

Performance Comparison of IST and Multi Scale Principal Component Analysis in the EEG Signal Processing

Dr.B.Krishna Kumar¹ and Dr.K.V.S.V.R.Prasad²

¹Professor, Dept.of ECE, Methodist College of Engineering and Technology, Hyderabad, TS, INDIA
e-mail: saisantu2004@yahoo.co.in

² Prof & Head Dept. of ECE, D.M.S.S.V.H Engg.College, Machilipatnam, Andhra Pradesh
e-mail: kvsvr@yahoo.com

ABSTRACT

The removal of Ocular Artifacts (OA) in Electroencephalogram (EEG) data is one of the key challenges in the analysis of brain recordings. Brain activity produces electroencephalogram signals, which consists of some of vital signs of neurological disorders such as epilepsy, tumor cerebrovascular lesions and the problems associated with the trauma. These signals can be acquired by placing the electrodes on the scalp at specified positions and exists in order of 1-5 μ v, whose frequency range is DC-64 Hz. Acquisition of these signals mainly suffers from different unwanted signals (artifacts or noise) resulting in less signal information for identification.

In this paper, two algorithms are proposed namely, Multi Scale Principal Component Analysis (MSPCA) and Iterative Soft Thresholding (IST) using wavelets in removing the Ocular Artifacts (OA) present in the EEG signals. This paper discusses not only the performance comparison of two algorithms on statistical parameters of EEG signals such as Signal to Noise Ratio, (SNR), SNRI or Noise Figure (NF) and Absolute Average Error (AAE) but also estimated the run time of each algorithm i.e., computational time of each algorithm.

Keywords: EEG, EOG, Wavelets, Denoising, Thresholding, PCA, Multi scale PCA(MSPCA), Ocular artifacts and Wavelet transform, Iterative Soft Thresholding

1. INTRODUCTION

Electrooculography (EOG) artifacts inevitably and frequently interfere with the electroencephalogram (EEG) signals. Eye-movement and eye-blink artifacts are the main sources of ocular artifacts. The frequency components of EEG signals are in the order of 1-5 μ V, and their frequency content differs among the different neurological rhythms, such as, alpha, beta, delta and theta rhythms [1-4]. However,

artifacts are the outstanding enemies of high quality EEG signals. These artifacts basically fall into two major categories: technical and physiological artifacts. The technical artifacts are often found in power line noise 50/60Hz results from poor electrode application on the scalp and transducer's artifacts. The physiological artifacts are often due to ocular, heart and muscular activity; are the EOG, ECG and EMG artifacts respectively [5].

To date, many methods were presented to remove EOG artifacts. Some methods were based on regression in the time domain or frequency domain [6-8] for removing eye blink artifacts. However, they always need a reliable reference channel. Moreover, EOG Reference channel often contains brain signals which will be also removed inevitably from the EEG by the procedure of regression. Therefore, the methods based on regression may not be the best way to remove EOG artifacts.

Principal Component Analysis is among the most popular methods for extracting the information from the data and has found applications in a wide range of disciplines [9]. PCA was introduced by Pearson in 1901[10] and developed by Hotelling [11] in the year 1933. Using Multi Scale Principal Component Analysis the EEG signal is decomposed into 6 levels. Later, estimation of the statistical parameters of EEG signals is done. Finally the performance comparison of these two algorithms is made.

In this paper, IST[12] using wavelets algorithm [13] the EEG signal is decomposed into a level 4, which gives us approximate and detail coefficients. The detail coefficients which are having more noise information are processed with IST and later

estimated the statistical parameters such as SNR, SNRI and AAE [13].

The noisy EEG signal is convolved with a low and high pass filter whose impulse response is determined by the wavelet chosen. The output of each filter produces the same number of samples as the original signal, so both outputs are down sampled by 2 resulting in the Approximate and detail coefficients each with half the number of points as that of the original signal.

The coefficients represent a correlation between the signal of interest and wavelet chosen at different scales and during translation. Because all of the coefficients are preserved, the original signal or any level of decomposition can be reconstructed using a filter scheme similar to decomposition shown in Figure 1. The process is reversed and now the coefficients are up sampled (interpolated), filtered, and summed.

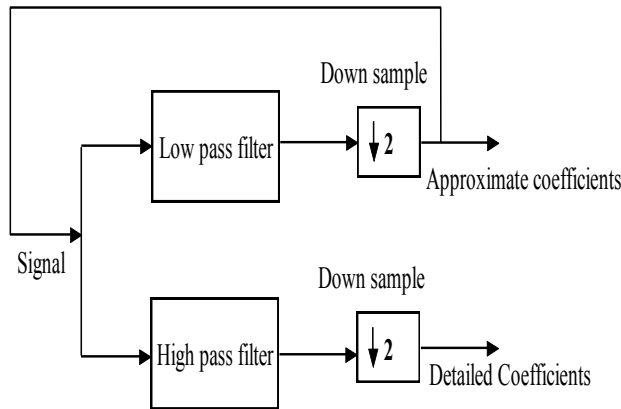


Figure.1. Block diagram to show DWT decomposition

The EEG signal is decomposed into 4 levels using dB10, Sym8 and Haar wavelets and corresponding approximate and detail coefficients are shown in the Fig.2. From derived coefficients high frequency components will be distributed in approximate coefficients and low frequency components presents in Detail coefficients cD4, cD3, cD2, cD.

2. ITERATIVE SOFT THRESHOLDING

This method describes an approach for the restoration of degraded signals using sparsity. This approach, which has become quite popular, is useful for numerous problems in signal processing: denoising, de convolution, interpolation, super-resolution, de

clipping, etc [12]. The observed signal y can be written as

$$y = Hx + n$$

Where x is the signal of interest which one would like to estimate, n is additive noise, and H is a matrix representing the observation processes. For example, if one observes a blurred version of x then H will be a convolution matrix. The estimation of x from y can be viewed as a linear inverse problem. A standard approach to solve linear inverse problems is to define a suitable objective function $J(x)$ and to find the signal x minimizing $J(x)$ [12]

Generally, the chosen objective function [12] is the sum of two terms:

$$J(x) = D(y, Hx) + \lambda R(x)$$

Where $D(y, Hx)$ measures the discrepancy between y and x and $R(x)$ is a regularization term (or penalty function). The parameter λ is called the regularization parameter and is used to adjust the trade-o between the two terms; λ should be a positive value. On one hand, we want to find a signal x so that Hx is very similar to y ; that is, one would like to find a signal x which is consistent with the observed data y . For $D(y, Hx)$, we will use the mean square error, namely

$$D(y, Hx) = \|y - Hx\|_2^2$$

The $\|v\|_2^2$ represents the sum of squares vector ' v '

2.1 Iterative Threshold Algorithm

The objective function minimization, previous case

$$J(x) = \|y - Hx\|_2^2 + \lambda \|x\|_1$$

The $J(x)$ can be minimized by combining Majorization-Minimization approach

To find a majorizer of $J(x)$, coincides $J(x)$ at x_k and which is easily minimized. Adding non-negative function to $J(x)$

$$G_k(x) = J(x) + (x - x_k)^t (\alpha I - H^t H)(x - x_k)$$

By design $G_k(x)$ coincides $J(x)$ at x_k . one needs to minimize $G_k(x)$ to get x_{k+1} .

From the above equation, we can get

$$G_k(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + (\mathbf{x} - \mathbf{x}_k)^t (\alpha \mathbf{I} - \mathbf{H}^t \mathbf{H}) (\mathbf{x} - \mathbf{x}_k)$$

Rewriting the equation using soft thresholding,

$$G_k(\mathbf{x}) = \alpha \left\| \mathbf{x}_k + \frac{1}{\alpha} \mathbf{H}^t (\mathbf{y} - \mathbf{H}\mathbf{x}_k) - \mathbf{x} \right\|_2^2 + \lambda \|\mathbf{x}\|_1 + K$$

Where, K is constant w.r.t x. Minimizing $G_k(x)$, equivalent to Minimizing $(1/\alpha)G_k(x)$ so x_{k+1} , obtained by minimizing:

$$\left\| \mathbf{x}_k + \frac{1}{\alpha} \mathbf{H}^t (\mathbf{y} - \mathbf{H}\mathbf{x}_k) - \mathbf{x} \right\|_2^2 + \frac{\lambda}{\alpha} \|\mathbf{x}\|_1$$

Therefore minimizing is achieved by the following soft-thresholding equation [12]

$$\mathbf{x}_{k+1} = \text{soft} \left(\mathbf{x}_k + \frac{1}{\alpha} \mathbf{H}^t (\mathbf{y} - \mathbf{H}\mathbf{x}_k) - \mathbf{x}, \frac{\lambda}{2\alpha} \right)$$

Where $\alpha \geq \text{maxeig}(\mathbf{H}^t \mathbf{H})$. This gives the Iterated Soft Threshold Algorithm.

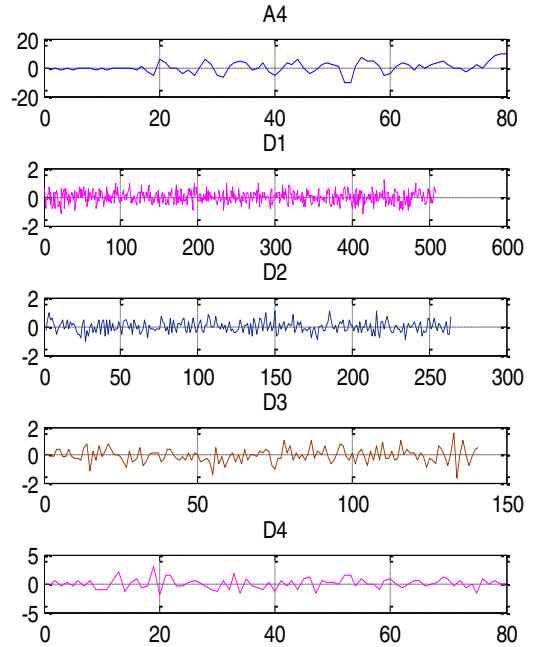
The iterative soft Thresholding algorithm is used for denoising the EEG signal. The model consists of three main steps i.e, normalization, wavelet decomposition and applying the IST for the detail coefficients. In this paper EEG signals are taken from Physionet database [15].

Data can be normalized to increase the accuracy of the application as follow:

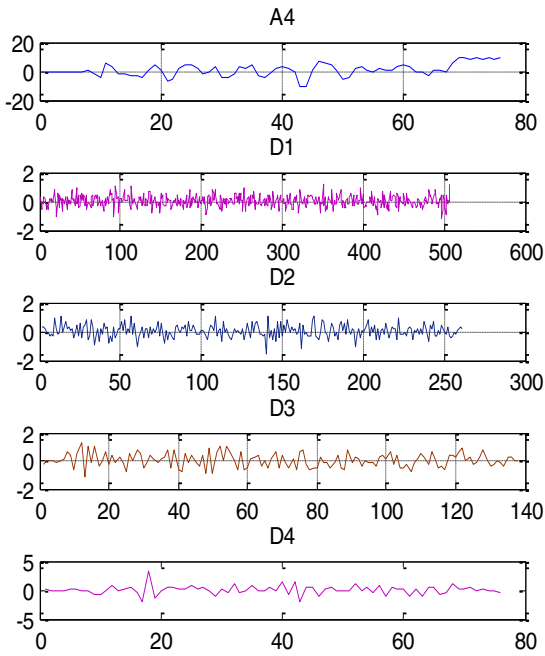
$$\mathbf{X}_{\text{norm}} = \frac{\mathbf{X} - \bar{\mathbf{X}}}{\text{std}(\mathbf{X})}$$

Normalizing of EEG signal is very important because of possible variations in signal acquisition from trial to trial, hence the normalization of data is necessary for valid results. EEG data is decomposed into 4 levels using dB10, Sym8 and Haar wavelets as shown in Figure 2(a,b and c) and all the detail coefficients are scaled by IST algorithm[13], finally by applying inverse wavelet transform results the enhanced EEG signal. Algorithmic flow result is shown in Figure 3

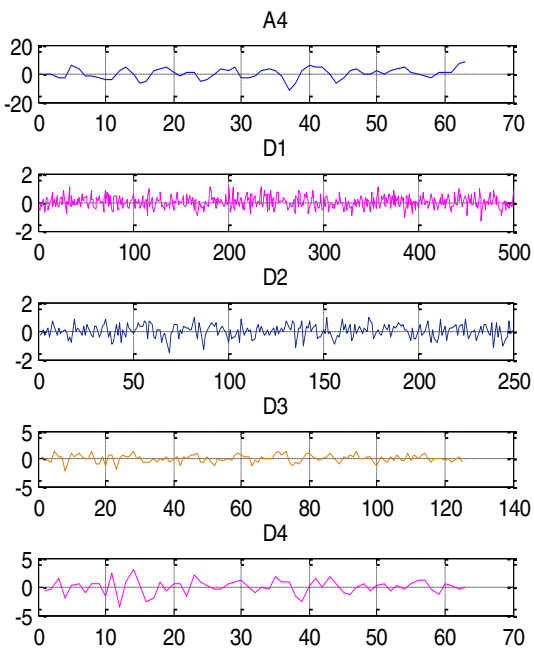
reveals that the presented method of modeling gives the good performance among all the previous techniques.



a) Approximate and detail coefficients using dB10



b) Approximate and detail coefficients using sym8



c) Approximate and detail coefficients using Haar

FIGURE 2

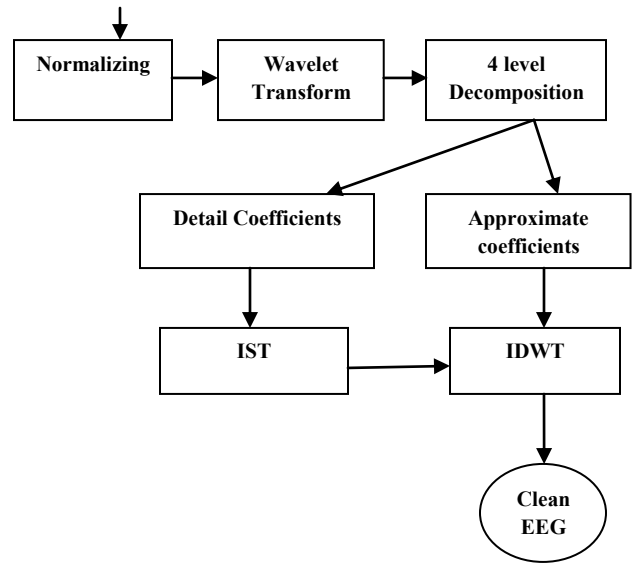


FIGURE 3: IST Using Wavelets Algorithm Flow

The noisy EEG signal and modeled EEG signal using IST algorithm is shown in the Figure 4.

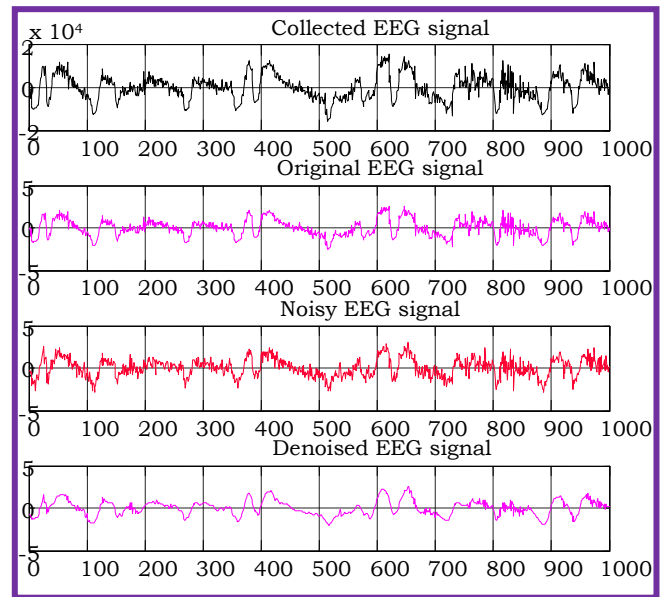


Figure.4: Noisy EEG Signal and Modeled signal

3. RESULTS OF IST

The results obtained from IST using wavelets algorithm [13] are provided in the Table.1

Table.1. statistical measures of EEG signal using IST

4. MULTISCALE PCA

PCA is a technique which is generally used for reducing the dimensionality of multivariate datasets i.e. reducing the number of dimensions, without much loss of information. Considering a vector of n random variables x for which the covariance matrix is Σ , the principal components (PCs) can be defined by $z = Ax$ Where z is the vector of n PCs and A is the n by n orthogonal matrix with rows that are the eigenvectors of Σ . The Eigen values of Σ are proportional to the fraction of the total variance accounted for the corresponding eigenvectors, so the PCs explaining most of the variance in the original variables can be identified. If, as is usually the case, some of the original variables are correlated, a small subset of the PCs describes a large proportion of the variance of the original data.

Multi scale PCA (MSPCA)[14,16] combines the ability of PCA to extract the cross correlation or relationship between the variables, with that of orthonormal wavelets to separate deterministic features from stochastic processes and approximately de correlate the autocorrelation among the measurements. To combine the benefits of PCA and wavelets, the measurements for each variable (column) are decomposed to its wavelet coefficients using the same orthonormal wavelet for each variable. This results in transformation of the data matrix, X into a matrix, WX , where W is an $n \times n$ orthonormal matrix representing the orthonormal wavelet transformation operator containing the filter coefficients.

MSPCA removes the artifacts like EMG and ocular artifacts. To remove the artifacts PCA is done twice on the same data by selecting only first four principal components, and the results were given in figures 5 for two different data sets taken from reference [15]. The results of MSPCA [14, 16] are provided in the Table.2.

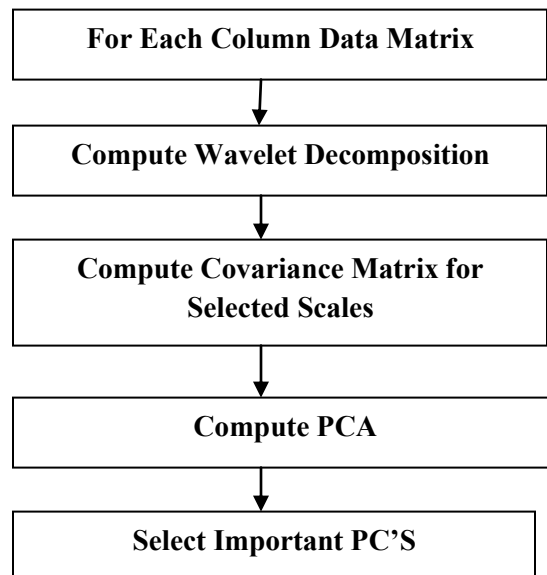
Table.2. statistical measures of EEG signal using MSPCA

Wavelet /Threshold	AAE(dB)	SNR(dB)	SNRI(dB)
Sym8 (soft)	23.58	34.95	25.34
Sym8 (hard)	19.26	24.52	48.00
Haar (soft)	24.89	20.29	10.75

Wavelet/Thresho Id type	AAE (dB)	SNR (dB)	SNRI or NF(dB)
Sym8 (soft)	23.58	34.95	25.34
Sym8 (hard)	19.26	24.52	48.00
Sym8 (IST)	18.26	38.27	50.20
Haar (soft)	24.89	20.29	10.75
Haar (hard)	18.18	25.10	46.08
Haar (IST)	17.22	39.40	52.30
dB10(soft)	14.88	28.89	34.58
dB10(hard)	19.98	23.74	47.69
dB10(IST)	14.44	29.43	49.43
Haar (hard)	18.18	25.10	46.08
dB10 (soft)	14.88	28.89	34.58
dB10 (hard)	19.98	23.74	47.69
MSPCA	13.72	37.43	46.91

From Table.1 and Table.2 it is clearly observed that both the proposed algorithms i.e., IST and MSPCA are providing better results over the conventional Thresholding techniques namely soft and hard Thresholding.

The algorithmic flow of MSPCA [14] is shown in the Figure.5.



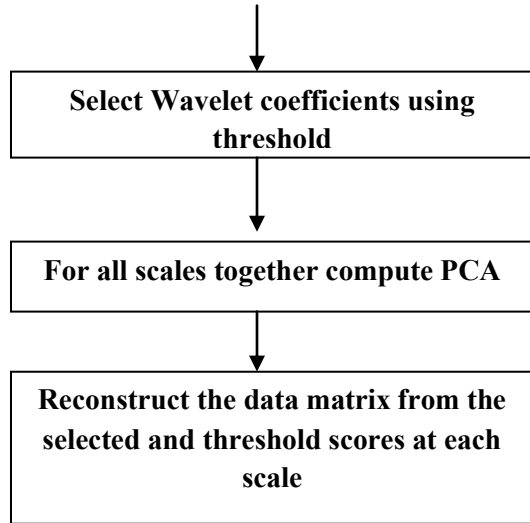


FIGURE.5: MSPCA Algorithm Flow

The noisy EEG signal and modeled EEG signal using Multi Scale Principal Component Algorithm is shown in the Figure 6.

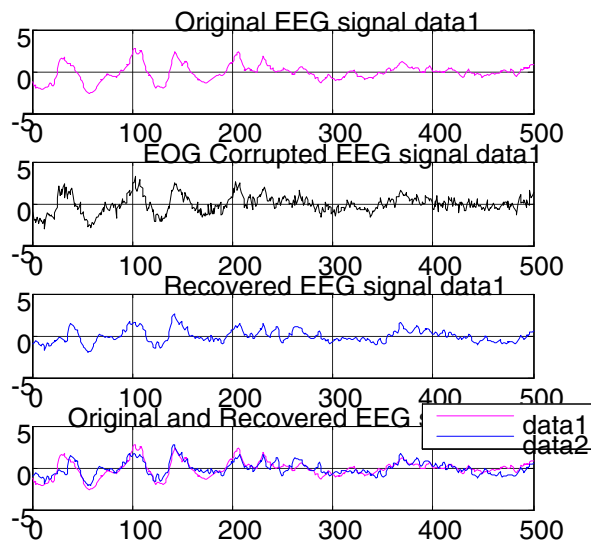


FIGURE.6: The noisy EEG signal and modeled EEG signal using MSPCA

5. RESULTS

Computational Efficiency of IST is better than MSPCA is provided in the Table.3.

Table.3. Computational Efficiency of IST and MSPCA

Algorithm	Computational Time(sec)
IST	1.391266
MSPCA	2.694674

6. CONCLUSIONS

From the above Table.1, Table.2 and Table.3 it is clear that IST not only providing better statistical parameters of EEG signal but also take less time in computing them when compared to MSPCA.

7. REFERENCES

- [1]. N.V. Thakor et al. "Multi Resolution Wavelet Analysis of Evoked Potentials", IEEE Transactions on Biomedical Engineering, Vol. 40, No 11, pp. 1085-1093, November, 1993.
- [2]J.Malmivuo and R. Plonsey, "Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields", New York: Oxford University Press, New York, Ch.28, 1995.
- [3] D. Nov'ak, L. Lhotsk'a, V. Eck, and M. Sorf, "EEG and VEP Signal Processing. Cybernetics", Faculty of Electrical Engineering, pp.50-53, 2004.
- [4] L. S'ormmo and P. Laguna, "Bioelectrical Signal Processing in Cardiac and Neurological Applications", Elsevier, Academic Press, 2005.
- [5]. S. Ventakaramanan, P. Prabhat, S.R Choudhury, H.B Nemade, and J.S. Sahambi "Biomedical Instrumentation Based On Electrooculogram (EOG) Signal Processing And Application To A Hospital Alarm System", Indian Institute Of Technology (IIT) Guwahati, Proceedings of IEEE ICISEP, pp.535-539, 2000.
- [6] Schlögl A, Keinrath C, Zimmermann D et al. "A fully automated correction method of EOG artifacts in EEG recordings" Clinical Neurophysiology.118, pp. 98-104, 2007.
- [7] Jung T-P, Makeig S, Humphries C, Lee T-W, McKeown MJ, Iragui V and Sejnowski TJ. "Removing Electroencephalographic artifacts by blind source separation" Psychophysiology.37, pp.163-178, 2000.
- [8] Gratton G, Coles MG and Donchin E "A new method for off-line removal of ocular artifact", Electroencephalography Clinical Neurophysiology 55(4), pp.484-486, 1983.
- [9] M.Aminghafari, N. cheze, J.M Poggi, "Multivariate denoising using wavelets and principal component Analysis," computational statistics & Data Analysis, vol. 50 pp. 2381-2398, 2006.
- [10] Pearson, K, "On Lines and Planes of Closest Fit to Systems of Points in Space", Phil. Magazine, vol.6, issue. 2, pp.559-572, 1901.
- [11] Hotelling. H," Analysis of a Complex of Statistical Variables Into Principal Components", J. Educ. Psychol., 24, pp.417-441, pp.498-520, 1933.
- [12] Ivan W. Selesnick, "Sparse Signal Restoration", pp.1-16, April 6, 2010.
- [13] B. Krishna Kumar, Dr. K.V.S.V.R. Prasad, Dr. K. Kishan Rao and J. Sheshagiri Babu, "Modeling of EEG Signals using Wavelets with Iterative Soft Thresholding", i-

manager's Journal on Future Engineering & Technology, Vol.6, No.4, pp.37-42, May - July 2011.

[14]B.Krishna Kumar, Dr.K.V.S.V.R.Prasad, Dr.K.Kishan Rao and Narsimha Baddiri, "Analysis of EEG Signals Using Multi-Scale Principal Component Analysis", IEEE Conference, Cape Institute of Technology, Levingapuram, Kanyakumari, pp.359-363, December15-18, 2011.

[15]<http://physionet.ph.biu.ac.il/physiobank/database>

[16] www.mathworks.com