Int J Pharm Bio Sci 2016 April; 7(2): (B) 687- 693

Original Research Article Neurobiology

International Journal of Pharma and Bio Sciences ISSN 0975-6299

SOFT THRESHOLDING TECHNIQUES WITH PCA AS POST CLASSIFIER FOR EPILEPSY RISK LEVEL CLASSIFICATION

SUNIL KUMAR PRABHAKAR[1] AND HARIKUMAR RAJAGURU[2]

[1]Research Scholar, Department of ECE, Bannari Amman Institute of Technology, India
[2]Professor, Department of ECE, Bannari Amman Institute of Technology, India

ABSTRACT

Due to the frequent electrochemical impulses in the neurons, recurrent and rapid disturbances of mental functions occur which is known as epilepsy and it has a drastic effect on the lives of human beings. The excessive discharge is termed as epileptic spikes which help in diagnosis and analysis of epilepsy. Electroencephalography (EEG) serves as a vital clinical tool for the analysis of epilepsy. Because of its self adaptive nature, Artificial Neural Network (ANN) has been widely used to classify the epilepsy risk levels from EEG Signals. This paper includes the concept of Soft Thresholding (ST) techniques followed by the implementation of Principal Component Analysis (PCA) for the Classification of Epilepsy Risk Levels from EEG Signals. The analysis is done in terms of benchmark parameters such as Performance Index (PI), Quality Values (QV), Sensitivity, Specificity, Time Delay and Accuracy.

KEYWORDS: EEG, epilepsy, ANN, PCA, PI

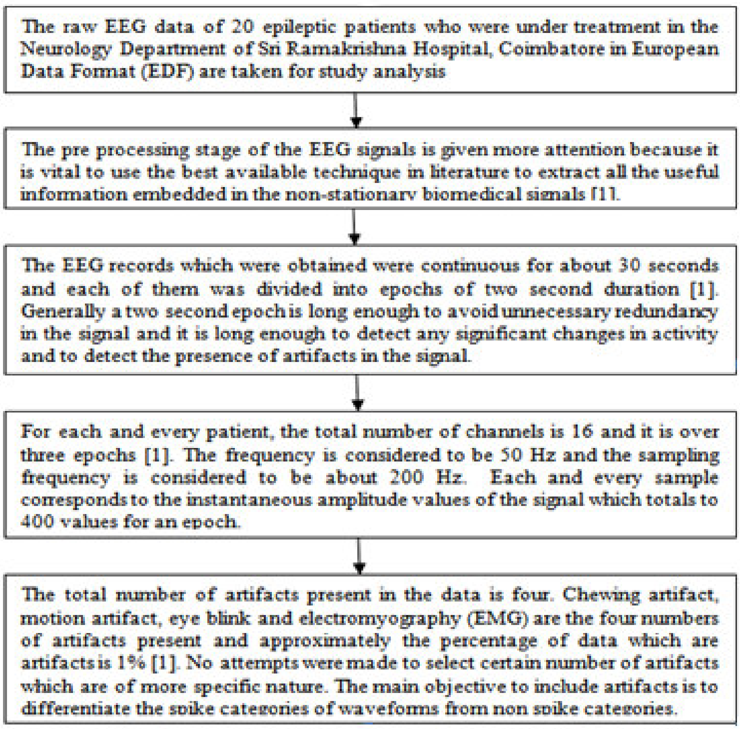
SUNIL KUMAR PRABHAKAR
Research Scholar, Department of ECE, Bannari Amman Institute of Technology, India

INTRODUCTION

Epilepsy is one of the collections of neurological disorders [1]. Due to the abnormalities in the electrical activity, sudden recurrent reactions in the brain functions occur and the manifestation is done clinically and termed as epileptic seizures. Since epilepsy is the recurrence of abnormalities in the electrical activities of the cortical regions, it is important to visualize and monitor the electrical activities in the cerebral cortex [5]. Examining the electrical activities is done by theElectroencephalogram (EEG) signals and it aids greatly in predicting the epileptic seizures. To improve the quality of life of epileptic patients, an automatic system has to be developed which would be able to herald seizures very earlier so that remedial and preventive measures could be taken in order to prevent mortality rates concerning the epilepsy [7]. The ambulatory recording of EEG signals seems to be the most conventional techniques, but the main drawback associated with it is that it takes a very long duration to record because of the huge length of the EEG signals in order to detect traces of epilepsy [6]. So, some sort of automatic detection techniques are used to detect and classify the epilepsy risk levels from EEG Signals [9]. In this paper, a soft thresholding concept is adopted followed by the classification using Principal Component Analysis Classifiers (PCA) for the perfect classification of epilepsy risk levels from EEG signals. This paper is organized as follows: In Section 2, the materials and methods are discussed, followed by the concepts of soft thresholding in Section 3 and PCA is used as a Post Classifier for the Classification of epilepsy risk levels from EEG signals in Section 4. Section 5 gives the results and conclusion.

MATERIALS AND METHODS

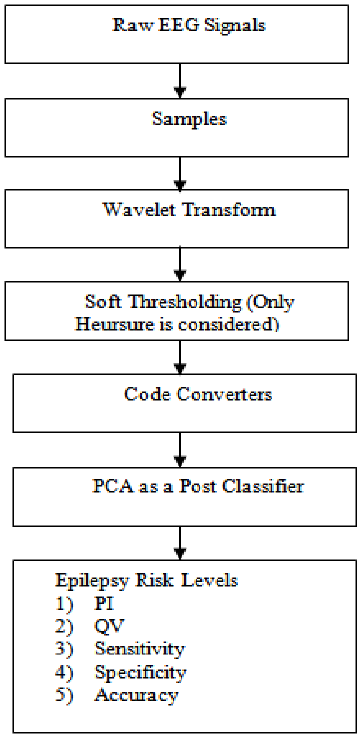
Figure 2.1
Acquisition procedure of the EEG Signals



For the detailed performance analysis of the epilepsy classification levels using soft thresholding technique and PCA as a Post Classifier, the flow chart is followed as shown in Figure 2.1

The figure 2.2 shows the block diagram of the procedure. Initially the raw EEG Signals are taken and it is sampled. Then wavelet transform is applied to the sampled signals. Then the concept of soft thresholding techniques is utilized and it is fed to code converters. Finally the output of the code converters is fed as an input to the Principal Component Analysis (PCA) and then the epilepsy risk levels are classified from the EEG signals in terms of Perfect Classification (PC), Missed Classification (MC), Performance Index (PI), Quality Value (QV), Time Delay (TD), Specificity, Sensitivity and Accuracy.

Figure 2.2
Block Diagram of the Procedure



Soft thresholding and code converters
For classification of the risk level of patients, certain parameters are considered as follows. The energy in a very efficient manner for each and every epoch is computed as follows

$$E = \sum_{i=1}^{n} x_i^2$$

Using the formula mentioned below, variance for each and every epoch is computed as

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}$$

The covariance of the entire duration is computed easily as follows

$$CD = \frac{\sum_{i=1}^{p}(D - t_i)^2}{pD^2}$$

Using the wavelet transforms, the four parameters are obtained as shown in reference [1]. Firstly, above the threshold, the total number of positive and negative peaks is found. Secondly, if zero crossing functions lies between 20 to 70 ms, then spikes can be detected and if zero crossing functions lies between 70 ms to 200 ms, then sharp waves can be detected. Thirdly the number of spikes and waves are together represented and determined as events. Finally, the duration for each and every event is determined.

Wavelet Transform
To extract the transient features, wavelet is the best choice. If a function $q(t)$ is considered, then wavelet transform of this function and is mathematically expressed as

$$wf(a,b) = \int_{-\infty}^{\infty} q(t)\varphi_{a,b}^{*}(t)dt$$

Where $\varphi^*(t)$ is the complex conjugate of the wavelet function $\varphi(t)$.
Soft thresholding [4] (wavelet shrinkage) is given by the following equation

$$\rho_T(x) = \begin{cases} x - T, if\,(x \geq T) \\ x + T, if\,(x \leq T) \\ 0, if\,|x| < -T \end{cases}$$

In this paper only Heursure soft threshold method is considered and compared with other soft thresholding methods such as Minimax, Rigrsure and Sqtwolog, Heursure soft threshold method is given much importance because it is a mixture of SURE and Sqtwolog threshold methods. Heursure selects the same thresholds as SURE at lower levels. While at higher levels (level3 and more), Heursure chooses higher thresholds which are close to the values selected by Sqtwolog. That means it tends to keep more co-efficients than minimax at low level. Expert's knowledge [1] aids us greatly to identify the different parametric ranges for five unique linguistic risk levels for the patients as shown in table 3.1.

Table 3.1
Risk Level Classifications

Risk levels / *Normalized Parameters*	Normal	Low	Medium	High	Very high
Energy	0-1	0.7-3.6	2.9-8.2	7.6-11	9.230
Variance	0-0.3	0.15-0.45	0.4-2.2	1.6-4.3	3.8-10
Peaks	0-2	1-4	3-8	6-16	12-20
Events	0-2	1-5	4-10	7-16	15-28
Sharp Waves	0-2	1-5	4-8	7-11	10-12
Average Duration	0-0.3	0.15-0.45	0.4-2.4	1.8-4.6	3.6-10
Covariance	0-0.05	0.025-0.1	0.09-0.4	0.28-0.64	0.54-1

Working on numerical is difficult, since it demands an exact decimal accuracy. So the inputs are generally encoded as string of alphabets and the five different classification levels are represented alphabetically as shown in Table 3.2

Table 3.2
Alphabetical Representation of Risk Levels

Risk Level	Representation
Normal	U
Low	W
Medium	X
High	Y
Very High	Z

With the encoding process, where each risk level is encoded to any one of the five states, a total of seven characters is obtained for each of the 16 channels of each epoch. The classification of the risk level varies in between the adjacent epochs. The sample output from each epoch is not identical and is highly varying in conditions and can be represented as [YYZXXYZ] to [YZXXYZZ] to [XYXXZYY] and so on.

PRINCIPAL COMPONENT ANALYSIS AS A POST CLASSIFIER

A particular set of data can be easily expanded into a set of orthogonal components using PCA analysis. On obtaining PCA, the maximum decorrelation of the signals can be easily achieved. PCA also aids in the separation of the entire data into both subspaces and domains such as noise and signal subspaces [2]. For a n-by-p data matrix, X always returns the co-efficients of the principal components which are termed as loadings. The original variables combined in a linear fashion together forms as the principal components. The orthogonality should be maintained between each and every variable thereby redundancy is completely eliminated here. The PCA can be defined if the data matrix is considered as follows

$$X = |x_{ij}|$$

in which the columns represent the *p* variables and rows represent measurements of *n* objects or individuals on those variables. The data can be represented by a cloud of *n* points in a *p*-dimensional space, each axis corresponding to a measured variable [3]. The derived variable defined by the operation is of the following form of equation

$$Y_1 = a_1x_1 + a_2x_2 + \cdots . a_px_p$$

With coefficients a_i satisfying the conditions of the following equation

$$\sum_{i=1}^{p} a_i^2 = 1$$

This process continous until *p* mutually orthogonally is found out. Each and every line defines a derived variable as is mathematically expressed as

$$Y_i = a_{1i}X_1 + a_{2i}X_2 \ldots \ldots a_{pi}$$

The Y_i thus obtained are called as the main Principal Components of the system and the process of acquiring them is called Principal Components Analysis.

RESULTS AND CONCLUSION

For the clear performance analysis of soft thresholding techniques (only Heursure) along with the Principal Component Analysis (PCA) based on the Quality values, Time Delay and Accuracy the results are computed in Table 5.1. The formulae for the Performance Index (PI), Sensitivity, Specificity and Accuracy are given as follows

$$PI = \frac{PC - MC - FA}{PC} \times 100$$

Where PC – Perfect Classification, MC – Missed Classification, FA – False Alarm,
The Sensitivity, Specificity and Accuracy measures are stated by the following

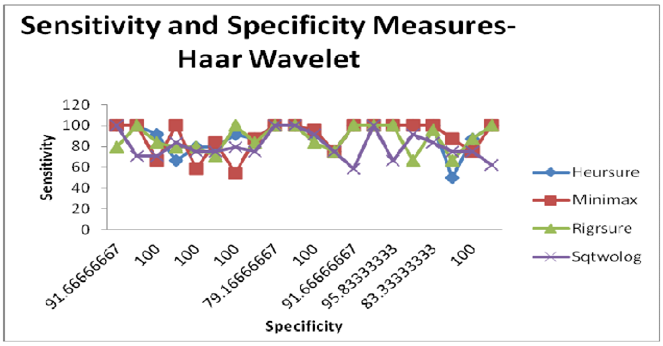
$$Sensitivity = \frac{PC}{PC + FA} \times 100$$

$$Specificity = \frac{PC}{PC + MC} \times 100$$

$$Accuracy = \frac{Sensitivity + Specificity}{2} \times 100$$

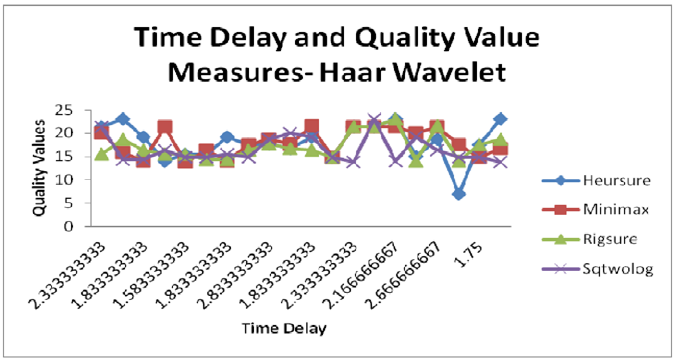
The Specificity and Sensitivity Analysis for the Soft Thresholding operations along with the PCA in comparison is shown in Figure 5.1. The Time Delay and Quality Value Measures for the Soft thresholding operations along with the PCA comparison is shown in Figure 5.2. The Performance Index and Accuracy for the Soft thresholding operations along with the PCA comparison for the classifiers in comparison isshown in Figure 5.3.

Figure 5.1
Specificity and Sensitivity Measures for Haar Wavelet



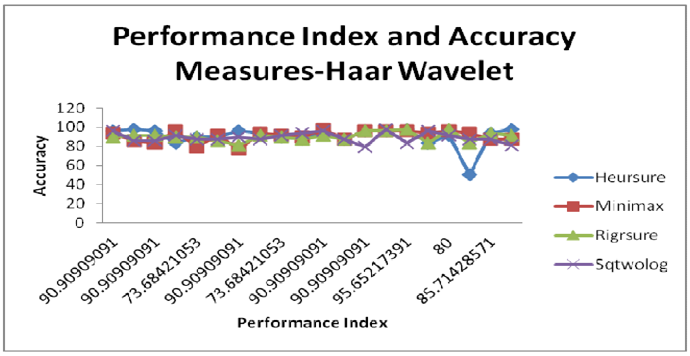
It is inferred from figure 5.1 that the specificity measures are not constant throughout and there are abrupt random variations throughout the series. An average specificity rate of 90.83% is obtained for Haar , 97.7% for dB2, 95.83% for Sym8 and 91.87% for dB4.

Figure 5.2
Time Delay and Quality Value Measures for Haar Wavelet



It is inferred from figure 5.2 that the time delay values are not constant throughout and an average time delay of about 2.15 seconds is found for Haar, 1,82 seconds for dB2, 1.94 for Sym8 and finally 2.13 seconds for dB4.

Figure 5.3
Performance Index and Accuracy Measures for Haar Wavelet



It is inferred from figure 5.3 that there are slight changes in performance index measures and an average PI of about 81.41% is obtained for Haar wavelet, 79.04% for dB2, 79.47% for Sym 8 and finally 75.96% for dB4 wavelet.

Table 5.1
Performance Comparison of Soft (Heursure) Thresholding with PCA Classifier

Parameters	Haar	dB2	Sym8	dB4
PC (%)	82.5	84.58	84.79	82.29
MC (%)	8.33	2.29	4.16	8.12
FA (%)	9.16	13.12	11.04	9.58
PI (%)	81.41	79.04	79.47	75.96
Sensitivity %)	90	86.87	88.95	90.41
Specificity(%)	90.83	97.70	95.83	91.87
Time Delay (sec)	2.15	1.82	1.94	2.13
Quality Values	18.06	18.53	18.59	17.82
Accuracy (%)	90.41	92.29	92.39	91.14

It is thus concluded that the average perfect classification rate is high for Sym 8 wavelets as about 84.79% when compared to the other wavelets. A high Quality Value of about 18.59 is found out when compared to the other wavelets and also the accuracy is also quite higher in Sym 8 as of 92.39 %. Thus, it is concluded that when soft thresholding technique (Heursure) is engaged with PCA as Post Classifier, then Sym 8 wavelets outperforms the other wavelet and thus it is considered to be a very versatile system. Future work may incorporate the possible usage of different types of post classifiers for this similar kind of analysis.

REFERENCES

1. Harikumar R, Vijayakumar T, Sreejith, MG, Performance Analysis of SVD and Support Vector Machines for Optimization of Fuzzy outputs in classification of Epilepsy risk level from EEG Signals. Proceedings of IEEE Conference on Recent Advances in Intelligent Computational Systems (RAICS).2011 :718-28.
2. Xu L, Cheung C, Yang H, Amari, S, 'Independent Component Analysis by the IC Approach with Mixture of Densities', IEEE Trans. Neural Networks, 1997, vol (5): 1821-26.
3. Jolliffe, IT, 'Principal Component Analysis', Springer-Verlag, New York, 1998
4. Hazarika N, Chen JZ, Tsoi AC & Sergejew A, 'Classification of EEG signals using the wavelet transform', Signal Processing, 1997 : 59(1) pp. 61-72.
5. Gotman J, 'Automatic recognition of epileptic seizures in the EEG', Electroencephalogr. Clin. Neurophysiol., 1982, vol. 54, pp. 530-40
6. Finley KH, DynesJB, 'Electroencephalographic studies in epilepsy: a critical analysis', Brain, 1942, vol.65, pp. 256-65.
7. Gabor AJ , Seyal M, 'Automated interictal EEG spike Detection using artificial neural networks', Electroenceph. Clin. Neurophysiol, 1992 ,vol. 83, pp. 271-80.
8. Gabor AJ 'Seizure detection using a self-organizing neural network: validation and comparison with other detection strategies', Electroencephalography and Clinical Neurophysiology, 1998, 107(1), pp. 16-27.
9. Geetha G, Geethalakshmi SN , 'Detectingepileptic seizure using electroencephalogram: A new and optimized method for seizure classification using hybrid extreme learning machine'. Proceedings of International conference on Process automation, control and computing (PACC), July 2011 July, pp. 1-6.