-Ming Lu & Sheng-He Sun. (2003). An Efficient Encoding Algorithm for Vector Quantization Based on Subvector Technique. *IEEE transactions on image processing,* 12, 460-472.
15. Gabor A J.(1998).Seizure detection using a self-organizing neural network: validation and comparison with other detection strategies. *Electroencephalography and Clinical Neurophysiology,* 107, 1, 27-38.
16. Marchesi B, Stelle A &Lopes L. (1997). Detection of epileptic events using genetic programming *Proceedings of the 19th Annual International Conference of the IEEE EMBS,* 3 ,527-535.
17. Osorio, M, Frei G & Wilkinson S B.(1992).Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset. *Epilepsia,* 39,140-153.
18. Kekre H B,VaishaliKulkarni. (2010).Speaker Identification by using Vector Quantization. *International Journal of Engineering Science and Technology,* 2(5), 1325-1331.
19. Mario A & Figueiredo.(2001).*Scalar and Vector Quantization*.DepartamentdeEngenhariaElectrotecnica e de Computadores, Instituto Superior Tecnico, Lisboa, Portugal.
20. J.Clin.(1994).*NeurophysiolAmerican Electroencephalographic Society.Guideline thirteen*. Guidelines for standard electrode position nomenclature.
21. Xu L, Cheung C, Yang H & Amari S.(1997).Independent Component Analysis by the Information-theoretic Approach with Mixture of Densities. *IEEE Trans. Neural Networks,* 7,1821-1826.
22. Oja E.(1999).Nonlinear PCA criterion and maximum likelihood in independent component analysis. *Proc. ICA'99 Int. Workshop Independent Component Analysis and Signal Separation ,Aussois, France,* 143-148.
23. Lu M, Pellouchoud E & Gevins. (1998). A rapid method for determining standard 10/10 electrode positions for high resolution EEG studies. *ElectroencephalogramClincNeurophysiol* ,106:554±558.
24. Cohen J A. (1960).A coefficient of agreement for nominal scales. *EduPsycholMeas* 20:30-46.
25. Everitt. (1968).Moments of stastics kappa and weighted kappa.*Br J Math Psychol,* 21:97-103.
26. Jiang E, Kannan R, Littman M & Vempala L. (1999).*Efficient Singular Value Decomposition Via Improved Document Sampling,* Tech. Rep. CS-99-5, Department of Computer Science, Duke University.
27. Stewart GW.(1993). On the early history of the singular value decomposition. SIAM REV,6,452-463.
28. Jolliffe IT.(2002) *Principal Component Analysis*. Springer; New York.
29. Abdi H.(2007). *Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition(GSVD). In:Salkind NJ, ed. Encyclopedia of Measurement andStatistics.* Thousand Oaks: Sage Publications.

30. Dray S.(2008).On the number of principal components: a test of dimensionality based on measurements of similarity between matrices. *Comput Stat Data Anal*,52:2228–2237.
31. Saporta G, Niang N.(2009). Principal component analysis: application to statistical process control. In: Govaert G, ed. *Data Analysis*. John Wiley & Sons London, 1–23.
32. Esteller R, Vachtsevanos G, Echauz , D'Alessandro M, Bowen C, Shor R & Litt B.(1999). Fractal dimension detects seizure onset in mesial temporal lobe epilepsy.*IEEE EMBS*,1 442-455.
33. Esteller R.(1999).Fractal dimension characterizes seizure onset in epileptic patients. *ICASSP '99: Proceedings of the Acoustics, Speech, andSignal Processing, 1999.IEEE International Conference. Washington,USA: IEEE Computer Society,* 2343-2346.
34. Ocak H.(2009).Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy,*Expert Syst. Appl.*, 36, 2,2027-2036.
35. McSharry P M. (2002). Linear and non-linear methods for automatic seizure detection in scalp electro-encephalogram recordings. Medical Rev,40,234-245.
36. Celka P & Colditz P.(2002).A computer-aided detection of EEG seizures in infants: a singular-spectrum approach and performance comparison.,*IEEE Transactionson Biomedical Engineering*,49,5, 455-462.
37. Ruey S. Huang, Chung J. Kuo, Ling-Ling Tsai & Oscal T.C. Chen. (1996). EEG pattern recognition arousal states detection and classification. *Proceedings of IEEE International Conference on Neural Networks*, 2, 641–646.
38. Obermaier B, Guger C & Pfurtscheller G. (1999). *HMM used for the offline classification of EEG data*. BiomedizinscheTechnik(Chapter 2).
39. Graja S & Boucher J M. (2003). Multiscale hidden Markov model applied to ECG segmentation. *IEEE International Symposium on Intelligent Signal Processing, Budapest, Hungary,* 1024-1028.
40. Easwaramoorthy & Uthayakumar R. (2010). Analysis of Biomedical EEG Signals using Wavelet Transforms and multi fractal Analysis, IEEE International Conference on Communication Control and Computing Technologies (ICCCCT), 544-549.