6th Annual International IEEE EMBS Conference on Neural Engineering
San Diego, California, 6 - 8 November, 2013

Sample Entropy Enhanced Wavelet-ICA Denoising Technique for Eye Blink Artifact Removal from Scalp EEG Dataset*

Ruhi Mahajan, *Student Member, IEEE, EMBS* and Bashir I. Morshed, *Member, IEEE, EMBS*

Abstract—**Scalp Electroencephalogram (EEG) recordings are usually contaminated with a variety of artifacts, which can be removed by threshold-based classifiers, Principal Component Analysis, Independent Component Analysis (ICA), wavelet-based multi-resolution analysis, or higher order statistics. In this paper we propose Sample Entropy, a self-sufficient statistical measure to identify the eye blink related independent components and Haar wavelet decomposition to subsequently denoise these components. The proposed method identified the blink artifactual components with an accuracy of 88% in our pilot study (N=4). The results demonstrated the improved performance of eye-blink artifacts removal with the neural activity intact in terms of Mutual Information (1.27 / 0.318 / 1.15), Correlation coefficient (0.574 / 0.369/ 0.569), and Standard deviation ratio (0.559/ 0.375 / 0.551) in comparison to standard Zeroing-ICA and wavelet-ICA based techniques, respectively. Instead of human expertise intervention to identify the eye blink components after Extended Infomax ICA decomposition, the algorithm offers potential for automation. This algorithm also offers advantage of being computationally fast and inexpensive, and does not require additional Electrooculographic signals for referencing.**

I. Introduction

Brain activities recorded using the Electroencephalogram (EEG) is strongly affected by various artifacts such as ocular activity (eye blinking, fixations and saccades), muscle activity, power line interferences, or heart beat activity. It is important to efficiently suppress these artifacts so that a clean artifact free EEG data can be obtained for the analysis. Power line interferences of 60 Hz can be removed from the signal by using notch filters and the subject may also be asked to avoid excessive muscle movement during data acquisition. But it is not feasible for the subject to avoid the eye blinks and eye movement as it is a part of the natural biological phenomenon. It becomes very crucial therefore to effectively remove these ocular artifacts (whose amplitude can be 10 to 100 times of EEG signals) from the data before further analysis, especially from frontal lobe channels.

The traditional approach to remove the eye blink artifacts is to use the linear filters for certain frequency bands that belong to artifact range [1]. However, this leads to significant loss of neurological activity from recorded data, as there is always spectral overlap between neurological and artifactual phenomenon [2]. Another common practice for correcting the ocular artifacts (OA) is by using regression analysis [3]. The performance of the regression based algorithm needs the design of optimally good regression models and requires recording of the Electrooculographic (EOG) channel [4]. Other methods including Principal Component Analysis (PCA) [5], Independent component analysis (ICA) [6], wavelet based denoising [7-9], Wavelet enhanced ICA (wICA) [10-11] and wavelet with higher order statistics [12] have achieved varying degrees of success. Most of these methods require recording of EOG signal for reference thresholding and efficient ways to select the artifactual independent components. Jointly using statistical tools like Kurtosis, Data improbability, Linear trends, Spectral pattern with the independent component scalp maps [13] and Kurtosis with Renyi's entropy [14] have shown to identify the eye blink components, but these approaches are time consuming and complex. Dan-hua Zhu et al. used Sample Entropy (SampEn) method to efficiently identify the blink independent components (IC) [15], but in their paper they replaced the entire eye blink IC with zero and reconstructed the EEG signal using rest of the ICs. It has been pointed out by Castellanos *et al.* [10] that ICs corresponding to blink artifacts (obtained after ICA decomposition) might also have significant amount of persisting neural signals, and if the entire IC is replaced with zero, there might be considerable loss of cerebral activity along with the artifactual activity.

In this paper, we seek a novel method that uses SampEn to identify the blink artifact ICs and instead of replacing the entire ICs with zero, we further decompose the blink ICs with Haar wavelet and threshold only the wavelet coefficients corresponding to the artifactual activity to zero. Thus, we minimize the loss of neural activities persistent in the blink ICs. This paper discusses the proposed method in detail.

II. Sample Entropy & wICA

A. Sample Entropy (SampEn)

Sample Entropy is a useful nonlinear tool to find the complexity of a time series. It is used here to identify the blink components in the EEG time series on the basis of entropy. SampEn is expected to be small for the eye blink activity as their wave pattern is more regular and predictable, but high for other activities [15]. SampEn is computed using the algorithm given in [16], expressed as:

$$\mathrm{SampEn}(m, r) = \log\left(\frac{B_r^{\,m}}{B_r^{\,m+1}}\right) \qquad (1)$$

For our analysis, maximum length of epochs for matching templates, m = 2 and tolerance, r = 0.2 × SD (optimal selection of m and r, denoted in [17]); where SD is the standard deviation of the data vectors, and B is the counter to track the template matches within the tolerance value r.

*This work was partially supported by Strengthening Communities Initiative (SCI) Capacity Building Grant, 2012.
Ruhi Mahajan and Bashir I. Morshed are with Department of Electrical and Computer Engineering, The University of Memphis, Memphis, TN 38152 USA. (Phone: 901-678-3650; E-mail: rmhajan1@memphis.edu, bmorshed@memphis.edu; Web: http://esarp.memphis.edu).

B. Wavelet enhanced ICA (wICA)

In order to efficiently separate EEG signals to source components, we have applied Extended Infomax ICA algorithm that can decompose both sub-Gaussian and super-Gaussian distributions into its statistically independent components [18]. For denoising IC, Haar wavelet is used as the basis function to recover the neural activities persistent in the ICs. The wICA method improves the performance of the artifact suppression over the Zeroing ICA technique as the wavelet decomposition provides an optimal resolution both in time and frequency domains [10]. After thresholding the wavelet coefficients for the neural activities, the EEG signals are reconstructed using the scalp projections.

III. Proposed Method

A. Data Acquisition

EEG recordings are collected using the wireless 14-channel device EPOC headset (Emotiv, Eveleigh, NSW, Australia) at the sampling rate of 128 sps. Data is collected from four subjects (2 males and 2 females of the age between 22-30 years) with the electrode locations (as per the International 10-20 electrode placement system) at AF3, AF4, F3, F4, F7, F8, FC5, FC6, P3 (CMS), P4 (DRL), P7, P8, T7, T8, O1 and O2. Due to the poor recording (loose electrode connection) from two electrode locations, we have discarded those channels from all the datasets and evaluated our algorithm on the remaining 12-channel recordings. The recordings have been conducted in a closed room with low ambient light for 1 min 45 seconds. During the data acquisition, the subject was asked initially to close the eyes for 30 seconds (to set the baseline), then open eyes and blink 9 times with a gap of 5 seconds in between. At the end of the recording session, the subject was again instructed to close eyes for 30 seconds.

B. Data Processing Procedure

After pre-processing 12- channel EEG data to remove DC offset, ICA is applied to the data using MATLAB (MathWorks, California, US) and EEGLAB (MATLAB Toolbox). As a statistical measure to identify the artifactual components, Sample Entropy is computed for all the components. The blink artifacts are the typical outlier data and can be identified using the threshold calculated with 95% Confidence Interval (CI) for the mean. The expected eye blink artifactual components have lower entropy values than neural activity components. Hence, we only compute the lower limit of 95% CI for mean, which reduces computational complexity of the algorithm. Mathematically:

$$\text{Lower Limit of 95\% CI} = \bar{x} - \frac{s}{\sqrt{n}} \times t_{n-1} \tag{2}$$

where $\bar{x}$, s and n are the sample mean, sample standard deviation and sample size of ICs, respectively. The two-sided test value of t_{n-1} for 11 degrees of freedom at 0.05 level of significance is 2.201 (as the sample size for the ICs in this study is < 30,we have used Students t-distribution test instead of classical Z-test). All the ICs with their SampEn less than the lower threshold limit of 95% CI for mean are marked as blink artifact components. We have also verified the marked eye blink artifact ICs with visual inspection method based on their spectral and temporal properties. For example, in Fig. 1, the scalp map projection, component activity and the power spectrum plot for the IC5 clearly show that it is a blink IC. Similarly, we have found IC2, 3 and 4 accounts for the eye blink in Fig. 2. Among all of the datasets, we were able to effectively relate the visual inspection based marked ICs with algorithm based marked ICs except for two cases: Dataset 1: IC4 with SampEn - 0.8925 and Dataset 3: IC3 with SampEn - 0.6015. These ICs pass the visual inspection check for the blink artifact but were not identified by the algorithm. Thus, out of 16 blink related ICs in the four datasets, the proposed algorithm with SampEn independently detected 14 ICs correctly in comparison with traditional visual check for the ICs which needs human expertise for identification, obtaining an accuracy of 88%.

The step–wise procedure for the proposed algorithm is outlined below:

- Apply the Extended Infomax ICA algorithm to the EEG data and obtain the weight matrix W.
- Compute the Sample Entropy of each independent component.
- Mark the ICs for denoising if the SampEn is lower than the threshold computed using Eqn 2.
- Compute the discrete wavelet transform (DWT) of the selected ICs using Haar wavelet as the basis function.
- Threshold the wavelet coefficients to zero if they exceed the value K computed by:

  $$K = \sqrt{2\log N}\ \sigma \tag{3}$$

  where σ^2 = median (|W (d, b)|)/0.6745 (4)

  is the estimator of the magnitude of the neural wide band signal, W(d, b) is DWT of ICs and N is the length of data to be processed (wavelet coefficients above the threshold K, represents the artifactual activity and are required to be denoised) [10].
- Compute the Inverse wavelet transform of the thresholded coefficients to recompose components of neural activity only.

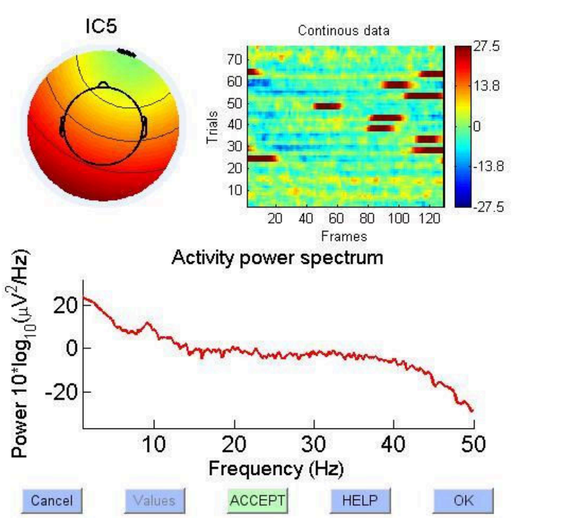

Figure 1. Scalp map, Component activity and power spectrum of IC5.

- Reconstruct the blink artifact free EEG signal by multiplying the Inverse of weight matrix with the inverse transformed wavelet coefficients.

IV. RESULTS

The preprocessed EEG signal of Dataset 1 from the prefrontal cortex location AF3 (for 78 sec duration) is plotted in Fig. 3 using MATLAB. Data labels for the amplitude of three eye blink activities at 24.22 sec, 43.99 sec and 64.41 sec are marked in the plot. The blink artifact clean EEG data from the same location using the proposed method is also plotted in Fig. 4 (due to space constraint, plots for other electrode locations are not included here). The corresponding data labels for the amplitude of selected eye blinks are also marked which shows the attenuation of the artifactual component in the data and persistent neural activity. The ICs marked for wavelet denoising are marked with the symbol + in Fig. 5. The lower limit threshold values for 95% CI for the four datasets (calculated using Eqn. 2) are found to be: 0.7732, 0.9884, 0.5237, and 0.7275. Different statistical metrics- Standard deviation ratio (signifies power of the signal affected with reconstruction), Correlation coefficient (how well the shape of the original signal is retained) and Mutual Information (MI) [19] are used to evaluate the performance of the proposed method over other traditional methods. The results are tabulated in Table I. Correlation coefficient and Standard deviation ratio have slightly improved with the proposed method. MI criterion also suggests that the reconstructed EEG signal (with the neural activities preserved) with the proposed method resembles better with the raw EEG signal compared to other methods.

V. DISCUSSION

Eye blinks causes one of the most common artifacts in the practical settings, and significantly affect some channels of EEG recordings (such as AF3). The proposed algorithm has shown tremendous potential to be suitable for detecting and suppressing blink related artifactual activities as shown with the pilot data. To attenuate other artifacts (eg. ocular activities including saccades and fixations, muscle activities, and heart beat activity), the algorithm needs to accordingly adjust the thresholding of the wavelet coefficients to remove those artifactual components.

TABLE I. PERFORMANCE ANALYSIS OF RECONSTRUCTED BLINK ARTIFACT FREE DATA AND RAW EEG DATA (FROM AF3 CHANNEL LOCATION) WITH THREE ALGORITHMS

Parameters	**Wavelet-ICA**	**Zeroing ICA**	**Proposed Method**
Dataset 1			
Standard Deviation ratio	0.5606	0.3265	0.5627
Mutual Information	1.3543	0.2148	1.4760
Correlation Coefficient	0.5832	0.3272	0.5897
Dataset 2			
Standard Deviation ratio	0.4617	0.3586	0.4676
Mutual Information	0.5750	0.3131	0.5887
Correlation Coefficient	0.4608	0.3550	0.4632
Dataset 3			
Standard Deviation ratio	0.6330	0.4428	0.6476
Mutual Information	1.5369	0.4274	1.7324
Correlation Coefficient	0.6638	0.4277	0.6704
Dataset 4			
Standard Deviation ratio	0.6036	0.1352	0.6064
Mutual Information	0.6748	0.0424	1.2672
Correlation Coefficient	0.6012	0.1387	0.6513

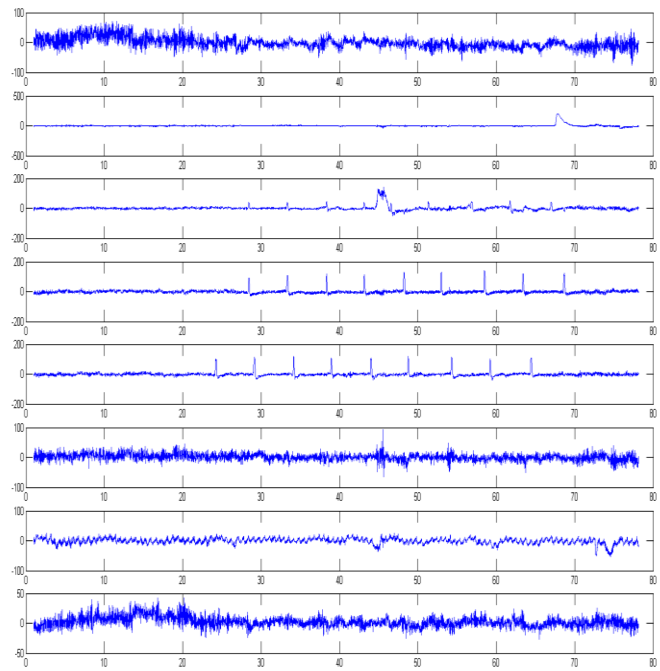

Figure 2. IC plots for the first eight components of the decomposed EEG data of Dataset 1 (X-axis is time in seconds and Y-axis is amplitude in microvolts). ICs 2,3,4 and 5 indicates the eye blink artifactual components.

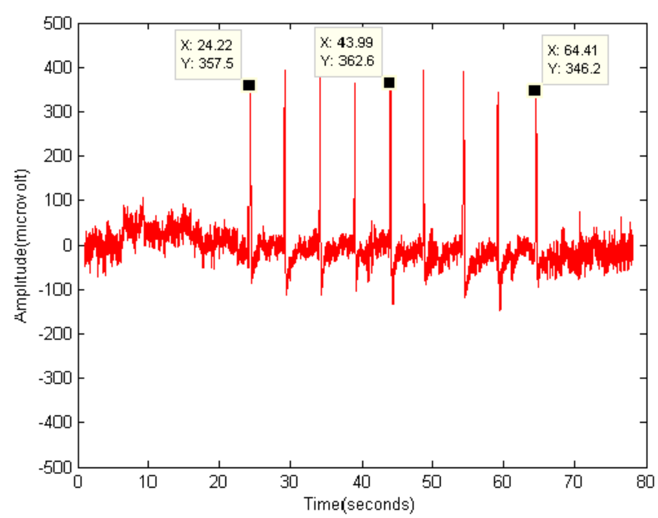

Figure 3. Raw EEG signal of Dataset 1 from prefrontal channel AF3 with 9 consecutive blinks with a gap of 5 seconds in between the blinks.

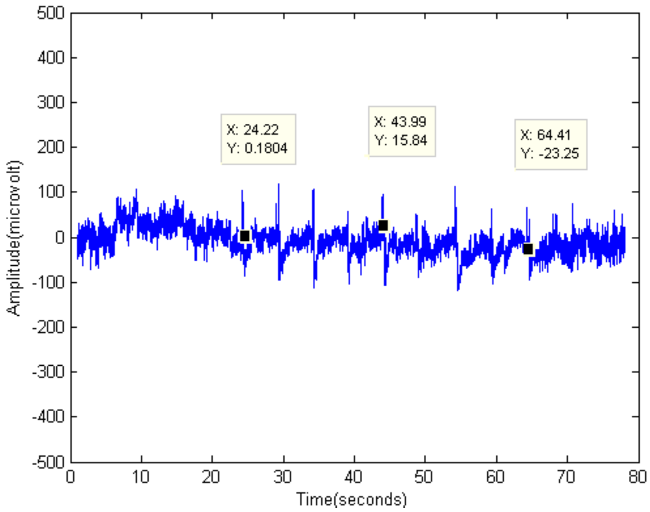

Figure 4. Reconstructed EEG signal from prefrontal channel location AF3 with Sample Entropy Enhanced Wavelet-ICA Denoising Technique.

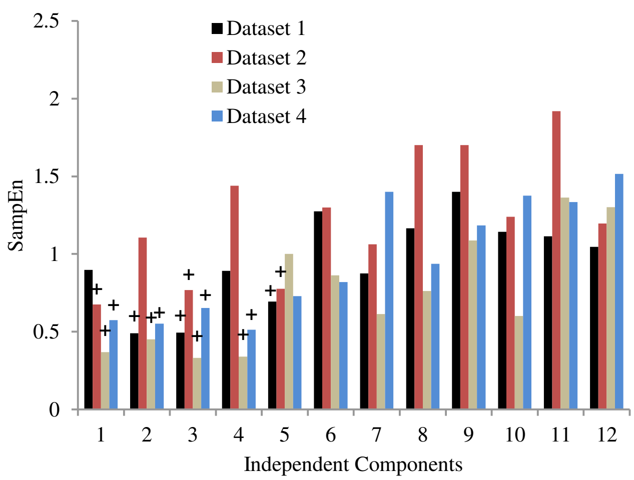

Figure 5. SampEn values of each IC for the four datasets of EEG data from AF3 channel location. ICs selected for denoising are marked with +.

VI. Conclusion

Among the existing algorithms, ICA is widely accepted as potentially the most robust tool. The challenge is to correctly and efficiently identify the blink related artifactual components. The proposed algorithm uses SampEn for identifying the blink artifactual independent components and applies Haar wavelets for denoising only the artifactual components. The algorithm can potentially be automated by adjustable threshold determined adaptively in real-time. Different statistical measures were used to evaluate the performance of proposed algorithm that shows the neuronal activity is maximally preserved after reconstructing EEG signals compared to other common algorithms (wICA and Zeroing ICA). The proposed method resulted in an accuracy of 88% with the pilot data for identifying the blink related ICs and the wavelet decompositions of only artifactual components steered to an efficient eye blink artifact removal technique. Removal of other artifacts is perceived to be feasible by applying similar technique.

References

[1] J. Gotman, D.R. Skuce, C.J.Thompson, P. Gloor, J.R. Ives, FW. Ray, "Clinical applications of spectral analysis and extraction of features from electroencephalograms with slow waves in adult patients," Electroencephalogr Clin Neurophysiol, vol. 35, pp. 225–35, 1973.

[2] N.A. de Beer, M. van de Velde, P.J. Cluitmans, "Clinical evaluation of a method for automatic detection and removal of artifacts in auditory evoked potential monitoring ," J Clin Monit, vol. 11, pp. 381–91, Nov. 1995.

[3] J.C. Woestenburg, M.N. Verbaten, J.L. Slangen, "The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain," Biol. Psychol., vol.16, pp. 127–147, March 1983.

[4] M. T Akhtar, W.Mitsuhashi and C.J. James, "Employing spatially constrained ICA and wavelet denoising, for automatic removal of artifacts from multichannel EEG data," Signal Processing, vol.92, pp.401-416, February 2012.

[5] T.D Lagerlund, F.W Sharbrough , N.E Busacker, "Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition," J Clin Neurophysiol, vol. 14, pp. 73–82, 1997.

[6] T.P Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, T.J. Sejnowski, "Analysis and visualization of single-trial event-related potentials," Hum Brain Mapp, vol. 14, pp.166–85, Nov. 2001.

[7] S.V.Ramanan, N.V. Kalpakam, J.S. Sahambi, "A novel wavelet based technique for detection and de-noising of ocular artifact in normal and epileptic electroencephalogram," International Conference on Communications, Circuits and Systems, vol.2, 2004, pp.1027-31.

[8] V. Krishnaveni, S. Jayaraman, S. Aravind, V. Hariharasudhan, K. Ramadoss, " Automatic Identification and Removal of Ocular Artifacts from EEG using Wavelet Transform," Measurement Science Review, vol.6, Section 2, 2006.

[9] P. Senthil Kumar, R. Arumuganathan1, K. Sivakumar , and C. Vimal, "Removal of Ocular Artifacts in the EEG through Wavelet Transform without using an EOG Reference Channel," Int. J. Open Problems Compt. Math., vol. 1, Dec. 2008.

[10] N.P. Castellanos and Valeri A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," Journal of Neuroscience Methods, vol.158, pp.300–312, Dec. 2006.

[11] M Zima, P Tichavsky, K Paul, and V Krajca, "Robust removal of short-duration artifacts in long neonatal EEG recordings using wavelet-enhanced ICA and adaptive combining of tentative reconstructions", Physiol Measurement, vol. 33, pp. N39-N49, 2012.

[12] G.Hosna and A.Erfanian, "A fully automatic ocular artifact suppression from EEG data using higher order statistics: Improved performance by wavelet analysis," Medical Engineering & Physics, vol. 32, pp. 720-729, September 2010.

[13] A. Delorme, T. Sejnowski, S. Makeig, "Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis," Neuroimage, vol.34, pp.1443–1449, 2007.

[14] A. Greco, N. Mammone, F. C. Morabito, M. Versaci, "Semi-Automatic Artifact Rejection Procedure based on Kurtosis, Renyi's Entropy and Independent Component Scalp Maps," International Enformatika Conference, Turkey,2005,pp. 22-26.

[15] D.H. Zhu, J.J Tong and Y. Chen, "An ICA-based Method for Automatic Eye Blink Artifact Correction in Multi-channel EEG," 5th International Conference on Information Technology and Application in Biomedicine, China, 2008, pp. 338-341.

[16] J.S. Richman, J. R. Moorman, "Physiological time-series analysis using approximate and sample entropy, " Am J Physiol Heart Circ Physiol, vol. 278, pp. 2039-2049, 2000.

[17] D.E. Lake, J.S. Richmann, M.P. Griffin, J.R. Moorman, "Sample entropy analysis of neonatal heart rate variability," Am. J. Physiol. Heart Circ. Physiol., vol.283, pp.R789–97, 2002.

[18] TW Lee, M. Girolami, T J. Sejnowski, "Independent Component Analysis using an Extended Infomax Algorithm for Mixed Sub-Gaussian and Super-Gaussian Sources," Neural Computation, 1999, vol. 11, no. 2, pp. 409-433.

[19] JW. Kelly, D P. Siewiorek, A. Smailagic, J. L. Collinger, D. J. Weber, W. Wang,"Fully Automated Reduction of Ocular Artifacts in High-Dimensional Neural Data," IEEE Transactions on Biomedical Engineering, vol.58, pp.598-606, 2011.