Denoising of Motor Imagery Signals using Wavelet Transform and Singular Spectrum Analysis

Poonam Sheoran, Department of Biomedical Engineering, Deenbandhu Chhotu Ram University of Sc. & Tech., Murthal, Sonepat, India, poonam.bme@dcrustm.org

***Abstract*—Motor Imagery (MI) signals play a fundamental role in designing brain computer interfacing (BCI) systems. Acquisition of MI electroencephalogram (EEG) signals is based on the visual instructions given to the subjects, that result in electrooculogram (EOG) based artifacts which degrade the quality of EEG signals. Automatic removal of EOG artifacts with preservation of relevant neural activity will help in effective feature extraction and classification for BCI applications. In this paper, we present the design of an automatic EOG artifact removing algorithm based on discrete wavelet transform (DWT) and singular spectrum analysis (SSA). Canonical correlation analysis (CCA) is used as blind source separation technique on wavelet coefficients. The uncorrelated coefficients thus obtained are used for elimination of noisy components using SSA. The results in terms of signal-to-noise ratio (3.61), correlation coefficient (0.9223) and classifier's output substantiate the efficacy of the algorithm in comparison to other state-of-the-art methods.**

Keywords—Discrete Wavelet Transform, Canonical Correlation Analysis, Singular Spectrum Analysis, Electroculogram, Denoising.

I. INTRODUCTION

Motor Imagery EEG signals are the induced electrical activities on the scalp or motor cortex in response to imagined movements [1]. Standard methods of designing BCI systems are based on MI signals. BCI connects brain to the external devices for communication and control [2]. Acquisition of MI signals is done by instructing the subjects to imagine the movement of specified body parts as per the visual cues being run on the computer screen. In this process, the prominent eyeballs' movements are involved which lead to interference of EOG components in motor imagery EEG signals [3]. So, to develop an effective BCI system, it is quite important to remove the EOG components from EEG signals.

As per literature, time-frequency decomposition of EEG signals aid in dimensionality reduction, obtaining relevant frequency bands and noise reduction as well. The commonly used time-frequency decomposition techniques are DWT, empirical mode decomposition and variational mode decomposition [4]. The empirical mode decomposition techniques suffer from mode mixing while variational mode decomposition (VMD) is a computationally taxing technique [5] [6]. DWT is a method that provides different frequency bands; hence, it has been used by researchers to identify the ocular artifact (OA) region. Further, DWT doesn't rely upon the reference OAs or visual inspection and thus, it is the preferred frequency decomposition technique for EEG signals [7]. For the purpose of signal denoising, signal source separation is an important task, as without unmixing of sources, noise separation and removal is very difficult. For that, blind source separation (BSS) techniques have widely been used for EEG denoising [8]. Principal component analysis (PCA), independent component analysis (ICA) and CCA are the major BSS techniques. PCA takes into account the orthogonality assumption that sometimes is inconsistent with the EEG data practically, and ICA being dependent on higher order statistics, proves computationally complex [9], while CCA, a method based on second order statistics, is fast and computationally simpler [10].

A host of other EEG denoising techniques have been proposed by researchers based on DWT with adaptive noise cancellor [7], ICA with DWT [11], minimum noise estimation filter [12], Hankel matrix based denoising [13], CCA and ensemble empirical mode decomposition (EEMD) based method [9], wavelet neural network based approach [14], non-negative matrix factorization for artifact rejection [15] etc. These methods either rely upon removal or rejection of intrinsic mode functions (IMFs) and independent components (ICs) or thresholding of decomposed components. Rejection of components may lead to loss of relevant underlying neural information and thresholding doesn't suffice for complete removal of noisy component. Thus, there is a need of such an algorithm that not only preserves the relevant neural information but also removes the noisy component sufficiently enough to extract good features.

Considering these issues, we propose a novel algorithm that makes use of DWT for multiresolution analysis, CCA for BSS and SSA for identification and rejection of identified noisy components. The paper is organized as follows: 'Materials and Method' section provides details about data used and the techniques implemented. 'Results and

 DOI : 10.35291/2454-9150.2019.0114

Discussion' section presents the results obtained using this algorithm and finally, the paper is concluded in 'Conclusion' section.

II. MATERIALS AND METHODS

A. EEG Data used

In this paper, the BCI data 1 from the publicly available datasets of BCI Competition IV consisting of datasets for seven healthy subjects is used. The motor imagery data of two classes is selected for each subject, containing 100 trials each. The sampling rate chosen by data providers is 100 Hz. The most relevant 15 motor imagery channels were chosen to conduct this study: F3, F1, FZ, F2, F4, C5, C3, C1, CZ, C2, C4, C6, P1, PZ and P2. For more details, please refer [16].

B. Discrete Wavelet Transform

DWT is a multiresolution method that decomposes the signal into different frequency bands. Using DWT, OAs can be found concentrated in the low-frequency band, so, it contributes in identification of noisy frequency band. It provides a time-frequency representation of the signal and hence, quite effective in analysis of non-stationary signals. Wavelet transform uses a variety of wavelet functions possessing different properties. Wavelet transform makes use of different basis functions with different properties which are not available in fourier transform. The mother wavelet functions expand, contract and shift over the length of the signal based on the frequency content of the signal and provide wavelet coefficients [14]. The signal is decomposed in lower and higher frequency bands at each level using low pass and high pass filter banks. The coefficients generated from high pass filter are called detailed coefficients and that from low pass filter are called approximate coefficients [17]. For the present work, we decompose the EEG signals using daubechies wavelet upto level 4. The decomposition scheme is presented as follows in Fig. 1:

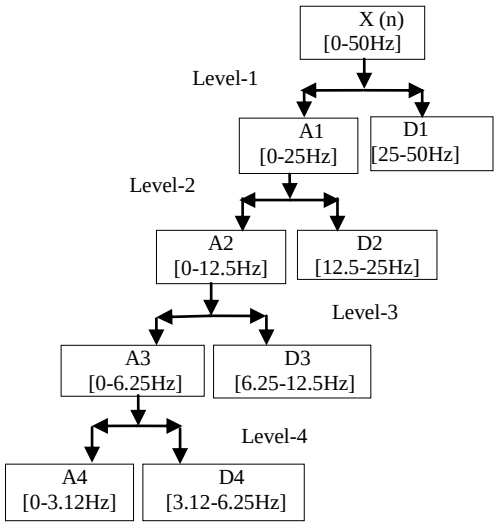

Fig. 1 Decomposition scheme of EEG signals.

C. CCA as BSS technique

The decomposed EEG signals (wavelet coefficients) with $t = 1,2,...N$ samples (number of samples are different for each coefficient) and k number of channels are represented as:

$$X(t) = [x_1(t), x_2(t), ..., x_k(t)]^T \quad (1)$$

$X(t)$ consists of uncorrelated unknown source signals $S(t) = [s_1(t), s_2(t), ..., s_k(t)]$ which are related to $X(t)$ using unknown mixing matrix, W, as:

$$X(t) = WS(t) \quad (2)$$

Now, CCA as a blind source separation (BSS) technique aims to obtain $S(t)$ by finding out W by using:

$$S(t) = W^{-1}X(t) \quad (3)$$

where, W^{-1} is called demixing matrix. The decomposed source signals computed using CCA are highly autocorrelated but mutually uncorrelated [9].

Consider $Y(t)$ as the delayed version of $X(t)$:

$$Y(t) = X(t-1) \quad (4)$$

Remove the mean of each row from the data matrices $X(t)$ and $Y(t)$ for data centralization. The linear combination in $X(t)$ and $Y(t)$ leads to following equations:

$$U = a_1x_1 + a_2x_2 + ... + a_kx_k = A^TX \quad (5)$$
$$V = b_1y_1 + b_2y_2 + ... + b_ky_k = B^TY$$

CCA intends to find the weighting vectors A and B that lead to maximization of correlation ρ between the variates U and V [10]:

$$\rho = \frac{A^TC_{xy}B}{\sqrt{(A^TC_{xx}A)(B^TC_{yy}B)}} \quad (6)$$

where, C_{xy} is the cross covariance of $X(t)$ and $Y(t)$ and C_{xx} and C_{yy} are the auto covariances of $X(t)$ and $Y(t)$. ρ maximization can be obtained by setting the derivatives of above equation to zero with respect to A and B. As X and Y contain almost same data, likewise, A and B too have same data. So, we only need to calculate A which is used to separate the observed signal into maximally autocorrelated and mutually uncorrelated source signals. Further, it is an estimation of W^{-1} which is used to get $S(t)$ using equation (3). Now, these uncorrelated source components are analysed for artifactual components by calculating their kurtosis values which signify the peaks of the signal [9]. To calculate the kurtosis values, following formula is used:

$$k_i = m_4 - 3m_2^2 \quad (7)$$

where, m_4 and m_2 are the 4th and 2nd central moments. There are m number of blink artifactual components having high kurtosis values which are used for singular spectrum analysis.

D. Singular Spectrum Analysis

Singular spectrum analysis (SSA) method is quite helpful in identifying and extracting of desired components automatically from the noisy signal. The SSA approach includes two complementary stages; decomposition and reconstruction. These stages further encompass two more stages. The decomposition stage involves embedding followed by singular value decomposition (SVD). Second

 DOI : 10.35291/2454-9150.2019.0114

stage consists of grouping and diagonal averaging to reconstruct the signal [18].

1) Decomposition

In univariate SSA, each of the time-series (uncorrelated wavelet coefficients) of $U = [u_1, u_2, , u_k]$ is first mapped in the form of trajectory matrix $X = [x_1, x_2, .., x_n]$ such that $x_i = [U_i, U_{i-2}, ..., U_{i+l-1}]^T$ where l is the window length. This trajectory matrix is called Hankel matrix whose diagonal elements are equal. This trajectory matrix or phase space reconstruction matrix is decomposed using singular value decomposition (SVD) into eigen subspaces. For that, compute the covariance matrix $C_x = XX^T$ and find the eigenvalues. Projecting the time-series onto each eigenvector gives the respective temporal principal component.

2) Reconstruction

Before reconstruction, reject the eigenvectors corresponding to the higher principal component due to the consideration that it carries a noisy fraction of the signal. For reconstruction, the elementary matrix X is split into many groups and then the matrices within each group are summed [19] as:

$$X = \sum_{g=1}^{g_t} \overline{X_g} \qquad (8)$$

where, $\overline{X_g}$ is the sum of elementary matrices within the group g, and g_t is the total number of groups. $\overline{U} = [\overline{U_1} \ \ \overline{U_2} ... \ \ \overline{U_s}]$ represents the reconstructed matrix of time-series with length s.

E. Proposed Methodology

Considering the pitfalls of various approaches proposed earlier for EEG denoising, we propose a new method that gives superior performance and is based on DWT and SSA which is made explicit in the form of flow chart of Fig. 2:

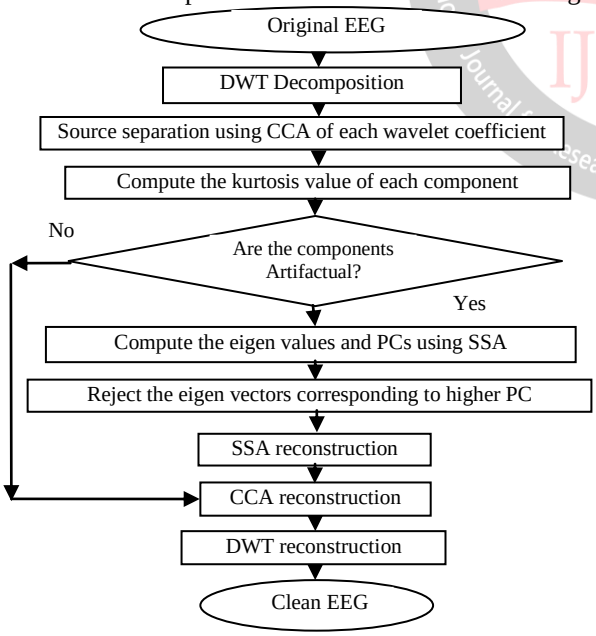

Fig. 2 Flowchart of proposed methodology

III. RESULTS AND DISCUSSION

BCI data set selected for this work has prominent EOG components. Using EEGLAB software, we plotted the EEG data by random selection and the same representing EOG activities is presented in Fig. 3. After wavelet decomposition, coefficients obtained are plotted and shown in Fig. 4 which represents the distribution of EOG components among the coefficients.

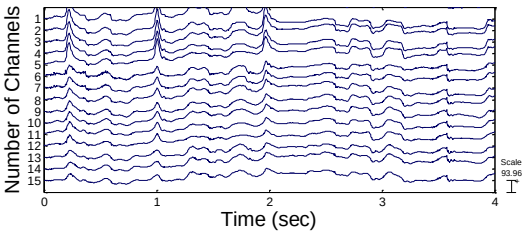

Fig. 3 Plot of original EEG contaminated with EOG

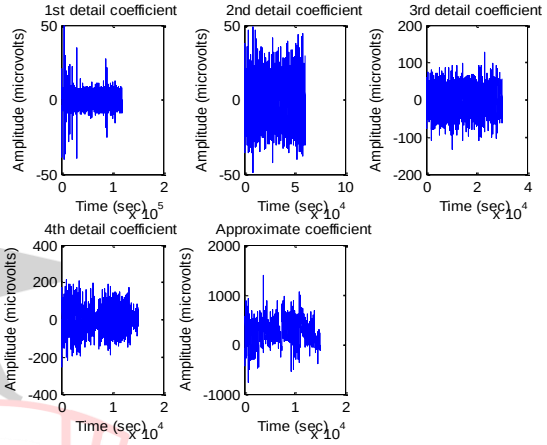

Fig. 4 Plot of wavelet coefficients of original EEG

The coefficients of fifteen channels are then uncorrelated using CCA. Observed that 4th detail coefficient and approximate coefficient have a higher range of amplitude (beyond the values of normal EEG), hence, are the main repository of EOG components. It can also be visualized in the plot of 4th detail coefficient after blind source separation presented in Fig. 5.

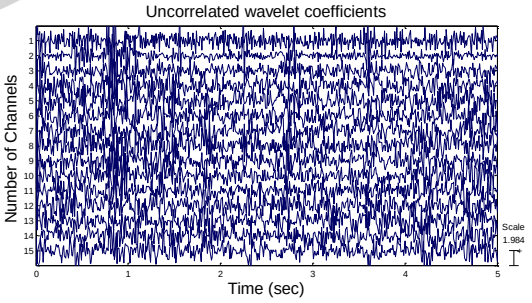

Fig. 5 Plot of 4th detail coefficients after BSS

Observe in Fig. 5 that the 4th detailed coefficients are effectively contaminated by EOG components. Further, the same is verified by computing the kurtosis values (as EOG activities have peaky distribution with high positive kurtosis). The components with kurtosis values more than the threshold are assumed to have noisy components and thus, the same are used for application of SSA for noise rejection. Every artifactual component is decomposed using SSA. After rejecting the principal components corresponding to highest

 DOI : 10.35291/2454-9150.2019.0114

eigen values, SSA reconstruction is performed. Thereafter, CCA reconstruction is done to get the clean wavelet coefficients. To get the clean EEG signals, wavelet reconstruction is done and the so obtained denoised EEG signals are shown in Fig. 6 in comparison with the original EEG signals.

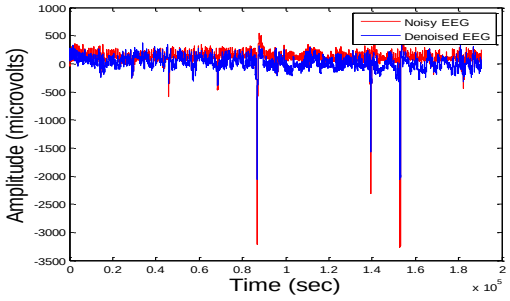

Fig.6. Comparative plot of noisy and denoised EEG signals

Further, mathematical analysis of denoising is also done by computing signal to noise ratio (SNR) and values of correlation coefficients. In the selected BCI dataset, the datasets of seven subjects possess both calibration and evaluation datasets both of which are used for denoising. Results in terms of SNR and correlation coefficient values are presented in Table I:

TABLE I. SNR and Correlation Coefficient values

Subjects	*SNR Values*	*Correlation Coefficient Values*
A-calibration	3.09	0.7680
B-calibration	3.36	0.8875
C-calibration	3.90	0.9345
D-calibration	3.78	0.9021
E-calibration	3.08	0.7548
F-calibration	3.88	0.9865
G-calibration	4.18	0.9808
A-evaluation	3.27	0.9265
B-evaluation	3.43	0.9345
C-evaluation	4.07	0.9887
D-evaluation	3.55	0.9567
E-evaluation	3.32	0.9475
F-evaluation	3.76	0.9674
G-evaluation	3.90	0.9772
Average	**3.61**	**0.9223**

The denoised EEG so obtained is used for feature extraction. Features computed are sample entropy and energy of the signals. Sample entropy is a measure of the extent and size of change of new information in signal sequence. Thus acts as an important feature. Likewise, energy of the signal is also representative of extent of information composed in the signal. Further details about sample entropy can be had from [20]. Using feature values, the signals are classified to find out the extent of preservation of neural information after denoising using linear discriminant analysis (LDA). LDA is a commonly used classifier and easy to use. It discriminates the data on the basis of inter- and intra-class variances (further information may be had from [21]). Results (in terms of resubstitution error, i.e., misclassification error) are compared with recently proposed CCA-EEMD method [9] and presented in Table II which shows that the proposed method outperformed the CCA-EEMD method in terms of SNR and classification results.

TABLE II. Comparison of results obtained with proposed method with CCA-EEMD method [9]

Methods	*SNR Values*	*Classification results*
CCA-EEMD	3.59	88% (classification accuracy (CA))
Proposed Method	3.61	89.06% (CA) &10.94% (resubstitution error)

IV. CONCLUSION

In this paper, a new method based on DWT and SSA for rejection of EOG activity from EEG is proposed. Here, DWT is observed as an effective multiresolution analysis tool and CCA performed well for separating the original EEG into source separated uncorrelated signal components that helped in identifying the noisy source. Then, SSA is used to reject the identified noisy components of EEG. The noise reduction is assessed by computing SNR values that are found to be fairly good as compared to recent methods. SSA reduced the noise with relevant data preservation, as supported by the classifier's output. The results also supported the aim of methodology. The algorithm may be incorporated in real time EEG analysis in future.

REFERENCES

[1] M. Lotze and U. Halsband, "Motor imagery," *J. Physiol. Paris*, vol. 99, no. 4–6, pp. 386–395, 2006.

[2] S. Articles, "Brain – computer-interface research : Coming of age," *Clin. Neurophysiol.*, vol. 117, pp. 479–483, 2006.

[3] X. Li, C. Guan, H. Zhang, and K. K. Ang, "Discriminative Ocular Artifact Correction for Feature Learning in EEG Analysis," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 8, pp. 1906–1913, 2017.

[4] P. Ofner, G. Muller-Putz, C. Neuper, and C. Brunner, "Comparison of Feature Extraction Methods for Brain-Computer Interfaces," in *5th International BCI Conference, Graz, Austria*, 2011, no. MI, pp. 9–12.

[5] X. Chen, C. He, and H. Peng, "Removal of muscle artifacts from single-channel EEG based on ensemble empirical mode decomposition and multiset canonical correlation analysis," *J. Appl. Math.*, vol. 2014, 2014.

[6] S. Lahmiri, "Comparative study of ECG signal denoising by wavelet thresholding in empirical and variational mode decomposition domains," *Healthc. Technol. Lett.*, vol. 1, no. 3, pp. 104–109, 2014.

[7] H. Peng, B. Hu, Q. Shi, M. Ratcliffe, Q. Zhao, Y. Qi, and G. Gao, "Removal of ocular artifacts in EEG - An improved approach combining DWT and ANC for portable applications," *IEEE J. Biomed. Heal. Informatics*, vol. 17, no. 3, pp. 600–607, 2013.

[8] S. M. H. Ahmed Kareem Abdullah, Zhang Chau Zhu, Lian Siyao, "Blind Source Separation Techniques Based Eye Blink Rejection in EEG Signals," *Inf. Technol. J.*, vol. 13, no. 3, pp. 401–413, 2014.

 DOI : 10.35291/2454-9150.2019.0114

[9] B. Yang, T. Zhang, Y. Zhang, W. Liu, J. Wang, and K. Duan, "Removal of Electrooculogram Artifacts from Electroencephalogram Using Canonical Correlation Analysis with Ensemble Empirical Mode Decomposition," *Cognit. Comput.*, vol. 9, no. 5, pp. 626–633, 2017.

[10] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2583–2587, 2006.

[11] R. Mahajan and B. I. Morshed, "Unsupervised eye blink artifact denoising of EEG data with modified multiscale sample entropy, kurtosis, and wavelet-ICA," *IEEE J. Biomed. Heal. Informatics*, vol. 19, no. 1, pp. 158–165, 2015.

[12] R. Foodeh, A. Khorasani, V. Shalchyan, and M. R. Daliri, "Minimum Noise Estimate filter: a Novel Automated Artifacts Removal method for Field Potentials," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 8, pp. 1143–1152, 2016.

[13] N. Alharbi, "A novel approach for noise removal and distinction of EEG recordings," *Biomed. Signal Process. Control*, vol. 39, pp. 23–33, 2018.

[14] C. Burger and D. J. Van Den Heever, "Removal of EOG artefacts by combining wavelet neural network and independent component analysis," *Biomed. Signal Process. Control*, vol. 15, pp. 67–79, 2015.

[15] C. Damon, A. Liutkus, A. Gramfort, and S. Essid, "Non-negative matrix factorization for single-channel EEG artifact rejection," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 1177–1181, 2013.

[16] B. Blankertz, G. Dornhege, M. Krauledat, K. R. Müller, and G. Curio, "The non-invasive Berlin Brain-Computer Interface: Fast acquisition of effective performance in untrained subjects," *Neuroimage*, vol. 37, no. 2, pp. 539–550, 2007.

[17] P. Sheoran and J. S. Saini, "Epileptic seizure detection using PCA on wavelet subbands," in *Proceedings of the 5th International Conference on Confluence 2014: The Next Generation Information Technology Summit*, 2014.

[18] S. Mahvash Mohammadi, S. Kouchaki, M. Ghavami, and S. Sanei, "Improving time–frequency domain sleep EEG classification via singular spectrum analysis," *J. Neurosci. Methods*, vol. 273, pp. 96–106, 2016.

[19] A. Maddirala and R. A. Shaik, "Removal of EOG Artifacts from Single Channel EEG Signals using Combined Singular Spectrum Analysis and Adaptive Noise Canceler," *IEEE Sens. J.*, vol. PP, no. 99, pp. 1–1, 2016.

[20] Y. Jia, H. Gu, and Q. Luo, "Sample entropy reveals an age-related reduction in the complexity of dynamic brain," *Sci. Rep.*, vol. 7, no. 1, pp. 1–10, 2017.

[21] F. Lotte, M. Congedo, L. Anatole, F. Lotte, M. Congedo, and L. Anatole, "A review of classification algorithms for EEG-based brain-computer interfaces", *J. Neural Eng.*, vol. 4, pp. 1–13, 2007.

 DOI : 10.35291/2454-9150.2019.0114