31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Focal Artifact Removal from Ongoing EEG – A Hybrid Approach Based on Spatially-Constrained ICA and Wavelet De-noising

Muhammad Tahir Akhtar*
The Education and Research Center for Frontier Science and Engineering, The University for Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, JAPAN
akhtar@ieee.org, akhtar@ice.uec.ac.jp

Christopher J. James
Signal Processing and Control Group, Institute of Sound and Vibration Research (ISVR), University of Southampton, Southampton, UK
C.James@soton.ac.uk

***Abstract*— Detecting artifacts produced in electroencephalographic (EEG) data by muscle activity, eye blinks and electrical noise, etc., is an important problem in EEG signal processing research. These artifacts must be corrected before further analysis because it renders subsequent analysis very error-prone. One solution is to reject the data segment if artifact is present during the observation interval, however, the rejected data segment could contain important information masked by the artifact. It has already been demonstrated that independent component analysis (ICA) can be an effective and applicable method for EEG de-noising. The goal of this paper is to propose a framework, based on ICA and wavelet denoising (WD), to improve the pre-processing of EEG signals. In particular we employ the concept of spatially-constrained ICA (SCICA) to extract artifact-only independent components (ICs) from the given EEG data, use WD to remove any brain activity from extracted artifacts, and finally project back the artifacts to be subtracted from EEG signals to get clean EEG data. The main advantage of the proposed approach is faster computation, as all ICs are not identified in the usual manner due to the square mixing assumption. Simulation results demonstrate the effectiveness of the proposed approach in removing focal artifacts that can be well separated by SCICA.**

I. INTRODUCTION

Ocular artifacts (OAs) (eye movements and eye blinks), muscle noise, heart signals, and line noise often produce large and distracting artifacts in electroencephalographic (EEG) recordings [1]. Rejecting EEG segments with artifacts offers an easy solution, however, the amount of data lost to artifact rejection may be unacceptable [2]. The EEG signals contain neural information below 100 Hz (in many applications the information lies below 30 Hz) [1], and conventional filtering methods can be used to remove, e.g., line noise and other high frequency components. The main problem is OAs and muscle artifacts that have a spectral overlap with the underlying EEG and cannot be removed using conventional filtering [3].

It has been shown by many researchers that independent component analysis (ICA) [4] can be used to efficiently separate the distinct artifactual processes from EEG data [5], [6]. In most of existing methods, after ICA, independent components (ICs) corresponding to artifacts are manually selected using visual inspection. The identified artifact ICs are rejected and remaining ICs are used to reconstruct clean EEG data. If some brain activity is leaked to artifact ICs, so rejecting these components results in loss of desired information. In order to solve this problem, and to have a fully automatic method, wavelet enhanced ICA (wICA) is proposed in [7]. The wICA combines ICA with wavelet denoising (WD), and makes use of wavelet thresholding for denoising of the demixed ICs. The thresholding allows conservation of the time-frequency structure of artifacts and recovering of the cerebral activity "leaked" into the components [7]. This method does not require manual identification of artifact ICs, as all ICs are wavelet denoised. However, extracting all ICs poses a problem mainly due to long processing time. In order to solve this problem we propose an approach based on spatially-constrained ICA (SCICA) [8], [9]. SCICA incorporates reference or constraint topographies in the ICA algorithms. This allows to extract only desired ICs - that is artifacts in our case. Since not all ICs are extracted, this approach is faster in computation. We then apply WD to artifact ICs to remove brain activity and get artifact-only signals. Finally these artifacts are projected back, and subtracted from, EEG data to get clean EEG signals.

The rest of the paper is organized as follows. Section II describes the proposed approach, and Section III presents some simulations results. Finally concluding remarks are given in Section IV.

II. PROPOSED APPROACH

The block diagram of the proposed approach for preprocessing of EEG data is shown in Fig. 1. As shown, EEG data is assumed to be generated according to ICA model [4] as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{v}(t), \tag{1}$$

where $\mathbf{x}(t) = [x_1(t), x_2(t), \cdots, x_M(t)]^T$ are a linear mixture of N sources $\mathbf{s}(t) = [s_1(t), s_2(t), \cdots, s_N(t)]^T$, $\mathbf{A}$ is $M \times N$ mixing matrix, and $\mathbf{v}(t) = [v_1(t), v_2(t), \cdots, v_M(t)]^T$ is additive noise at the EEG sensors. Here the number of sources N, their waveforms $s_i(t)$, and mixing matrix $\mathbf{A}$ are all unknown. For the sake of simplicity we consider the square mixing problem, i.e., $M = N$. The source signals $s_i(t)$ can be considered as being generated from different brain regions and artifacts. These artifacts mask the brain-activity information, and are harmful for further analysis and processing. Thus it is very important to process EEG data $\mathbf{x}(t)$ so that contribution of artifacts is removed, without affecting the brain-activity information, and is the main task of the work presented in this paper. As shown in Fig. 1., the proposed approach comprises following key steps:

- Preprocessing using conventional filtering.

*This work was carried out whilst Dr. Akhtar was visiting researcher at the Institute of Sound and Vibration Research(ISVR), University of Southampton, Southampton, UK, and supported by research fund of the the University of Electro-Communications, Chofu, Tokyo, JAPAN.

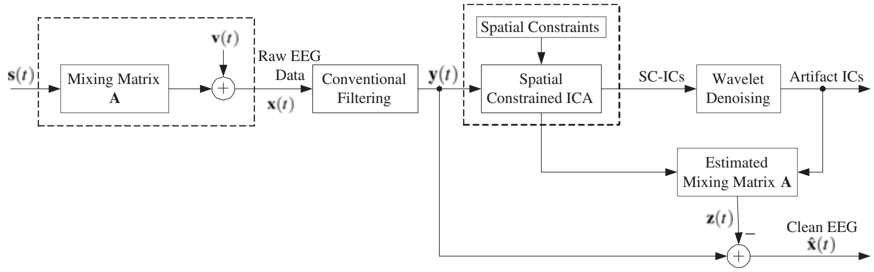

Fig. 1. The block diagram of proposed approach for preprocessing of EEG data.

- Apply SCICA to get SC-ICs representing artifacts in EEG data.
- Apply WD to remove any brain activity leaked to these artifact ICs.
- The extracted artifact-only signals are projected back, and subtracted from, EEG data to get clean EEG for further analysis and processing.

The objective of conventional filtering is to process raw EEG data $\mathbf{x}(t)$ to remove 50 Hz line noise, baseline values, artifacts occupying very low frequencies and high frequency sensor noise $\mathbf{v}(t)$, and this stage might include combination of various conventional notch, lowpass, and/or highpass filters.

A. Spatially-Constrained ICA (SCICA)

The key step in the proposed approach is the application of SCICA to get artifact ICs from filtered and baseline corrected EEG data $\mathbf{y}(t)$. Here we briefly describe SCICA, and for details, the reader is referred to [8], [9]. The main idea is to define a spatial constraint (SC) on the mixing matrix $\mathbf{A}$ to represent specific prior knowledge or prior assumptions regarding the spatial topography of some source sensor projections, i.e., the SC operates on selected columns of $\mathbf{A}$ and is enforced with reference to a set of predetermined constraint sensor projections, denoted by $\mathbf{A}_c$. Thus, the spatially constrained mixing matrix comprises two types of columns

$$\mathbf{A} = [\hat{\mathbf{A}}_c, \mathbf{A}_u], \tag{2}$$

where $\hat{\mathbf{A}}_c \approx \mathbf{A}_c$ are columns subject to the constraint, and $\mathbf{A}_u$ are otherwise unconstrained columns. Depending on the application, the predetermined sensor projections (or constraint topographies) could be obtained by manual selection of sources extracted from a previous data segment using conventional ICA methods or derived from the predictions of some mathematical model of the signal generating process (e.g., biophysical or physiological) under investigation [8]. Depending upon the confidence level about the accuracy of the constraint topographies $\mathbf{A}_c$, and the extent to which constrained columns $\hat{\mathbf{A}}_c$ may diverge from reference $\mathbf{A}_c$, there are three types of constraints: 1) *hard* constraints indicating fixed column, 2) *soft* constraints allowing divergence within a small angular threshold α, and 3) *weak* constraints that only provide an initial guess for otherwise unconstrained estimation. We consider the spatially-constrained-FastICA (SCFastICA) algorithm of [8] with soft SCs.

The SCFastICA algorithm seeks to maximize the statistical independence of the unconstrained sources while minimizing the divergence between the spatially constrained source sensor projections and their corresponding reference topographies. Since we are interested in only SC-ICs, we use a deflationary approach to extract only desired components, and thus the output of the SCFastICA algorithm is SC-ICs (which are artifact signals in our case), and an estimate of corresponding mixing matrix. This results in fast computational time, as compared with if all ICs are extracted.

B. Wavelet Denoising (WD) of SC-ICs

It is worth mentioning that SC-ICs extracted by SCFastICA are expected to correspond to artifacts only, however, some brain activity might leak to these extracted signals. Since artifacts have a frequency overlap with the brain signals, conventional filtering cannot be used, and hence we propose to use WD to remove any brain activity from extracted SC-ICs.

The discrete wavelet transform (DWT) analyzes a finite length time domain signal by breaking up the initial domain in two parts: the detail and approximation information [10]. The approximation domain is successively decomposed into detail and approximation domains. The basic principle is that the decomposition of a noisy signal on a wavelet basis (by DWT) have the property to "concentrate" the informative signal in few wavelet coefficients having large absolute values without modifying the noise random distribution. After transformation the noise coefficients have small values, inversely to the informative signal (normal or pathologic neural activity and artifacts). Therefore, denoising can be achieved by thresholding the wavelet coefficients [11]. We have implemented WD using the MATLAB wavelet tool box as follows: (the corresponding MATLAB `function` is given in parenthesis)

- Choosing the value of the threshold (`ddencmp`)
- Apply DWT (`swt`) to the SC-IC signal to obtain details and approximations
- Threshold (`wthresh`) the detailed components obtained in the previous step
- Apply inverse DWT (`iswt`) to obtain artifact-only signal

Once "clean" artifacts are obtained, these are projected back to EEG sensors by using mixing matrix $\mathbf{A}$ estimated by SCFastICA, and artifacts in the EEG data are obtained, as denoted by $\mathbf{z}(t)$ in Fig. 1. Finally, the clean EEG data $\hat{\mathbf{x}}(t)$ is obtained by subtracting artifacts $\mathbf{z}(t)$ from EEG data $\mathbf{y}(t)$.

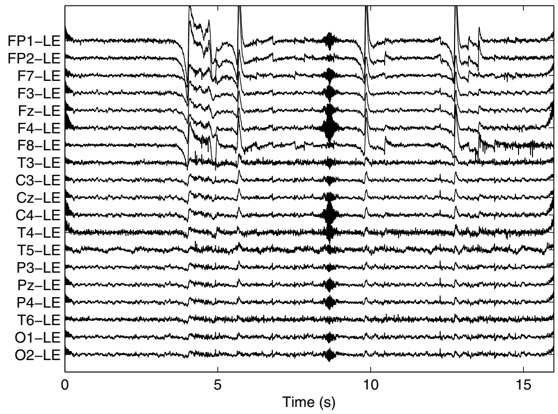

Fig. 2. Raw EEG data after preprocessing through conventional filtering.

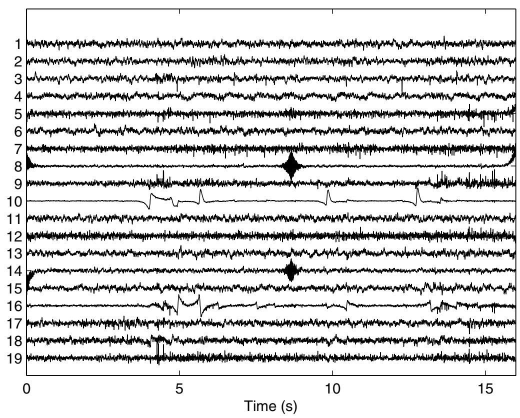

Fig. 3. Independent components (ICs) obtained by unconstrained FastICA algorithm.

A comparison of the proposed approach with the existing methods is given below:

- In [8], SCFastICA is proposed to extract all ICs including those corresponding to incorporated SCs. The SC-ICs are rejected before EEG data is reconstructed. This results in an almost automatic approach for artifact removal, however, rejecting SC-ICs also removes brain activity masked in these components, and might result in appreciable loss of desired information. In our approach, we exploit the concepts of SCICA and WD to reject artifacts, and get the clean EEG data. As compared with [8], less brain information is rejected and hence better EEG signals are extracted.
- The wICA proposed in [7] computes all ICs, and uses WD to reject artifacts. This results in improved artifact rejection, however, requires computation of all ICs which might pose problem of long computational time. In comparison with wICA [7], only a few ICs corresponding to artifacts (thanks to SCICA) are extracted and processed and hence a faster computation time is achieved. This is quite advantageous, especially in long term ongoing EEG recordings.

III. RESULTS AND DISCUSSION

Here we present some simulation results to verify the effectiveness of the proposed approach.

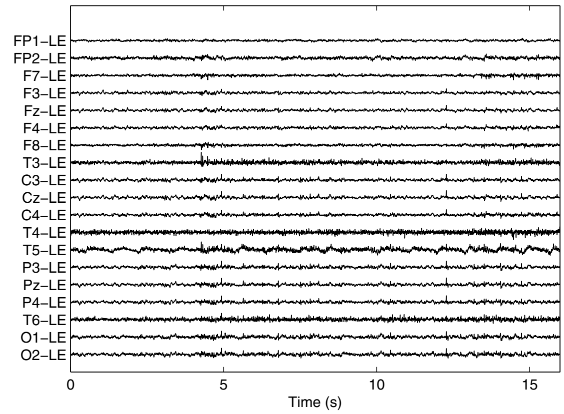

Fig. 4. Reconstructed EEG data after rejecting artifact ICs (as shown in Fig. 3, ICs 8, 10 , 14 and 16 mainly contain artifacts).

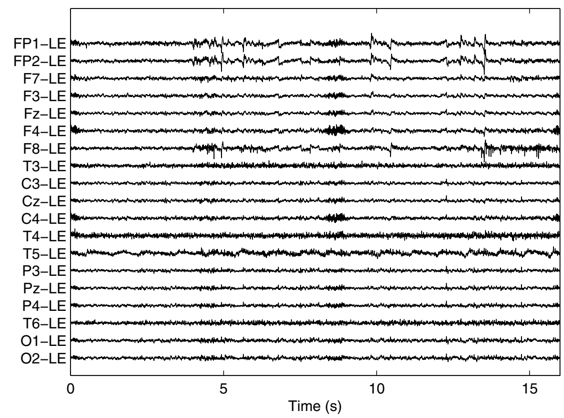

Fig. 5. Clean EEG data obtained by applying wavelet enhanced ICA (wICA) of [7].

Fig. 2 shows raw EEG data after being processed by conventional filtering. The result of unconstrained ICA using the FastICA algorithm is shown in Fig. 3. We see that, in this particular example, the FastICA algorithm can well separate the artifacts in ICs 8, 10, 14 and 16. A close observation reveals that these ICs contain some brain activity as well. Rejecting these ICs, and an artifact free EEG data can be reconstructed as shown in Fig. 4. Fig. 5 shows EEG data reconstructed by applying wICA [7], where ICs shown in Fig. 3 are denoised using DWT before EEG being reconstructed.

The results of the proposed approach are shown in Figs. 6 and 7. Here results of unconstrained FastICA are used to obtain an initial guess for spatial constraints to be used in SCFastICA algorithm [8]. It is worth mentioning that running unconstrained ICA is not the main part of the proposed approach, and is used merely to initialize the algorithm[1]. Fig. 6(a) shows the SC-ICs extracted by SCFastICA. It is evident that these SC-ICs correspond to artifact ICs give in Fig. 2. These extracted SC-ICs are wavelet denoised to remove any

[1] As stated earlier, SCICA incorporates predetermined sensor projections (or constraint topographies). In real EEG recordings, these constraints can be derived from the predictions of an appropriate mathematical model, or could be obtained by manual selection of sources extracted from previous data segment using conventional ICA [8]. In this paper, we have used simulated EEG data, and hence, unconstrained ICA to get the reference topographies to be used as constraints with SCICA.

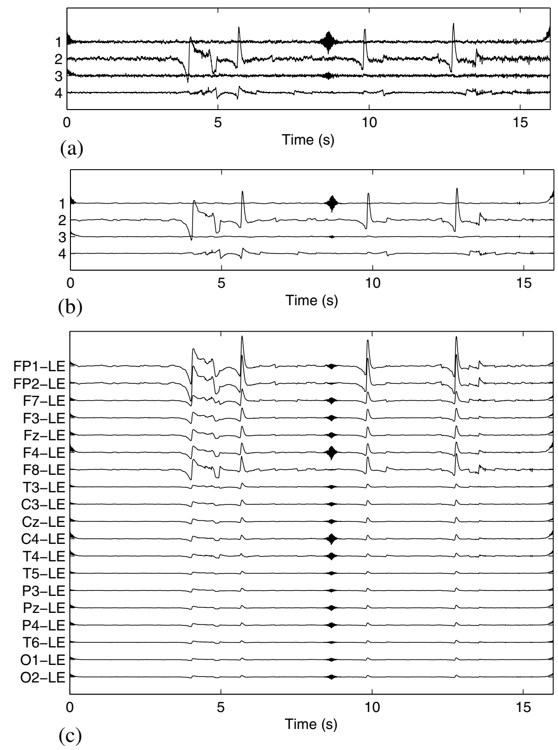

Fig. 6. Results obtained by spatially constrained Fast ICA. (a) Extracted SC-ICs corresponding to artifacts. (b) Wavelet denoised SC-ICs. (c) Artifacts projected to the scalp.

brain activity. A level 5 decomposition and `haar` wavelet is used with MATLAB function `swt` to compute stationary wavelet transform, where `ddencmp` is used to get the default threshold. The artifact-only signals as shown in Fig. 6(b). Finally these artifacts are projected back to the EEG sensors to get artifacts projected at the scalp. Subtracting artifacts from the EEG data gives the denoised EEG signals as shown in Fig. 7. In order to compare the performance of various methods, artifact free epoch are shown in Fig. 8, which shows that rejecting artifacts ICs results in loss of brain information, and gives a distorted version of EEG signal, whereas wICA and proposed approach are able to retain the brain information.

IV. CONCLUDING REMARKS

In this paper we have proposed a new approach for removing artifacts from multichannel EEG data. In contrast to existing approaches, where the emphasis is to reject artifacts directly, the proposed approach attempts to first identify the artifacts and then perform the denoising. The main advantage is computational efficiency, as less data is processed as compared with the existing approaches. Our experiments show that the proposed approach is advantageous in the case of focal artifacts that can be well localized by SCICA, and hence can be applied to removing ocular artifacts, for example.

In the future, it would be interesting to explore the removal of other types of artifacts. In this paper we have considered only simulated EEG data, and denoising long term ongoing EEG recordings is also a task for future work.

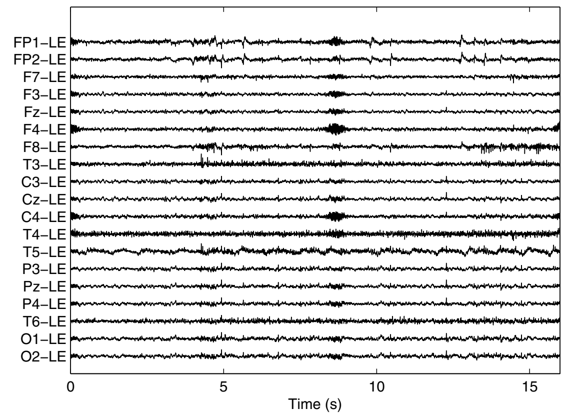

Fig. 7. Clean EEG data obtained by proposed approach.

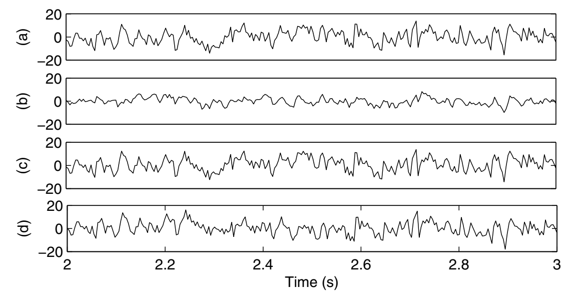

Fig. 8. Zoomed signals during artifact free epoch at FP1 electrode. (a) Raw EEG signal. (b) Rejecting artifact ICs. (c) Wavelet enhanced ICA. (d) Proposed approach.

ACKNOWLEDGMENTS

The authors of [7] are highly acknowledged for providing their MATLAB implementation of wICA and test EEG data for simulation studies.

REFERENCES

[1] S. Sanei and J.A. Chambers, *EEG Signal Processing*, John Wiley and Sons, 2007.

[2] T.P. Jung, S. Makeig, C. Humphries, T.W. Lee, M.J. Mckeowm, V. Iragui, and T.J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, pp. 163–178, 2000.

[3] L. Shoker, S. Sanei, W. Wang, and J.A. Chambers, "Removal of eye blinking artifact from the electro-encephalogram, incorporating a new constrained blind sourse separation algorithm," *Medical Biological Engineering Computing*, vol. 43, pp. 290–295, 2005.

[4] A. Hyvarinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Netw.*, vol. 13, pp. 411–430, 2000.

[5] Y. Tran, A. Craig, P. Boord, and D. Craig, "Using independent component analysis to remove artifact from electroencephalographic measured during stuttered speech," *Med. Biol. Eng. Comput.*, vol. 42, pp. 627–633, 2004.

[6] C.J. James and C.W. Hesse, "Independent component analysis for biomedical signals," *Physiol. Meas.*, vol. 26, pp. R15–R39, 2005.

[7] N.P. Castellanos and V.A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," *J. Neuroscience Methods*, vol. 158, pp. 300–312, 2006.

[8] C.W. Hesse and C.J. James, "The FastICA Algorithm With Spatial Constraints," *IEEE Signal Processing Lett.*, vol. 12, no. 11, pp. 792–795, 2005.

[9] C.W. Hesse and C.J. James, "On Semi-Blind Source Separation Using Spatial Constraints With Applications in EEG Analysis," *IEEE Trans. Biomedical Engineering*, vol. 53, no. 12, pp. 2525–2534, 2006.

[10] S. Mallat, *A wavelet tour of signal processing*, Academic Press, 1999.

[11] R. Romo-Vazquez, R . Ranta, V. Louis-Dorr, and D. Maquin, "EEG ocular artefacts and noise removal," *in Proc. 29th Annual International Conf. of IEEE EMBS*, Cite Internationale, Lyon, France, August 23-26, 2007, pp. 5445–5448.