

# Spectral Subtraction Denoising Improves Accuracy of Slow Cortical Potential Based Brain-Computer Interfacing

Meena M. Makary and Hani M. Bu-Omer  
Systems & Biomedical Engineering  
Cairo University  
Giza, Egypt  
meena.magharious@eng.cu.edu.eg

Yasser M. Kadah  
Electrical and Computer Engineering  
King Abdul Aziz University  
Jeddah, Saudi Arabia  
ykadah@kau.edu.sa

**Abstract**—Spectral subtraction denoising as a new preprocessing block for Slow Cortical Potential(SCP)-based brain computer interface is proposed. This method adaptively reduces the noise attached to the signal using spectral subtraction denoising. We show that this technique provides better performance when low number of electrodes is used. This suggests its potential application in portable BCI real time applications where low power and minimum weight are required for practical use. Here, the classification accuracy was used as performance measure. The developed method provides relatively better performance compared to the more commonly used wavelet shrinkage signal denoising.

**Keywords**—Brain-computer interface, Spectral Subtraction, Wavelet Shrinkage, Signal Denoising, Slow Cortical Potential.

## I. INTRODUCTION

Brain computer interfacing (BCI) is a tool for direct communication between subjects brain and computer by interpreting brain activity recorded by EEG. Since the performance of such tool relies on the quality of EEG recording, the problem of improving the signal-to-noise ratio (SNR) of EEG signals has been the focus of several research groups. Generally speaking, methods to improve SNR fall into two broad categories; namely, temporal and spatial domain techniques [1]. In temporal domain techniques, finding the time domain similarities of a channel signal is aimed to be used for identification and suppression of noise component in that signal. Methods ranging from improving the SNR by averaging of signals to transform domain based techniques such as different variants of wavelet shrinkage can be used for signal denoising [2], [3]. In contrast, spatial domain techniques [4], [5] use the data acquired from different electrodes to distinguish the true component of the EEG signal from the noise component which assumed to be independent among those channels. Such methods include forms of spatial averaging [6]–[8]. However, it is essential to have large number of electrodes (or channels) to use this approach, which may not be practical in many cases. Furthermore, BCI signals which are preprocessed by techniques in temporal domain have somewhat high computational complexity. Subsequently, it is highly desirable to introduce a method that allows using reduced number of electrodes and improves the BCI systems utility beyond current methods [9].

This work aims to develop a denoising technique for SCP-based BCI data. This new adaptive denoising technique is applied to SCP-based experimental datasets with the performance

measure taken as the classification accuracy and compared to the case without denoising and when the widely used wavelet shrinkage denoising technique is used.

## II. METHODS

The classification of SCP trials in EEG data is difficult due to the inherent low SNR. To solve this problem, averaging many consecutive trials together is common practice that effectively reduces the random noise. Unfortunately, this causes the communication rate to be significantly diminished. Since the noise coming from background brain activity is inherent to the EEG recording systems, noise reduction methods such as adaptive denoising can increase the SNR, and hence the classification accuracy, of SCP trials.

In this work, we develop a new approach for adaptive denoising that addresses this problem based on the well-known spectral subtraction method [10]. This technique is implied from the fact that accurate information about the random noise model parameters can be easily obtained adaptively from EEG datasets. The technique used the uncorrelatedness of the SCP signal's deterministic component and the random noise to separate them in the original power spectrum.

### A. Spectral Subtraction Denoising Level

EEG signal nature shows that the variation in the estimate of power spectrum may be only due to the random part compromised on it. From the characteristics of the periodogram statistical estimate and the noise variation expected from the predicted model, the noise at each frequency bin of the power spectrum can be expressed as a Gaussian random variable with statistical parameters equal to the noise model. So, we would not effectively remove all parts of the noise power spectrum. That is, the upper half of Gaussian distribution of the noise would still exist in the processed signal [11].

To overcome this problem, we add a slight adjust to obtain direct control over the level of the removed noise. The adjusted equation is

$$P_{rr}(\omega) = P_{ss}(\omega) - \gamma P_{nn}(\omega). \quad (1)$$

Here, we add  $\gamma$  as a factor for controlling the noise removal confidence level. We can statistically express the problem of selecting the value of  $\gamma$  in the form of z test by assigning  $\gamma$  as the p-value of this test. That is, the likelihood of finding a noise part in the output power spectrum is inversely related the selected  $\gamma$  value. On the other hand, as we increase this value the likelihood of removing true signal parts would increase.

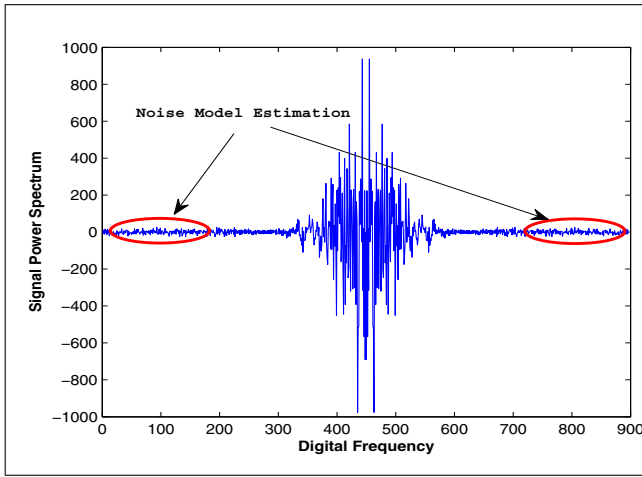


Fig. 1. Estimation of Noise power spectrum from the background part of the original signal power spectrum which don't have any components from the true signal.

Therefore, it is important to select the optimum value of  $\gamma$  that improves the performance of this new method. In this paper, we studied the effect of several values of  $\gamma$  and the resulted classification accuracy was used to choose the best value of it. The optimization criteria of this factor is left for further research.

### B. Implementation Steps

The practical implementation steps applied for spectral subtraction adaptive denoising method are:

- 1) Removing the mean of the signal.
- 2) Selecting a temporal part within the background/first area of the FT of the SCP signal to obtain the model of noise as illustrated in Fig. 1.
- 3) Computing the Fourier transform for each channel of the SCP signal and separately saving both the magnitude and phase of the result.
- 4) Using the periodogram to compute the original/true power spectrum of this signal as the magnitude square of the FT resulted from step 3.
- 5) Computing the denoised/filtered power spectrum by subtracting the power spectrum of the noise resulted from step 2 with taking into consideration the denoising control factor  $\gamma$  as in (1).
- 6) Computing the FT of the denoised EEG signal as the square root of the denoised/filtered power spectrum obtained from step 5 multiplied by the phase retained in step 3.
- 7) Computing the denoised/filtered signal from the real part of the inverse FT of the result of step 6.

Then spectral subtraction denoising method is compared with Wavelet Shrinkage denoising method as shown in Fig. 2.

## III. EXPERIMENTAL VERIFICATION

To verify our method, we used BCI competition II dataset and compared the classification accuracy with the best achieved results for this dataset. To ensure the consistency of the comparison, the same sequence of processing, feature

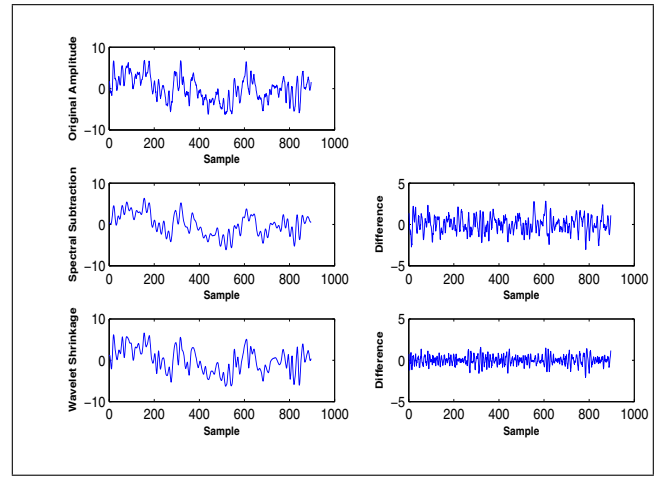


Fig. 2. Results comparison between wavelet shrinkage denoising and adaptive spectral subtraction denoising method.

extraction & selection and classification of [12] was followed with and without an initial preprocessing block based on our method. At the  $k^{\text{th}}$  sample and  $j^{\text{th}}$  level, WPD coefficients can be obtained from:

$$d_j^n(k) = \sum_m h_0(m-2k) d_{j-1}^{\frac{n}{2}}(m), \quad (2)$$

in case of  $n$  is even. and

$$d_j^n(k) = \sum_m h_1(m-2k) d_{j-1}^{\frac{(n-1)}{2}}(m), \quad (3)$$

in case of  $n$  is odd.

where,  $k$  and  $j \in \mathbb{Z}$ , the values of  $n = 0, 1, \dots, 2^j - 1$  and the couple of quadruple mirror filters  $h_0(k), h_1(k)$  are related by:

$$h_1(k) = (-1)^{(1-k)} h_0(1-k). \quad (4)$$

After decomposition of the signal by  $j$  levels, the ranges of frequency of all sub-bands at the  $j^{\text{th}}$  level are:

$$\left\{ \left[ 0, \frac{f_s}{2^{(j+1)}} \right]; \left[ \frac{f_s}{2^{(j+1)}}, \frac{2f_s}{2^{(j+1)}} \right]; \dots; \left[ \frac{(2^j - 1)f_s}{2^{(j+1)}}, \frac{f_s}{2} \right] \right\}, \quad (5)$$

where,  $f_s$  is the sampling frequency.

Since combining information from both time domain and frequency domain can provide more complete features, we used coefficients mean and sub-band energies to create the final feature vector. Fisher distance criterion was used as selection criteria and dimensionality reduction method.

(1) **Average coefficients:** As the useful frequency range of the EEG signal is from 0-50 Hz, we calculated the sub-band means ( $MEAN_{j,n}$ ) at the  $j^{\text{th}}$  level in this frequency range as primary features using this equation:

$$MEAN_{j,n} = \frac{2^N}{2^j} \sum_k d_j^n(k), \quad (6)$$

After WPD to the six level ( $j = 6$ ), recommended by [12], and as  $f_s = 256$  for this dataset so according to (5) the sub-band frequencies will be  $\{[0-2], [2-4], \dots\}$ . The useful EEG frequency range is in the first 25 sub-bands which correspond to [0-50]Hz. So we have 25 mean features for each channel.

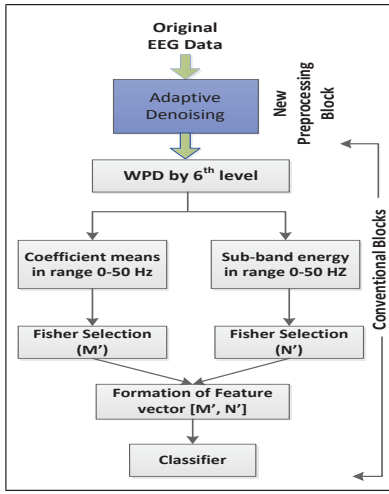


Fig. 3. Feature vector formation procedure including denoising block.

So our initial mean feature vector contains 150 features (6 channels \* 25 features each)  $M = [m_1, m_2, \dots, m_{150}]$ .

(2) **Sub-band Energy:** Similarly as the average coefficients vector, the sub-band energy vector ( $ENERGY_{j,n}$ ) is calculated at the  $j^{th}$  level which correspond to frequency range of [0-50] Hz as preliminary energy features using the following equation:

$$ENERGY_{j,n} = \sum_k (d_j^n(k))^2. \quad (7)$$

Similarly, for the six channels the feature vector will be  $N = \{n_1, n_2, n_3, \dots, n_{150}\}$ . Finally, Fisher distance criterion is used as feature selection criteria to select the most promising features among the mean and energy initial features. The function of Fisher criterion,  $F(w)$ , can take the form:

$$F(w) = w^T S_B w / w^T S_I w, \quad (8)$$

where  $S_B$  and  $S_I$  are the between-class scatter and the within-class scatter matrices respectively. Fisher weight vector  $w$  can be calculated by maximizing the value of Fisher function  $F(w)$ . To measure the separability of features, we used Fisher criterion using the following equation:

$$J = \text{tr}(S_w^{-1} S_b), \quad (9)$$

where  $S_b$  and  $S_w$  are the between and within classes dispersion matrices respectively. The larger the value of  $J$ , the better classification separability obtained from this feature. Creation steps of the feature vector including denoising step is shown in Fig. 3.

After calculating  $J$  for each feature in  $M$  and select the most promising ones with largest values of  $J$  say the first  $d$  values, the new features of mean vector can be written as:  $M' = [m'_1, m'_2, \dots, m'_d]$ . Similarly, for energy features  $N$ , by selecting the first  $l$  values with largest values of  $J$ . The new vector will be  $N' = [n'_1, n'_2, \dots, n'_l]$ . The final feature vector for classification step can be written as  $X = [M', N']$ .

#### IV. RESULTS AND DISCUSSION

Decomposition of EEG signal to the six level was done using Daubechies Wavelet (db4). There are 64 sub-bands of wavelet packet at the sixth level which are corresponding to

frequency ranges of  $\{[0,2],[2,4],\dots,[126,128]\}$ Hz, among them there are 25 sub-bands belong to the useful frequency range [0-50] Hz, so the final feature vector consists of 150 features (6 channels x 25 sub-bands in each one). Each component separability  $J$  is calculated using Fisher's criterion. As we can see in Fig. 4, the separability of the mean vector  $M$ , it has two peaks, so the final mean feature vector will be of only two dimensions i.e.  $M' = [m'_1, m'_2]$ . Similarly for the energy vector  $N$ , separability of each component is shown in Fig. 5, it can be shown that it has a lot of peaks. Selecting the first 15 value (the classification accuracy will not be affected if we increased this value) to construct the final energy vector of 15 dimension in  $N$  i.e.  $N' = [n'_1, n'_2, n'_3, \dots, n'_{15}]$ . The final feature vector  $X = [M', N']$  is of dimension 17 features.

To study the effect of the new method, we construct the feature vector in the same way as described before but with including the spectral subtraction denoising step. Fig. 6 and Fig. 7 show the separability ( $J$ ) of each component in the Mean vector ( $M$ ) and the Energy vector ( $N$ ) respectively. As we can see from the figures the peaks in the mean and energy vector are the same as before denoising but with larger amplitude which means these features became more promising in classifying the data. Furthermore other peaks appeared which can be added to the final feature vector.

After repeating all the previous steps after adaptive denoising to the EEG data with denoising parameter  $\gamma = 1000$ , we got a classification accuracy (classification is done by four layers probabilistic neural network and the effect of regularized

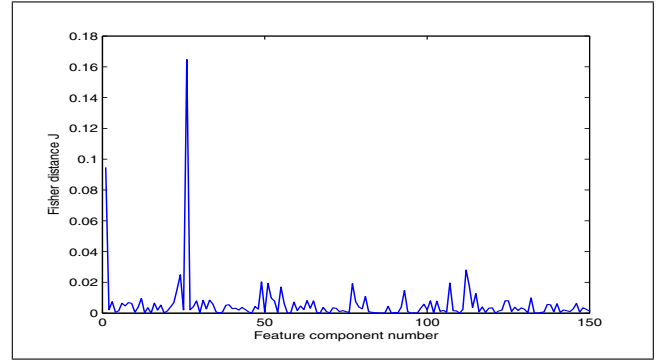


Fig. 4. Each component separability in the mean vector  $M$  without denoising step.

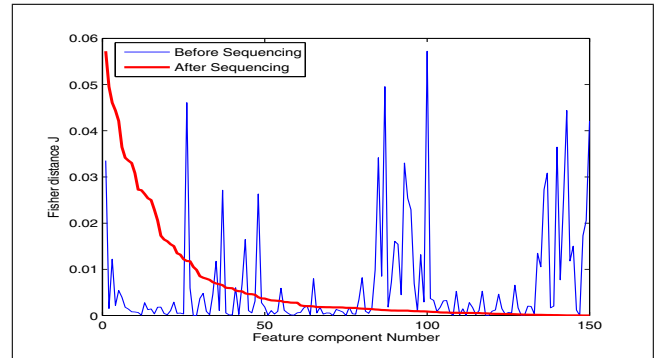


Fig. 5. Each component separability in the energy vector  $N$  without denoising step.

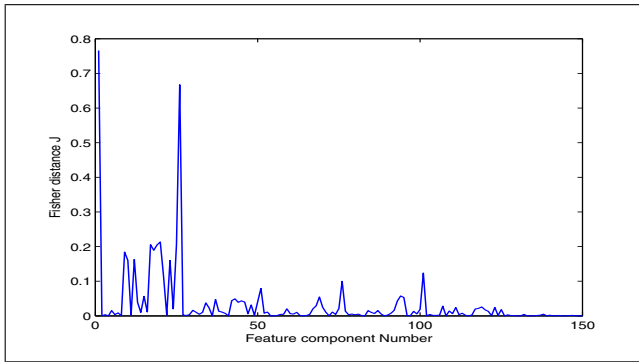


Fig. 6. Each component separability in the mean vector M after denoising step.

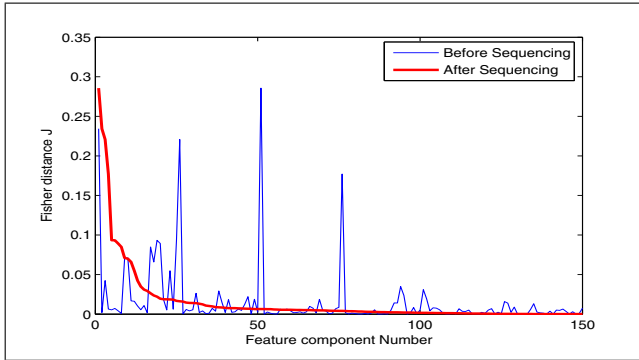


Fig. 7. Each component separability in the energy vector N after denoising step.

neural networks is left for further investigations [13]) of 91.4% which is better than the best results obtained in the literature for this data set, also we got 94% after adding 6 more features from energy vector to the final feature vector with insignificant increase in processing time. Table I shows results comparison.

TABLE I. RESULTS COMPARISON OF CLASSIFICATION ACCURACY BEFORE AND AFTER DENOISING

Method	Accuracy (%)
Best achieved result of competition	88.7
Results of Wu Ting et al. [12]	90.8
Results after Adaptive Denoising	91.4
Adaptive Denoising with adding 6 more features	94

## V. CONCLUSION

A new preprocessing technique for *adaptive denoising* was developed and its impact on the classification accuracy of SCP-

based BCI experiments was verified. Adaptive denoising provided a better performance than wavelet shrinkage denoising method in removing the undesired random noise component of the signal and resulted in an increased classification accuracy. The results of adaptive denoising suggests the potential for practical utility to include the new block as a standard preprocessing step to enhance present BCI systems.

## REFERENCES

- [1] Y. M. Kadah, "Adaptive denoising of event-related functional magnetic resonance imaging data using spectral subtraction." *IEEE Trans Biomed Eng.*, vol. 51, no. 11, pp. 1944–53, Nov. 2004.
- [2] M. Ahmadi and R. Quian Quiroga, "Automatic denoising of single-trial evoked potentials." *NeuroImage*, vol. 66C, pp. 672–680, Nov. 2012.
- [3] A. Effern, 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." *Psychophys*, vol. 37, no. 6, pp. 859–65, Nov. 2000.
- [4] G. Pires, U. Nunes, and M. Castelo-Branco, "Statistical spatial filtering for a P300-based BCI: tests in able-bodied, and patients with cerebral palsy and amyotrophic lateral sclerosis." *J Neurosci Meth*, vol. 195, no. 2, pp. 270–81, Feb. 2011.
- [5] A. de Cheveigné and J. Z. Simon, "Denoising based on spatial filtering." *J Neurosci Meth*, vol. 171, no. 2, pp. 331–9, Jun. 2008.
- [6] 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.
- [7] 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." *Sig Proc*, vol. 92, no. 2, pp. 401–416, Feb. 2012.
- [8] 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 Eng.*, vol. 30, pp. 1064–1071, 2012.
- [9] M. M. Makary and Y. M. Kadah, "Processing Methodologies For Brain Computer Interface". ISBN: 978-3-659-56019-4. LAP LAMBERT Academic Publishing, GERMANY, 2014.
- [10] M. J. Alhaddad, M. I. Kamel, M. M. Makary, H. Hargas, and Y. M. Kadah, "Spectral subtraction denoising preprocessing block to improve P300-based brain-computer interfacing." *Biomedical engineering on-line*, vol. 13, no. 1, p. 36, Jan. 2014.
- [11] M. M. Makary and Y. M. Kadah, "Improving P300 and SCP-Based Brain Computer Interfacing by Spectral Subtraction Denoising," in *Middle East Conference on Biomedical Engineering (MECBME)*. Doha, Qatar: IEEE, 2014, pp. 228–231.
- [12] W. Ting, Y. Guo-zheng, Y. Bang-hua, and S. Hong, "EEG feature extraction based on wavelet packet decomposition for brain computer interface," *Measur*, vol. 41, no. 6, pp. 618–625, Jul. 2008.
- [13] M. M. Makary and I. A. Yassine, "Study of Effect of Regularized Neural Network on the Accuracy of Handwriting Recognition," in *The 5th International Conference on Information and Communication Technologies for Amazigh (TICAM)*, Rabat, Morocco, 2012.