Compression of EEG Signals Using Two-level Principal Component Analysis

B. Krishna kumar[1], Dr. K.V.S.V.R. Prasad[2], Prof.R.K.Singh[3] and D.Alekhya[4]
[1]Prof. Dept of ECE, GNITC, Telangana, INDIA.
[2]Professor & Head, Dept. of ECE, DMSSVH College of Engineering, AP, INDIA.
3Asociae Director, GNITC,Telangana, INDIA.
[4] IV B.Tech Scholar(ECE) ,GNITC,Telangana, INDIA.

Abstract---**Electro-encephalogram (EEG) is the electrical activity of brain cell groups in the cerebral cortex or the scalp surface. It plays an important role in studying the patient mental condition and Human Machine interfacing. Normal EEG signals can avail in the band of DC to 100Hz frequencies with a few hundreds of micro volts of strength. Ocular artifacts and muscular noise with similar statistical properties are the major challenges which make the analysis more complex and may yields wrong interpretation. Here un this paper we implemented a tool called Two Scale PCA to enhance the EEG data in the noise environment. Two scale PCA combines the ability of PCA to de correlate the variables by extracting a linear relationship, with that of wavelet analysis to extract deterministic features and approximately de correlate auto correlated measurements. Due to its multi scale nature, Two scale PCA is appropriate for modeling of data containing contributions from events whose behavior changes over time and frequency so that it can be utilized efficiently to process the EEG data. It has been investigated that the experimental results show the superiority of the proposed method over wavelet based methods.**

***Index Terms* – EEG, EOG, PCA, Two scale PCA, Ocular artifacts and Wavelet transform.**

I. INTRODUCTION

Biomedical signals such as ECG (Electro cardio Gram), EEG (Electro Ecephalo Graphy), EOG (Electro Occulo Gram) and PPG (Poly Plethismo Graphy) etc. are very helpful in diagnosis and as well as in artificial intelligence like Brain Computer Interference. All these signals are non stationary and low amplitude signals with DC to 100 Hz frequency range. It is complex to analyze the signals since they suffer from the interference of the other biomedical signals. Generally EEG signals are classifieds into different wave's i.e. Delta, Theta, Alpha, Beta and Gamma with respect to their oscillations in the range of 0-80Hz [1]. In this paper a novel technique called two level Principal Component Analysis is used to process the EEG signal to compress the EEG signals. This method of two level PCA is shows the superiority over the single level PCA in compressing the EEG data. This compressed data is very much useful in feature extraction as well as to estimate the RMSE of EEG signals effectively[2].

Principal component analysis (PCA) is among the most popular methods for extracting information from data, and has found application in a wide range of disciplines [3]. In chemical process operation, control, data rectification gross error detection, disturbance detection and isolation, statistical process monitoring. PCA transforms the data matrix in a statistically optimal manner by diagonalizing the covariance matrix by extracting the cross correlation or relationship between the variables in the data matrix. If the measured variables are linearly related and are contaminated by errors, the first few components capture the relationship between the variables, and the remaining components are comprised only of the error. Thus, eliminating the less important components reduces the contribution of errors in the measured data and represents it in a compact manner.

Wavelet transform is a power full tool with its multi resolution property for analyzing localized variations of power within a non stationary time series. By decomposing a time series into time–frequency space, one is able to determine both the dominant modes of variability and how those modes vary in time. Two level PCA combines the ability of PCA to extract the crosscorrelation or relationship between the variables [4], with that of orthonormal wavelets to separate deterministic features from stochastic processes and approximately de correlate the autocorrelation among the measurements. The Two level PCA approach is analogous to multi block PCA (Wold et al., 1996), with the sub-blocks being defined by the wavelet coefficients at each scale, and the super-block by the selected scales together.

Here in this paper compression of EEG signals are done with the wavelet decomposition and Two level PCA by decomposing the signal into 6 levels. Comparative analysis of compressed EEG signal with single level PCA and Two level PCA is performed

This paper is organized as follows, Introduction about the paper is given in section I and PCA and Wavelet transforms are explained in section II. Section III deals with the Two level PCA proposed method description. Results and discussions are given in section IV. Finally conclusions are made in section V

II. PRINCIPAL COMPONENT ANALYSIS

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 $\boldsymbol{x}$ for which the covariance matrix is $\boldsymbol{\Sigma}$, the principal components (PCs) can be defined by

$$z = \mathbf{A}x$$

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 by 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.

The data matrix X is size of $m \times n$, where n is the SVR computed periodicity and m is the number of periods considered.

$$X(t) = [x_1(t), x_1(t), x_1(t), \ldots\ldots x_1(t)] \quad (2.1)$$

is the time ordered collection of the feature at all beats into a single matrix to which PCA can be applied. The means of the x_i are removed and the covariance matrix computed. Then covariance is defined as

$$\Sigma = \frac{1}{n}\left[X X^T\right] \quad (2.2)$$

Σ is an $m \times m$ square symmetric matrix, Eigen values (a_j) and corresponding eigenvectors (λ_j) will be calculated, In general, once eigenvectors are found from the covariance matrix, the next step is to order them by Eigen value, highest to lowest. This gives you the components in order of significance. The lesser Eigen values can be ignored; this will form the basis for compression. Principal components are ordered eigenvectors of the covariance matrix. The PCs were obtained using

$$z_j = a_j x \qquad j = 1, 2, \ldots\ldots n$$

The PCs are a linear transformation of the beats with transformation coefficients given by the eigenvectors a_j. It is the eigenvectors which provide the surrogate respiratory signal in our analysis.

PCA can be solved using two methods, one is using covariance matrix and other is using singular value decomposition (SVD).

Let X be an arbitrary $n \times m$ matrix and X^TX be a rank r, square, symmetric $m \times m$ matrix. $\{\hat{v}_1, \hat{v}_2, \hat{v}_3 \ldots\ldots \hat{v}_r\}$ is the set of orthonormal m×1 eigenvectors with associated Eigen values for $\{\lambda_1, \lambda_2, \lambda_3, \ldots\ldots \lambda_r\}$ the symmetric matrix $X^T X$.

$$(X^T X)\hat{v}_i = \lambda_i \hat{v}_i \quad (2.3)$$

$\sigma_i = \sqrt{\lambda_i}$ are positive real and termed the singular values

$\{\hat{u}_1, \hat{u}_2, \hat{u}_3 \ldots\ldots \hat{u}_r\}$ is the set of n×1 vectors defined by

$$\hat{u}_i = \frac{1}{\sigma_i}\left[X \hat{v}_i\right]$$

$$\hat{u}_i \, \hat{u}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad \text{Eigenvectors are orthonormal.} \quad (2.4)$$

$$\| X \hat{v}_i \| = \sigma_i$$

The scalar version of singular value decomposition is

$$X \hat{v}_i = \sigma_i \hat{u}_i$$

X multiplied by an eigenvector of X^TX is equal to a scalar times another vector. The set of eigenvectors $\{\hat{v}_1, \hat{v}_2, \hat{v}_3 \ldots\ldots \hat{v}_r\}$ and the set of vectors are $\{\hat{u}_1, \hat{u}_2, \hat{u}_3 \ldots\ldots \hat{u}_r\}$ both orthonormal sets and bases in r dimensional space.

$$\Sigma = \begin{pmatrix} \sigma_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ & \sigma_{\tilde{\gamma}} & \\ 0 & \cdots & 0 \end{pmatrix} \quad (2.5)$$

$\sigma_1 \geq \sigma_2 \geq \sigma_3 \ldots \geq \sigma_{\tilde{\gamma}}$ are the rank-ordered set of singular values. Likewise we construct accompanying orthogonal matrices,

$$V = \left[\hat{v}_1, \hat{v}_2, \hat{v}_3 \ldots\ldots \hat{v}_r\right] \quad (2.6)$$

$$U = \left[\hat{u}_1, \hat{u}_2, \hat{u}_3 \ldots\ldots \hat{u}_r\right] \quad (2.7)$$

Matrix version of SVD

$$XV = U\,\Sigma \quad (2.8)$$

where each column of V and U perform the scalar version of the decomposition (Equation 3). Because V is orthogonal, we can multiply both sides by $V^{-1} = V^T$ to arrive at the final form of the decomposition.

$$XV = U\,\Sigma V^T$$

III. WAVELET TRANSFORM

A wavelet transform is the representation of a function by wavelets. The wavelets are scaled and translated copies (known as "daughter wavelets") of a finite-length or fast-decaying oscillating waveform (known as the "mother wavelet"). Wavelet transforms have advantages over traditional Fourier transforms for representing functions that have discontinuities and sharp peaks, and for accurately deconstructing and reconstructing finite, non-periodic and/or non-stationary signals.

Wavelet Transform can be represented as a linear transformation i.e. $Y = WX$, where X, Y are input and output of the transformation and W is orthogonal mother wavelet transformation matrix. Mother wavelet is defined as

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right) \quad (3.1)$$

Wavelets are oscillating functions of time that must satisfy several conditions: A wavelet ψ has zero time average and unit energy corresponds to ortho normality property of wavelets. The amplitudes of a wavelet have large fluctuations within a designated time period and extremely small values outside of that time while being band-limited in terms of their frequency content. The CWT of a signal $f(t)$ can be calculated using equation

$$F(u,s) = \int f(t)\frac{1}{\sqrt{s}}\psi^*\left(\frac{t-u}{s}\right)dt \quad (3.2)$$

By varying the values for *s* and *u* results in an infinite number of combinations, can be used to decompose the signal *f (t)*. Here *u* and *s* are the translation and dilation respectively.

A much more computationally efficient approach is the use of the discrete wavelet transform (DWT), which was developed by Mallat. Knowing only the values of the DWT coefficients, the waveform can be perfectly reconstructed. All of the extra coefficients of the CWT create a redundancy in calculation because they are highly correlated with the ones of the DWT. In implementation, the DWT performs even better because waveforms are already digitally sampled and have finite duration so the number of coefficients is limited DWT or CWT can be seen as a number on the time scale plane representing the correlation between the *signal* vector and the wavelet function at a given time-scale point. The DWT produces as many wavelet coefficients as there are samples in the original signal by using a filter scheme shown in Fig. 2.

The original 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 approximation and detail coefficients each with half the number of points as the original signal.

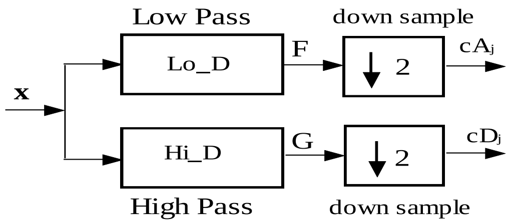


Figure 1. Wavelet Decomposition

The EEG signal will be corrupted by additive white noise during the process of signal acquisition. The corrupted EEG signal (observed) is given as

$$y = y_i = s_i + v_i \qquad i = 1............n. \qquad (3.3)$$

Where s_i is original EEG signal, v_i represents the independently and identically distributed random variable representing the amplitudes of the white Gaussian noise with $N(0, \sigma^2)$ here the problem is to remove or attenuate the maximum no. of v_i from the output signal 'y'.

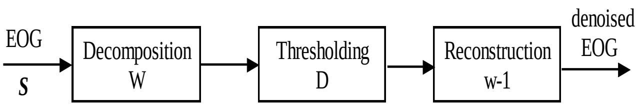


Figure 2. Wavelet denoising

Noisy corrupted signal is decomposed into 4 levels with all the mother wavelets. The sub bands thus formed contains the frequencies in the bands of 0-10 Hz, 10-20 Hz, 20-40 Hz and 40-80 Hz. These sub bands contains almost all the energy contained by the signal. Since the mother wavelet resembles the signal's' and large coefficients are generated corresponding to the eye moments and low coefficients corresponding to the noise. There are different thresholding methods such as soft and hard thresholding.

- Soft thresholding:

$$\begin{aligned} y &= \operatorname{sgn}(x).(|x| - t) && |x| \geq th \\ &= 0 && |x| \leq th \end{aligned} \qquad (3.4)$$

- Hard thresholding:

$$\begin{aligned} y &= x && |x| \geq th \\ &= 0 && |x| \leq th \end{aligned} \qquad (3.5)$$

The threshold is defined as $th = \sqrt{2\sigma^2 \ln(N)}$ where σ^2 is the variance of the signal, N is the total number of samples and x & y are the wavelet coefficients before and after threshold respectively.

Proposed Method

Two level PCA

Two level PCA 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 ´ n orthonormal matrix representing the orthonormal wavelet transformation operator containing the filter coefficients,

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 nXn orthonormal matrix representing the orthonormal wavelet transformation operator containing the

$$W = \begin{bmatrix} h_{L,1} & h_{L,2} & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & h_{L,N} \\ g_{L,1} & g_{L,2} & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & g_{L,N} \\ g_{L-1,1} & \cdot & \cdot & \cdot & g_{L-1,\frac{N}{2}} & & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & & g_{L-1,\frac{N}{2}+1} & \cdot & \cdot & \cdot & g_{L-1,N} \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot \\ g_{1,1} & g_{1,2} & 0 & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & & \cdot & \cdot & 0 & g_{1,N-1} & g_{1,N} \end{bmatrix} = \begin{bmatrix} H_L \\ G_L \\ G_{L-1} \\ \cdot \\ \cdot \\ G_m \\ \cdot \\ \cdot \\ \cdot \\ G_1 \end{bmatrix}$$

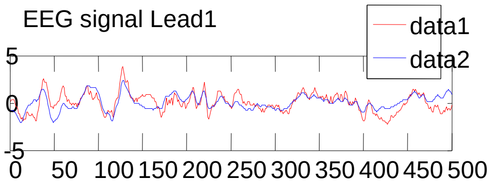
- (4.1)

Where, Gm is the 2log2nxm n matrix containing wavelet filter coefficients corresponding to scale m = 1, 2, ..., L, and HL is the matrix of scaling function filter coefficients at the coarsest scale. The matrix, WX is of the same size as the original data matrix, X, but due to the wavelet decomposition, the deterministic component in each variable in X is concentrated in a relatively small number of coefficients in WX, while the stochastic component in each variable is approximately de correlated in WX, and is spread over all components according to its power spectrum.

Artifacts are the major problems for analysis of the EEG signal and it makes the diagnosis more complex. Many methods are proposed to remove artifacts till now.

In this paper we proposed a method called Two level PCA to compress the EEG data for better feature extraction as well as for estimation of RMSE of EEG signals. The compression results using single level PCA and Two level PCA were tabulated in Table.1for two different data sets, which are taken from the references' [8] and [9].and the noisy and compressed EEG signals of two data sets are shown in the Figure 3.

Data1 corresponds to noisy EEG signal and Data2 corresponds to compressed EEG signal.



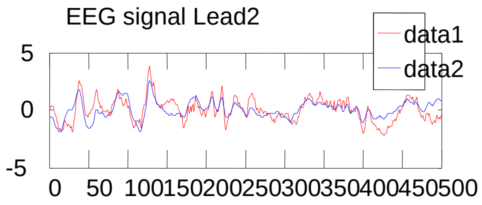


Figure.3. Noisy and Compressed EEG signal of two data sets

IV. Comparative Analysis

Table 1 Performance Results

S.No	EEG DATA SET1	EEG DATA SET 2
RMSE Using single level PCA	22.190493345 8124%	73.949742800 8459%
RMSE Using Two level PCA	99.9981%	99.9991%.

From Table.1, it is clear that RMSE is approaching near to 100% using a two level PCA than single level PCA for the two EEG data sets which is very much desirable feature in estimating the SNR of an EEG signal and also for feature extraction.

V. Conclusions

The results reveal the superiority of the proposed method in the processing of EEG signals for better compression of EEG signal. This approach is very helpful for analyzing the non stationary signals as it is designed with the advantages of wavelets and principal component analysis which are mostly uses for the same in general.

References

[1] **F. S. Tyner, J. R. Knott, and W. B. Mayer, Fundemantals of EEG Technology. Vol. 1. Basic Concepts and Methods. New York: Raven Press, 1983.**

[2] **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.**

[3] **M. Aminghafari, N. cheze, J.M Poggi,"Multivariate denoisious using wavlets and principal component Analysis," *computational statistics & Data Analsis,* vol. 50 pp. 2381-2398, 2006.**

[4] **B. Bakshi," Muliscale PCA with application to MSPC monitoring", *AlchE J*, vol. 44 pp 1596-1610, 1998.**

[5] **S. Ventakaramanan, P. Prabhat, S.R Choudhury, H.B Nemade, J.S. Sahambi, "Biomedical Instrumentation Based On Electrooculogram (EOG) Signal Processing And Application To A Hospital Alarm System", Indian Institute Of Technology (IIT) Gunawati, *Proceeding of IEEE ICISEP 2000*, pp.535-539.**

[6] **J. V. Basmajian and C. J. De Luca, Muscles Alive. Their Functions Revealed by Electromyography. Baltimore: illiams &: Wilkins, 1985.**

[7] **J. S. Barlow, "Artefact processing (rejection and minimization) in EEG data processing," in Handbook of Electroencephalography and Clinical Electrophysiology: Clinical Applications of Computer Analysis of EEG and Other Neurophysiological Signals (F. H. Lopes da Silva, W. Storm van Leeuwen, and A. Rmond, eds.), ch. 1, pp. 15- 62, Elsevier, 1986.**

[8] **http://sisec2010.wiki.irisa.fr/tiki-index.php?page=Artifact+removal+in+EEG+data.**

[9] **http://physionet.ph.biu.ac.il/physiobank/databas.**