

Improving P300 and SCP-Based Brain Computer Interfacing by Spectral Subtraction Denoising

Meena M. Makary¹ and Yasser M. Kadah², Senior Member, IEEE

Abstract—A new denoising technique for preprocessing of P300 and Slow Cortical Potential (SCP)-based Brain computer interface data is proposed. This new technique adaptively removes the superimposed noise using a modified version of spectral subtraction method. A better performance is achieved especially when less number of electrodes is used which accordingly reduce weight and consumed power for portable BCI applications. Classification accuracy and bitrate estimate were used as quantitative performance measures. Results showed better performance when compared to preprocessing without denoising and with using the relevant and widely used wavelet shrinkage denoising method. Results proved the practical utility of this method and we suggest adding it to different BCI experiments.

I. INTRODUCTION

Brain computer interfacing (BCI) is considered an important tool used for direct communication between subject's brain and computer by reading the brain activity. This information can be used to perform different actions controlled by the subject and hence provide an extra means of communication beside normal communication channels present in normal people. Different methods including electroencephalography (EEG), magnetoencephalography (MEG), and functional magnetic resonance imaging (fMRI) can be used to measure the brain activity at different locations. EEG-based system provides a simple, relatively inexpensive electrode cap attached to relatively small processing unit with low signal-to-noise ratio (SNR) and at the same time it allows subject mobility. The high levels of noise combined with the highly variable EEG signal lead to difficult extractable features and also decrease the information transfer rate (bitrate). So, approaches to remove such noise and artifacts are highly desirable to achieve better practical utility of EEG-based BCI systems. To decrease such amount of noise and increase the bitrate of EEG-based BCI system accordingly, we aim to study the effect of adaptive denoising on the accuracy of P300 as well as SCP-based BCI experiments. The problem of improving the quality of EEG signals for BCI applications was addressed by several research articles in past decade. This improvement can be recognized in two broad categories; namely, temporal domain techniques and spatial domain techniques [1]. In the spatial domain techniques [2] the data acquired from multiple spatially-distinct channels are used to distinguish the true component of the signal

projected onto all those channels from the noise which is assumed to be totally independent among such channels. Such methods include spatial averaging and blind source separation methods (BSS) such as independent component analysis (ICA) [3]–[5]. On the other hand, In temporal domain techniques, finding similarities within the time domain of a channel signal is attempted to be used for identification and suppression of noise component in that signal. Methods ranging from averaging of signals that improves the SNR to transform domain based techniques such as different variants of wavelet shrinkage can be used for signal denoising [6]–[9]. Such methods have achieved significant improvements, but there are still different limitations that need further research to be reduced. For example, the availability of many electrodes (or channels) is an essential condition for spatial domain methods, which would increase the cost, the weight and the power consumption of a portable BCI system. Moreover, in temporal domain techniques, signals need to be integrated into the preprocessing chain of BCI signals which have somewhat high computational complexity. Therefore, it is highly desirable to introduce a technique that provides the use of a small number of channels and at the same time improves the BCI system performance beyond the present methods. The aim of this work is to develop a denoising method based on spectral subtraction for P300-based and SCP-based BCI data that provides better performance when lower number of channels are used. The new denoising method will be applied to two different experimental datasets, one with P300-based and another with SCP-based, and then the classification results are compared to that of the same datasets using the same preprocessing and classification steps to allow direct comparison of results. Also, the new method will be compared to a relevant and widely used method for denoising, namely, standard wavelet shrinkage based denoising. Classification accuracy and bit rate will be used as quantitative performance measures for these different experiments.

II. METHODS

The classification of P300 as well as SCP trials in EEG data is made difficult due to the low SNR of their responses. To overcome this problem, it is common practice to average together many consecutive trials, which effectively diminishes the random noise, but unfortunately when more repeated trials are required for applications such as the P300 speller and SCP-based experiments, the communication rate is greatly reduced. Since the noise results from background brain activity and is inherent to the EEG recording methods,

¹Meena M. Makary is with Biomedical Engineering Department, Faculty of Engineering, Cairo University, Giza 12613, Egypt. meena.magharious@eng.cu.edu.eg

²Yasser M. Kadah is with the Biomedical Engineering Department, Faculty of Engineering, Cairo University, Giza 12613, Egypt. ykadah@eng.cu.edu.eg

noise reduction techniques like adaptive denoising increases SNR and accordingly the classification accuracy of P300 and SCP trials.

A. Theory of Adaptive Denoising

Generally speaking, we can model the EEG temporal signal as a summation of three components; a true activation response, a physiological baseline fluctuation component, and a random noise component [1]. The physiological baseline fluctuation component can be considered as a deterministic yet unknown signal such as motion artifact and can be dealt with using existing preprocessing methods as windsorizing [10]. On the other hand, the random noise component comprised of two parts: the thermal noise coming from the data acquisition system electronics which is well known to be Gaussian white noise, and the superimposed signals from neighboring neurons not involved in the true response at all. The later component can be dealt with applying the central limit theorem to the summation of many signals of random activation patterns. Therefore, we will consider the model of EEG temporal signal as the sum of one deterministic component comprised of both the true activation signal and the physiological noise and an uncorrelated stochastic one . That is

$$s(t) = d(t) + n(t), \quad (1)$$

As these two components are assumed to be independent, the corresponding power spectra are related by

$$P_{ss}(\omega) = P_{dd}(\omega) + P_{nn}(\omega), \quad (2)$$

where cross terms vanish because the uncorrelatedness of these two components. Hence, we can estimate the power spectrum of the deterministic component using the form [1]

$$P_{dd}(\omega) = P_{ss}(\omega) - P_{nn}(\omega). \quad (3)$$

That is, spectral subtraction of the noisy signal power spectrum and that of noise is used to obtain the power spectrum of the denoised signal. In order to compute the deterministic signal component from its power spectrum, we should determine the magnitude and phase of Fourier transform of the signal. The former part can be obtained as the square root of the power spectrum, while an estimate of the later one can be obtained from phase of the Fourier transform of the original signal $s(\omega)$. Hence, the Fourier transform of the processed/denoised signal $s_d(\omega)$ can be expressed as

$$s_d(\omega) = \sqrt{P_{dd}(\omega)} \cdot \exp(j \cdot \text{phase}(s(\omega))). \quad (4)$$

The enhanced/denoised deterministic signal $s(\omega)$ is the real part of the inverse Fourier transformation of this expression. It should be noted that this method is related to the optimum Wiener filter [11]. A block diagram for the proposed strategy is shown in Fig. 1.

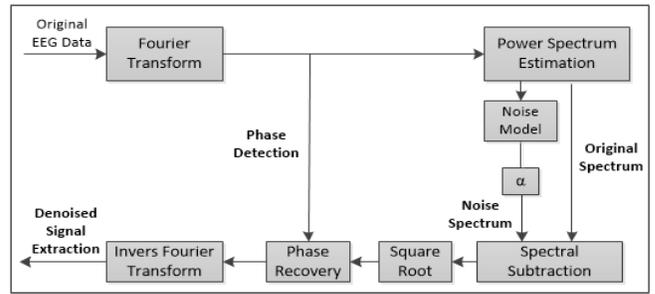


Fig. 1. Block diagram of the proposed technique.

B. Level of Denoising

We can observe from the nature of the EEG signal that the variance in the estimate of power spectrum may only result from the random component. Given the statistical characteristics of the periodogram estimate and the fact that the expected value of the noise variation is known from the derived model, the noise at each of the power spectrum frequency bins can be expressed as a Gaussian random variable with mean and variance both equal to the noise model [12], [13]. As a result, the subtraction in (3) would not effectively remove all parts of the noise power spectrum. To overcome this problem, we add a slight modification to provide direct control over the level of noise removed. The modified equation takes the form

$$P_{dd}(\omega) = P_{ss}(\omega) - \alpha \cdot P_{nn}(\omega). \quad (5)$$

Here, we add α as a control factor of the confidence level of noise removal. The problem of choosing the value of this factor can be expressed when α is chosen to be the p-value of a statistical z test. That is, when we increase the value of α , the probability that the output power spectrum contains a noise component will be decreased. On the other hand, the probability of removing some parts of the signal would increase by increasing this value. Therefore, selecting the optimum value of α is an important issue to improve the performance of this new technique. In this paper, we studied the effect of several values of α and chose the best value based on the resulted classification accuracy. The optimization of this parameter will be left for further investigation.

C. Practical Implementation Procedure

The practical implementation steps can be done as shown in the block diagram Fig. 1. The noise model is obtained from the background part of the fast Fourier transform of the EEG signal as shown in Fig. 2. Then spectral subtraction denoising method is compared with Wavelet Shrinkage denoising method as shown in Fig. 3.

III. EXPERIMENTAL VERIFICATION

A. P300-based Experiment

In the first part of this work, we tested the developed denoising method using the data of Hoffmann et al. [10] and compared it to both results of no denoising case and the case of denoising by standard wavelet shrinkage denoising method

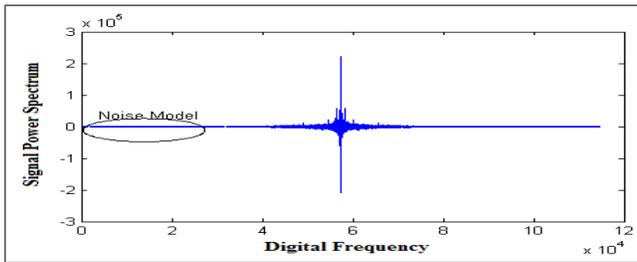


Fig. 2. Noise power spectrum estimation from the upper part of the signal power spectrum known to have no true signal components.

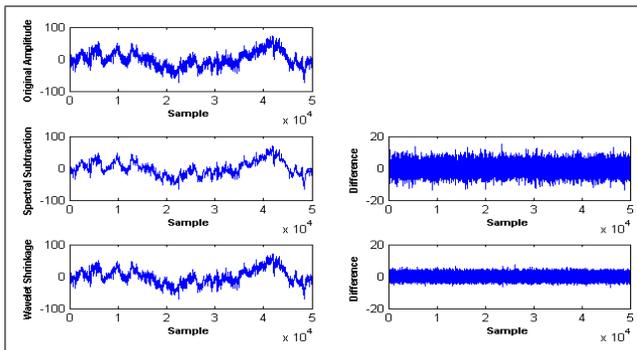


Fig. 3. Illustration of the results of spectral subtraction denoising as compared to the original signal and wavelet shrinkage denoising.

[7], [14]. We followed the same steps of preprocessing and classification in this paper to allow a direct comparison between the classification results of the two cases of preprocessing with and without the proposed denoising technique. For detailed description of the data set and the experiment, refer to [10]. The effect of different electrode configurations and machine learning algorithms on classification accuracy was tested in an offline procedure. To estimate average classification accuracy four-fold cross-validation was used. The preprocessing operations applied were: referencing using the two mastoid electrodes, bandpass filtering with cut-off frequencies set to 1.0 Hz and 12.0 Hz using a sixth order forward-backward Butterworth filter, downsampling by a factor of 64, single trials were extracted, windsorizing to remove eye blinks, eye movement and muscle activity, and finally amplitude normalization. The number of electrodes was selected as 4, 8, 16 or 32 depending on the experiment with the same electrode configurations in [10]. Then, Feature vector construction is done. The samples from the selected electrodes were concatenated to construct the feature vectors. The dimensionality of the feature vectors was $N_e \times N_t$, where N_e is the number of electrodes (selected as 4, 8, 16, or 32) and N_t is the number of temporal samples in one trial (32 samples in our experiments). Bayesian linear discriminant analysis (BLDA) is used as a classifier and the software developed by [10] was used to perform this step. For the comparison with wavelet denoising method, standard wavelet shrinkage denoising was used using MATLAB with the basic wavelet chosen as “Coiflet-3” as suggested by [14]. The universal threshold was selected and no multiplicative threshold rescaling.

B. SCP-based Experiment

In this paper, we also studied the effect of adaptive denoising on the classification accuracy of BCI competition II experiment. For more details about the data set refer to [15]. As we just want to test the effect of adaptive denoising as a preprocessing step on the SCP-based experiment, we followed the sequence of Wu Ting et al. work and compare the results with and without the denoising step. Actually we selected this work because their results were better than the best results of the competition for this dataset [16]. The feature vector is constructed as a concatenation of the average and the energy of Wavelet Packet decomposition sub-bands. Fisher criterion [16] was used to select the most promising features. A four layers Probabilistic Neural Network (PNN) was used as a classification method: one input layer, one output layer, one hidden layer and one added layer.

IV. RESULTS AND DISCUSSION

A. Choice of Denoising Parameter α

To select the best value of α we repeated the experiment for different denoising values of α and choose the value which gives the best classification accuracy. Fig. 4 shows the classification accuracy and bitrate vs. time after denoising for four electrode configurations with different values of α . As we can see the best value of denoising parameter α is 1000 which gives best classification accuracy (reach 100% accuracy faster). Also this value gives the best classification accuracy curve for other electrode configurations.

B. Classification Accuracy and Bitrate

Theoretically, adaptive denoising has a significant effect when a small number of electrodes is used because in that case SNR is relatively low and hence adaptive denoising is needed too increase it. As BCI should be a real time application, a small number of electrodes is highly recommended to reduce the processing complexity and time consumed. For P300-based experiment, Fig. 5 shows the classification accuracy and bitrate versus time before denoising (blue line for accuracy and green line for bitrate) and after denoising (red lines) with using the most promising value of $\alpha = 1000$ for subject 1 & subject 2 from disabled subjects and subject 8 & subject 9 from normal subjects. As we can see, adaptive denoising always increases classification accuracy and bitrate

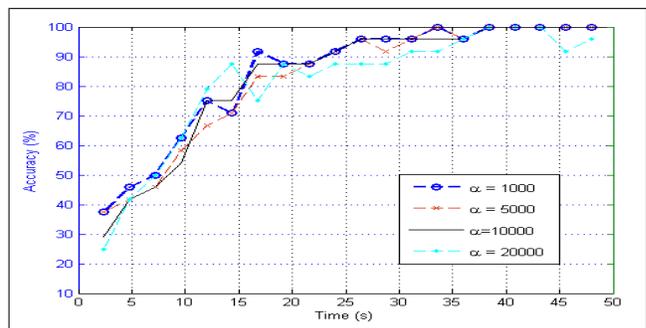


Fig. 4. The choice of the most promising α based on the value which gives the best classification accuracy.

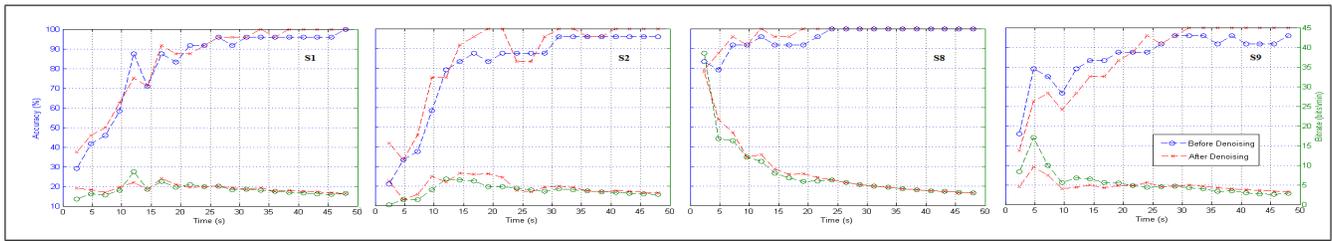


Fig. 5. Classification accuracy and bitrate plotted vs. time before and after denoising for 4 electrode configuration ($\alpha = 1000$) for subjects 1, 2 from disabled subjects and subjects 8, 9 from normal subjects.

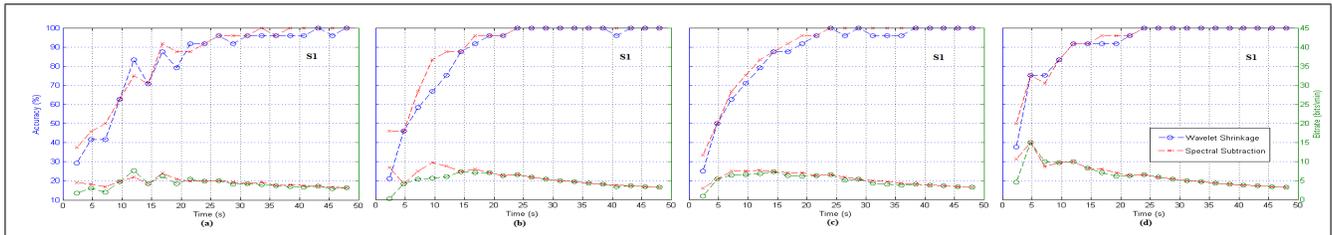


Fig. 6. Classification accuracy and bitrate plotted vs. time comparison between Wavelet and Spectral Subtraction methods for: (a) 4 electrode (b) 8 electrode, (c) 16 electrode, (d) 32 electrode configurations.

for all subjects. As a comparison with wavelet shrinkage denoising method, spectral subtraction denoising proved a better performance for all electrode configurations as shown in Fig. 6 and it also proved better denoising for the EEG signal as shown in Fig. 3. For SCP-based Experiment, TABLE I shows results comparison for the BCI competition III dataset. The best result of the competition was (88.7%). After spectral subtraction denoising we got a classification accuracy of (91.4%) which is better than the best accuracy achieved for this data set (90.8%) [16].

TABLE I
RESULTS COMPARISON OF CLASSIFICATION ACCURACY

Method	Accuracy (%)
The best result of competition	88.7
Results of Wu Ting et al. [16]	90.8
Results after Adaptive Denoising	91.4

V. CONCLUSION

In this work we introduced the theory of a new preprocessing technique and studied its effect on classification accuracy of different BCI experiments. In P300-based experiment, results indicated that adaptive denoising has a significant effect when small number of electrodes is used (4 and 8 electrode configurations). Adaptive denoising proved a better performance than wavelet shrinkage denoising method. In SCP-based experiment, classification accuracy was also increased using this denoising method. Finally, the effect of adaptive denoising in different experiments suggests the potential for practical utility of the method.

REFERENCES

- [1] Y. M. Kadah, "Adaptive denoising of event-related functional magnetic resonance imaging data using spectral subtraction." *IEEE transactions on bio-medical engineering*, vol. 51, no. 11, pp. 1944–53, Nov. 2004.
- [2] A. de Cheveigné and J. Z. Simon, "Denoising based on spatial filtering." *Journal of neuroscience methods*, vol. 171, no. 2, pp. 331–9, Jun. 2008.
- [3] R. R. Ramírez, B. H. Kopell, C. R. Butson, B. C. Hiner, and S. Baillet, "Spectral signal space projection algorithm for frequency domain MEG and EEG denoising, whitening, and source imaging." *NeuroImage*, vol. 56, no. 1, pp. 78–92, May 2011.
- [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, no. 2, pp. 401–416, Feb. 2012.
- [5] E. P. Sumesh, G. Geetha, and S. Geethalakshmi, "Artifact Removal from EEG using Spatially Constrained Independent Component Analysis and Wavelet Denoising with Otsu's Thresholding Technique," *Procedia Engineering*, vol. 30, pp. 1064–1071, 2012.
- [6] M. Ahmadi and R. Quian Quiroga, "Automatic denoising of single-trial evoked potentials." *NeuroImage*, vol. 66C, pp. 672–680, Nov. 2012.
- [7] D. Donoho, "De-noising by soft-thresholding." *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
- [8] R. Quian Quiroga and H. Garcia, "Single-trial event-related potentials with wavelet denoising." *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 114, no. 2, pp. 376–90, Mar. 2003.
- [9] A. Effer, K. Lehnertz, T. Grunwald, G. Fernández, P. David, and C. E. Elger, "Time adaptive denoising of single trial event-related potentials in the wavelet domain." *Psychophysiology*, vol. 37, no. 6, pp. 859–65, Nov. 2000.
- [10] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient P300-based brain-computer interface for disabled subjects." *Journal of neuroscience methods*, vol. 167, no. 1, pp. 115–25, Jan. 2008.
- [11] B. Lathi, *Modern Digital and Analog Communication Systems*, 2nd ed., Englewood Cliffs, NJ: Prentice Hall, 1994.
- [12] S. Vaseghi, *Advanced digital signal processing and noise reduction*, New York: Wiley, 1996.
- [13] R. W. C. S. Burrus, J. H. McClellan, A. V. Oppenheim, T. W. Parks, Schafer, and H. W. Schuessler, *Computer-Based Exercises for Signal Processing Using MATLAB*, Englewood Cliffs, NJ: Prentice Hall, 1994.
- [14] C. Saavedra and L. Bougrain, "Wavelet Denoising for P300 Single-Trial Detection," in *Proc. of the 5th French conference on computational neuroscience (Neurocomp'10)*, vol. 10, 2010, pp. 227–231.
- [15] "Bci competition ii website." [Online]. Available: <http://www.bbci.de/competition/ii/>
- [16] W. Ting, Y. Guo-zheng, Y. Bang-hua, and S. Hong, "EEG feature extraction based on wavelet packet decomposition for brain computer interface," *Measurement*, vol. 41, no. 6, pp. 618–625, Jul. 2008.