Original Article

Enhancing P300 Wave of BCI Systems Via Negentropy in Adaptive Wavelet Denoising

Zahra Vahabi[1], Rassoul Amirfattahi[1,2], Abdolreza Mirzaei[3]

[1]Digital Signal Processing Research Lab, [2]Medical Image and Signal Processing Research Center, Isfahan University of Medical Sciences, [3]Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran

ABSTRACT

Brian Computer Interface (BCI) is a direct communication pathway between the brain and an external device. BCIs are often aimed at assisting, augmenting or repairing human cognitive or sensory-motor functions. Electroencephalogram (EEG) separation into target and non-target ones, based on presence of P300 signal, is a difficult task mainly due to their natural low signal to noise ratio. In this paper, a new algorithm is introduced to enhance EEG signals and improve their signal to noise ratio. Our denoising method is based on multi-resolution analysis via Independent Component Analysis Fundamentals. We have suggested combination of negentropy as a feature of signal and sub-band information from wavelet transform. The proposed method is finally tested with dataset from BCI Competition 2003, and has given results that compare favorably.

Key words: *Brian computer interface, denoising, independent component analysis, negentropy, P300 speller, wavelet transform*

INTRODUCTION

The Brian Computer interface (BCI) system is a set of signal processing components and sensors that allows acquiring and analyzing brain activities with the goal of establishing a communication channel directly between the brain and an external device, such as a computer, neuroprosthesis, and etc, by analyzing electroencephalographic (EEG) activities that reflect the functions of the brain.[1,2]

There are many BCI systems based on EEG rhythms, such as Alpha, Beta, Mu, Slow Cortical Potentials (SCPs), Event Related Synchronization/Desynchronization (ERS/ERD) phenomena, Steady-State Visual Evoked Potential (SSVEPs), P300 component of the Evoked-Related potentials (ERP's) and so on.[3]

P300 Based BCI was first introduced by Farwell and Donchin in 1988, for controlling an external device.[4-8]

The P300 (P3) wave is an event related potential elicited by task-relevant, infrequent stimuli. This wave is considered to be an endogenous potential as its occurrence links to a person's reaction to the stimulus, not to the physical attributes of a stimulus. This is a positive ERP, which its occurrence is over the parietal cortex with a latency of about 300 ms after rare or task relevant stimuli. Mainly, the P300 is thought to reflect tasks involved in stimulus evaluation or categorization.[9]

Within BCI, P300 potentials can provide a means of detecting a person's intention concerning on the choice of object. Detecting the P300 peaks in the EEG accurately is the main goal. In a P300 speller, therefore, variety of feature extraction and classification procedures have been implemented, improving the performance. A good preprocessing step could enhance the signal to noise ratio (SNR) and help to have simple, more accurate algorithm for feature extraction and classification. More reliable and ever fast signal processing methods for preprocessing the recorded data are crucial in the improvement of practical BCI systems.

Single-trial ERP detection is understood to be challenging, as P300 waves and other task related signal components have a large amount of noise (artifacts-ongoing task and unrelated neural activities).[9,10]

A preprocessing step for P300 detection is applied to enhance the SNR, and to remove both interfering physiological signals as those related to ocular, muscular and cardiac activities, and non-physiological artifacts, such as electrode movements, broken wire contacts, and power line noise. To detect the specific patterns,

Address for correspondence:
Dr. Zahra Vahabi, Digital Signal Processing Research Lab, Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, 84156-83111, Iran, E-mail: z.vahabi@ec.iut.ac.ir

we extract features in brain signals encode the patient's motor intentions or reflect the user commands. Translating feature into control signals is aimed at last step to sent to an external device.[1,11]

Several methods, based on Independent Component Analysis (ICA), Fourier Transform, and Wavelet Transform (WT), were thus proposed to enhance the SNR and to remove the artifacts from EEG signals.[10,12]

The major drawback of ICA-based method is that they are supervised and they are not specifically designed to separate brain waves. For instance, ICA is a popular method to EEG denoising, but after the decomposition in independent components (IC), it is necessary to select (by spatio-temporal prior or manually) the ICs which contained the evoked potentials.

Wavelet denoising techniques have been adopted for signal enhancement applications. EEG signal is first mapped to discrete wavelet domain by means of multiresolution analysis. Details of coefficients related to additive noise could be eliminated using a contrast or adaptive threshold level. Finally, signal is reconstructed back to time domain using remaining wavelet coefficients.

Donoho[13] suggested a denoising technique in the wavelet domain in order to enhance the domain. It is by thresholding the wavelet coefficients in the orthogonal wavelet domain. Two thresholding algorithms were proposed, namely the soft thresholding and hard thresholding.

So, to denoising the signal, choosing a good threshold and algorithm would help having more clear signals.

In this paper, we propose a new unsupervised algorithm to automatically estimate noise subspace from raw EEG signal. The aim is to provide a new method to increasing the spelling debit.

In particular, our algorithm is an adaptive wavelet denoising, which is one of the best methods for signal denoising. The threshold on each WT levels are different and dependence of negentropy. Negentropy is the mainly parameter that ICA work based on. Negentropy of each level in WT is calculated and the lowest level negentropy is introduced for certifiable parameter. Other levels negentropy compare with this one and independency of such levels and the main (the lowest level) acquired. This independency, determines the threshold. So the threshold must be increased to eliminate the noise.

This paper is organized as follows: Section 2 describes the P300 subspace and the BCI Enhancement. Section 3 describes Independent Component Analysis and Wavelet Transform whereas. Section 4 presents our proposed algorithm. In the last parts, Sections 5 and 6 have simulated results and conclusion.

METHODOLOGY

On the P300 based BCI, the user was presented with a screen that is a 6*6 character matrix with 36 symbols. The user then, one by one, focuses on letters of an expected word. The columns and rows of the matrix are randomly flashed. Concerning the letter of the word, there are 12 illuminations (six columns and six rows) which provide the visual stimulus. Two flashes (one column and one row) out of the twelve intensifications decide a character which the user wants to say. It is expected that the evoked waveforms are different from others.[1,11] Each row and column in the matrix was randomly illuminated for 100 ms. After illumination of a row/column for 75 ms, the matrix was blank. For each character, sets of 12 illuminations were repeated 15 times. So there were 180 illuminations for each character. After all illuminations, the matrix is blank for 2.5 s.[11]

To improve and validate signal processing and classification methods for BCIs, some BCI groups organized an online BCI data bank, known as the BCI competition datasets. These datasets consist of continuous single-trials of EEG activity, one part is training data and another part is test data, which is unlabeled. Initially, the labels for test data were not available for the purpose of the competition. The labels for testing sets are released and the data sets became available for developing new methods towards improving BCI studies. This paper uses EEG signals of the BCI competitions 2003 dataset-IIb which are recorded from a P300/ERP based BCI word speller.[1]

Reducing noise of signals will help P300 detection. We proposed methods to enhance EEG with adaptive WT via ICA concepts. It is demonstrated that channel Cz has mainly data and more accurate to detect P300, so most of the studies rely on this record.[9]

INDEPENDENT COMPONENT ANALYSIS

To alleviate the influence of noise, in this research, an independent component analysis ICA-based denoising scheme is proposed and integrated with adaptive wavelet denoising.

The methodology of Independent Component Analysis (ICA) was first introduced in the context of neural networks.[14] Delorme *et al*.[15,16] utilized ICA as an important algorithm for EEG analysis. The goal of ICA is to separate instantaneously mixed signals into their independent sources, without knowledge of mixing process. One practical application of ICA decomposition is the Evoked Related Potential and electroencephalogram analysis. The recorded signal can be considered as different sources in the brain and various artifacts that generate electrical signals. ICA is presented of interesting sources[17] or automated classification of epileptiform activity[18] and artifact removal.[19]

Hidden information is called the independent components (ICs) of the data. The noise information usually cannot be directly obtained from the observed data. Thus, ICA can be used to detect and remove the noise via the identification of the ICs of the time series data, and improve the performance of the specific patterns in brain activity.

Let $X = [x_1, x_2, x_3, ..., x_m]^T$ be a multivariate data matrix of size $m * n$, $m \leq n$ consisting of observed mixture signals x_i (size 1*n, i=1,2., n).

In the ICA model, the matrix X can be introduced as:[18,19]

$$x = AS = \sum_{i=1}^{m} a_i s_i \tag{1}$$

Where s_i is the i-th row of the m*n source matrix S and a_i is the i-th column of the m*m unknown mixing matrix. Vectors s_i are the latent source signals that cannot be observed from the mixture signals x_i. The ICA algorithm aims at finding an m*m de-mixing matrix W such as:

$$Y = [y_i] = wx \tag{2}$$

Where y_i is the i-th row of the matrix Y (i=1,2, … m). The vectors y_i must be as statistically independent as possible, and are called as independent components. When de-mixing matrix W completely is the inverse of mixing matrix A, ICs (y_i) can be used to estimate the latent source signal s_i.

The ICA algorithm is formulated as an optimization problem by setting up the measure statistical independence of ICs as an objective function. To do this, using some optimization techniques for solving the de-mixing matrix W is suggested. In general, the ICs are obtained by using the de-mixing matrix W to multiply the original matrix. This matrix can be determined by using an algorithm which maximized the statistical independence of ICs. The statistical independence obtained by ICs with non-gaussian distribution.[18] The negentropy can measure the non-Gaussianity of the ICs such as:

$$J(y) = H(y_{gauss}) - H(y) \tag{3}$$

Where y_{gauss} is a Gaussian random vector having the same covariance matrix like y. H is the entropy of random vector y defined by H with density p(y):

$$H(y) = -\int p(y) \log p(y) dy \tag{4}$$

The negentropy is zero if and only if the vector has a Gaussian distribution and it is always non-negative parameter. Since the problem in using negentropy is computationally very difficult, an approximation of this parameter is proposed as follows:[1,10,20]

$$J(y) \approx [E\{G(y)\} - E\{G(v)\}]^2 \tag{5}$$

Where E stand for entropy, v is a Gaussian variable of zero mean and unit variance, and y is a random variable with same mean and variance. G is a non-quadratic function, and is given by $G(y) = \exp(-y^2 / 2)$ in this study.[3]

We can find the de-mixing matrix with optimization. This maximization in described method, FASTICA, is achieved using an approximate Newton iteration. After every iteration, to prevent all vectors from converging to the same maximum, (that would yield several times the same source), the p-th output has to be de-correlated from the previously estimated sources. A deflation scheme based on a Gram-Schmidt orthogonalization is a simple way to do this.[9,10,21]

WAVELET TRANSFORM

It has been shown that EEG is a classical non stationary signal. Short time fourier transform (STFT), which was a time-frequency analysis method, was applied to analyze brain signals, but it has been noted that the transform depends critically on the window. Wavelet Transform (WT), that is a multi-resolution analysis method, brings solution to this task and give a more accurate temporal localization.[22] In the research of brain signal enhancement, a more accurate local band denoising required Wavelet Transform that could help to eliminate noise, which we were interested in, and that can generate spectral resolution.[22,23]

In particular, the Wavelet Transform forms a signal representation which is local in time and frequency domains. The WT relies on smoothing the time domain signal at various scales; thus, if $\psi_s(x)$ represents as wavelet at scales, the WT of such a function like $f(x) \in L^2(R)$ is defined as a convolution:[24]

$$Wf(s,x) = f * \psi_s(x) \tag{6}$$

The scaled wavelet are constructed from a ‘Mother’ wavelet, $\psi(x)$:

$$\psi_s(x) = (1/s)\psi(x/s) \tag{7}$$

With a Gaussian function, G(x), the ‘Mother’ wavelet will be defined as:[6]

$$\psi(x) = \frac{dG(x)}{dx} \tag{8}$$

From knowledge of the modulus maxima of WT, the signal f(x) may be reconstructed to a good approximation. These maxima propose a compact representation idea for noise removal.

The existing research uses a lipschits exponent method based on least square. Lipschitz exponent-α (L.E.α) is a criteria to quantize locally regulation of function in mathematics.[13,25] L.E.α, at one point, can reflect the signal singularity in that area of function. This means that its smoothness increases and signal singularity diminishes with α.[26]

Order $0 \le \alpha < 1$, and assuming that a constant (C) exists that makes:

$$\forall x \in R, \left| f(x) - f(x_0) \right| \le C \left| x - x_0 \right|^{\alpha} \tag{9}$$

The lipschitz exponent at x_0 point is called α. If the function is nth differentiable, but nth derivative of $f(x)$ is uncontinuously, the lipschitz exponent of $\int f(x)dx$ is up to $\alpha+1$ when $n \le \alpha < n+1$ (n being the nearest integer to α).[1,23]

So it is possible to characterize isolated singularities when they occur in a smooth signal. Furthermore, for such point like x in the neighbor of x_0, where f(x) is α-Lipschitz at x_0, the modulus maxima of the WT evolve with scale s, according to:

$$\left| Wf(s,x) \right| \le AS^{\alpha} \tag{10}$$

Where A is a constant.[1]

To determine α, WT of one typical EEG segment calculated and by formula (10) it's Lipschitz have to be zero [Figures 1-3].

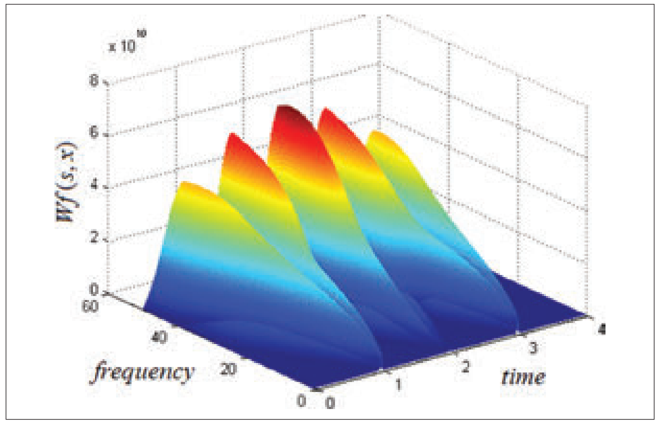

Figure 1: Wavelet transform of EEG

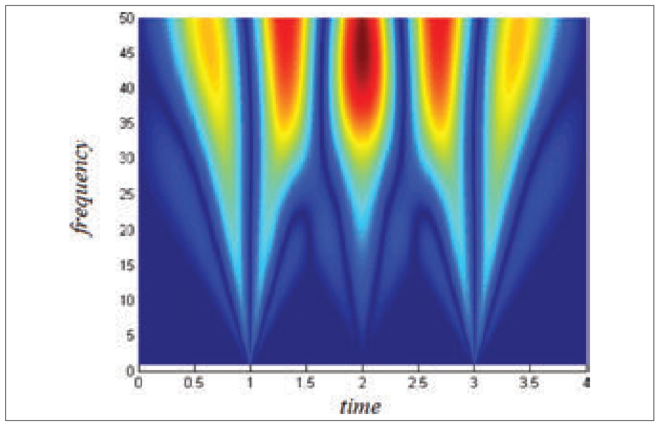

Figure 2: XY plane to know the point of singularites

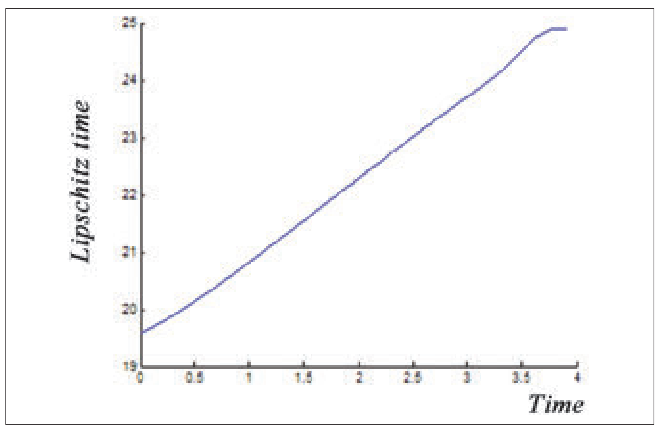

Figure 3: A+.5 line with m=1.4566 and a=0.9566; so we have a=0 for determining the wavelet function

NOISE REMOVAL

Adaptive Wavelet Denoising has many good results in biomedical signals, and ICA has many advantages in EEG denoising. Therefore, it is possible to combine them to reach the better SNR. The SNR was defined as the ratio of standard deviations of the clean ERP signal and one of the surrogate EEGs. Note that, for lower SNR, the ERPs are hardly recognizable in the single-trial. ICA approach helps us to use adaptive thresholding via wavelet to choose thresholds of denoising [Figure 4].

As shown in Figure 4, adaptive wavelet denoising is a method of noise removal. So we applied Wavelet Transform on Signal. Then, the level C_{-3}, the lowest frequency is kept, which is more noiseless and similar to original signal. Other levels (d_{-n}) must denoised with sufficient thresholds (such as T_n). With adaptive wavelet thresholding, user can denoised each level with interested threshold and function.[13]

To have WT coefficients only for singularities and no other points, a proper wavelet need to detect singularities in a signal (sufficient number of moments must be zero). EEG signal has peaks at some points, like discontinuous regions; so we choose α (lipschitz Regularity) to be zero.

$$\int \psi(t)dt = 0$$
$$\int t\psi(t)dt \neq 0 \qquad (11)$$
$$so\ n = 0 \quad and \quad \psi(t) = (-1)^n \frac{d^n\theta(t)}{dt^n} = \theta(t) = e^{-t^2/2\delta^2}, \delta = 1$$

Now, the signal has to be input of WT because SNR of EEG is low; we propose to have three analysis levels by soft thresholding. Usually the levels of transform can be changed (most three to five), but in more practical denoising, three levels work well. Our experiment certify this suggestion.[13,23]

Soft thresholding is a useful tool to denoising signals with low SNRs, but hard thresholding could generate more noise. It can be simply defined by[13] [Figure 5]:

$$\delta_{sof\ tthresholding}(d_J, \lambda) = (d_J - sign\lambda)\ \mathrm{II}(|d_J| > \lambda) \qquad (12)$$

where II (.) *is heavyside function*

d_J's are the elements in each level of WT and λ is the threshold for denoising each element.

We change above function fewness, to have a better signal power.

$$\delta_{soft\ thresholding(new)}(\mathrm{d_J}, \lambda) = \left(d_J - k * \frac{\lambda^2}{d_J}\right)\Pi\left(|d_J| > \lambda\right) \qquad (13)$$

where $\Pi(.)$ *is heavyside function*, $k = 2$.

By WT, we have subspaces of Cz channel complex with noise. ICA separate ICs by maximizing negentropy of sources. We want to use this idea to find proper threshold in each branch.

Applying some preprocessing steps before using an ICA algorithm is very useful to separate ICs. Therefore, we discuss some techniques that make the ICA estimation so simpler and in better condition.

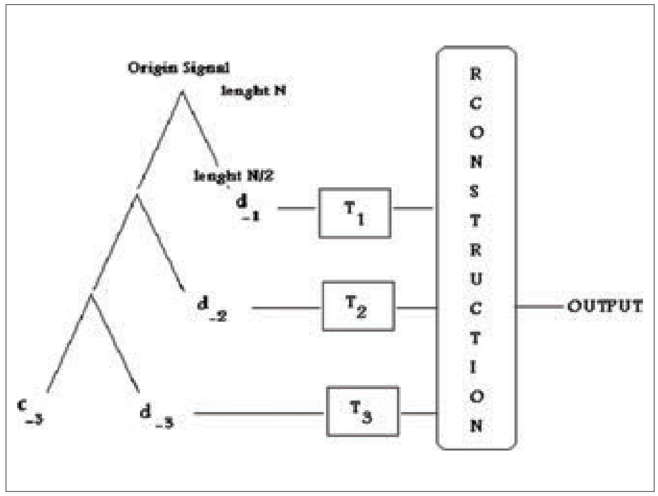

Figure 4: Adaptive wavelet thresholdings

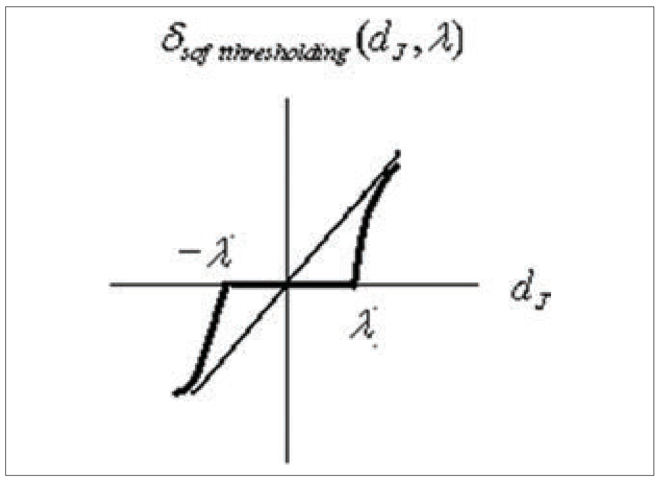

Figure 5: Function of $\delta_{soft\ thresholding(new)}$ ($\mathrm{d_J}$, λ)

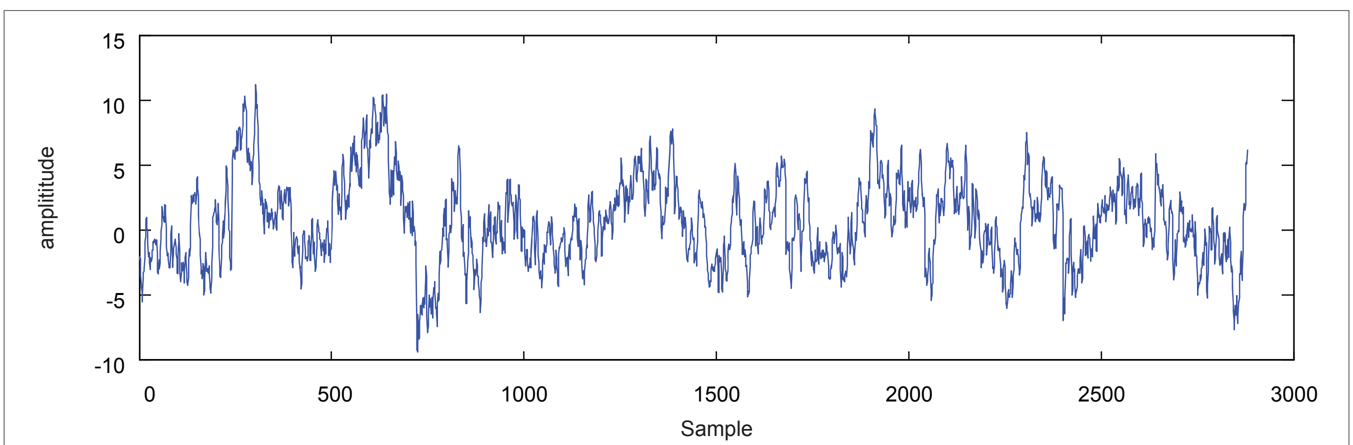

Figure 6: EEG signal via Cz channel in ERP

The most basic preprocessing step is to center x, i.e. to make x a zero-mean variable, subtract its mean vector m=$E\{x\}$. This process does not mean that the mean couldn't be estimated.

We can complete the estimation of mixing matrix A with centered data, by adding the mean vector of s, back to the centered estimates of it. $A^{-1}m$ gives the mean vector of s (in the preprocessing m was subtracted as the mean).[12]

Whitening the observed variables is another useful preprocessing strategy in ICA. This means that after centering and before the application of the ICA algorithm, we linearly transform the observed vector X to obtain a new white vector X, i.e. its components are uncorrelated and their variances equal one. In other words, the covariance matrix of X is the identity matrix:

$$E\left\{XX^{T}\right\}=I \tag{14}$$

Using the eigen-value decomposition (EVD) of the covariance matrix, $E\{xx^{T}\}=EDE^{T}$ is one of the popular method for whitening, where E is the orthogonal matrix of eigenvectors of $E\{xx^{T}\}$ and D is the diagonal matrix of its eigenvalues, D=diag($d1$.,., dn).[3,12] Note that $E\{xx^{T}\}$ can be estimated from the available sample x(1), ., x(T). We can see that whitening can now be done by:

$$\tilde{x}=Ed^{-1/2}E^{T}x \tag{15}$$

Where, the matrix $D^{-1/2}$ is computed by a simple component-wise operation as $D^{-1/2}=diag(d_{1}^{-1/2},...,d_{n}^{-1/2})$. The mixing matrix transforms into a new one, $\tilde{A}$, where it is orthogonal:[3,12]

$$\tilde{x}=ED^{-1/2}E^{T}As=\tilde{A}s \tag{16}$$

$$E\{\tilde{x}\tilde{x}^{T}\}=\tilde{A}E\{ss^{T}\}\tilde{A}^{T}=\tilde{A}\tilde{A}^{T}=I \tag{17}$$

Centering and whitening prepare subspace signals to denoising. In universal denoising, the parameter has to be introduced by:[13,23]

$$\lambda=\delta\sqrt{2*\log_{e}^{N}} \quad where \quad \delta=\frac{median\left(\left|d_{-1,k}\right|\right)}{0.6795} \tag{18}$$

Where in this function, δ stand for variance of noise, $d_{-1,k}$'s are the first subspace elements ($1\le K\le N$), N imply length of trail and λ is the threshold. In universal thresholding, the same threshold (above formula) was used for all sub-bands.

But we compute negentropy for each branch. One of the sub-bands that have lower frequencies is like to signal, so its negentropy is important. Other subspace negentropies have compared to this one and their differences change the threshold of denoising. To adaptive wavelet thresholding, SURE has introduced to fine the best threshold in each band. In this approach, threshold from zero to universal value changed and SURE like follow formula computed then threshold that associated minimum SURE determined and used for that band. We have improved SURE by negentropy feature. So the SURE can change like this:

$$SURE_{Convential}=N_{J}-2*\sum \mathrm{II}(\left|d_{J,K}\right|<\lambda)+\sum \min(\left|d_{J,K}\right|,\lambda)$$
$$N_{J}=lenght\ of\ signal$$
$$d_{J,K}=signal\ elements$$
$$\lambda=0\ to\ universal\ \lambda$$
$$SURE_{New}=N_{J}-2*\sum \mathrm{II}(\left|negentropy_{J,K}\right|<negentropy_{C}) +\sum \min(\left|d_{J,K}\right|,\lambda) \tag{19}$$
$$negentropy_{J,K}=negentropy\ of\ signal\ elements$$
$$negentropy_{C}=negentropy\ of\ C\ subbsnd$$
$$\lambda=0\ to\ universal\ \lambda*(1-N_{c}/N_{d})$$

If SNR is less, we can increase the Wavelet Transform and denoise each level accurately. But in practice, the WT levels for denoising are three to five.[13,23] In this application, we found that three levels could be sufficient. So negentropy of three levels wavelet decomposition from EEG were computed; then each one has denoised by adaptive soft thresholding obtained by their negentropy. These subspaces, reconstructed in time/domain transform. This approach suggests very good denoising method and SNR results.

Denoising Algorithm of EEG Signals via Adaptive Wavelet Thresholding by ICA Concepts

1. Averaging Cz (premier) channel after one complete

Table 1: Comparing different methods on achieved signals to noise ratios of some signals

	SNR of signal 1	SNR of signal 2	SNR of signal 3	SNR of signal 4	SNR of signal 5	SNR of signal 6	SNR of signal 7	SNR of signal 8	Averages of SNR's (with standard deviation)
ICA denoising method	4.4992	4.5524	4.5623	4.6095	4.5726	4.6102	4.5880	4.5941	4.5735 (0.036)
Wavelet denoising method	4.5411	4.5333	4.6402	4.6529	4.7203	4.7387	4.6219	4.6745	4.6403 (0.074)
Wavelet denoising method via ICA	4.6521	4.6578	4.7291	4.6899	4.8256	4.7126	4.8005	4.7565	4.7280 (0.063)

SNR - Signal to noise ratio; ICA - Independent component analysis

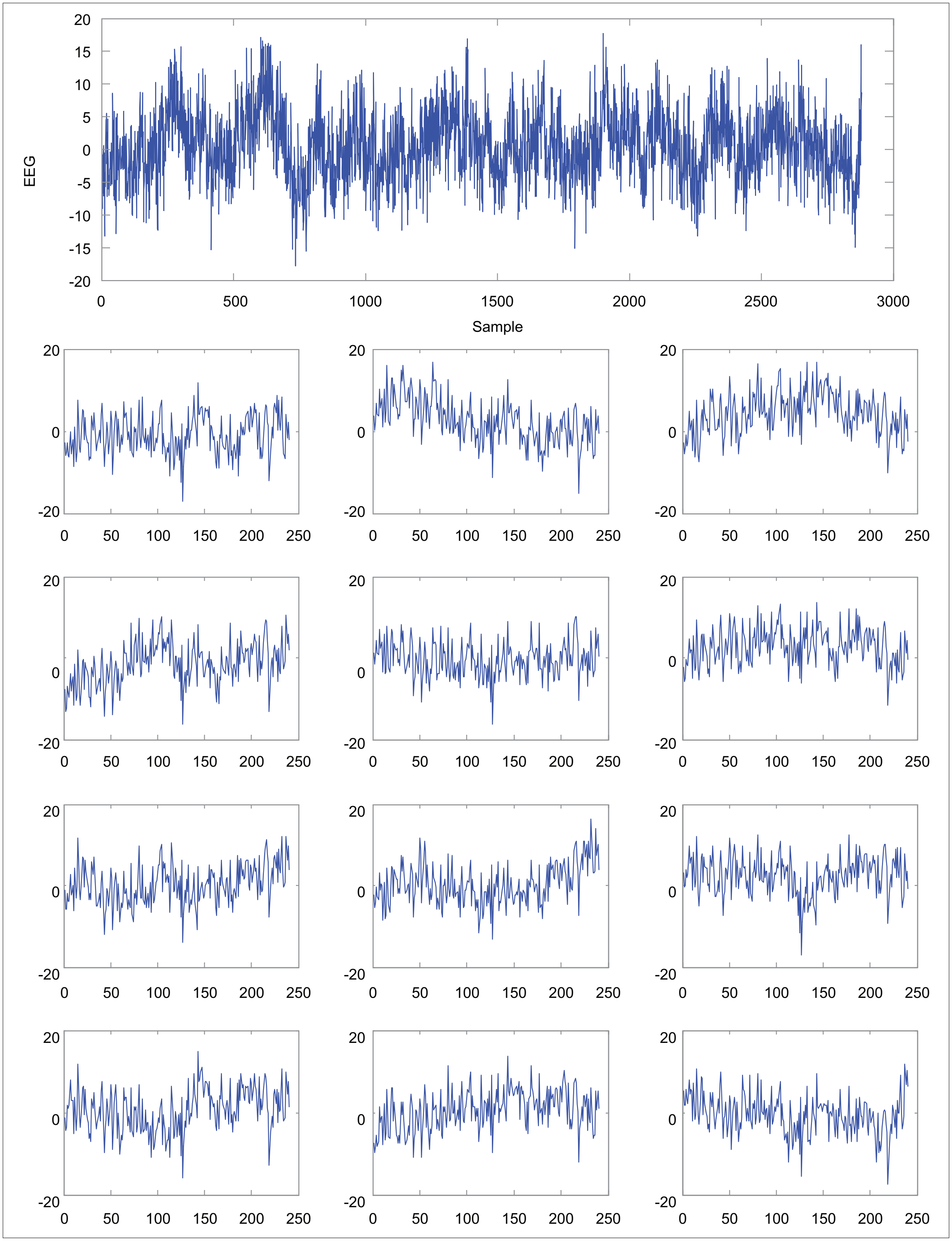

Figure 7: Noisy EEG of Cz channel (top) and its 12 segments (down)

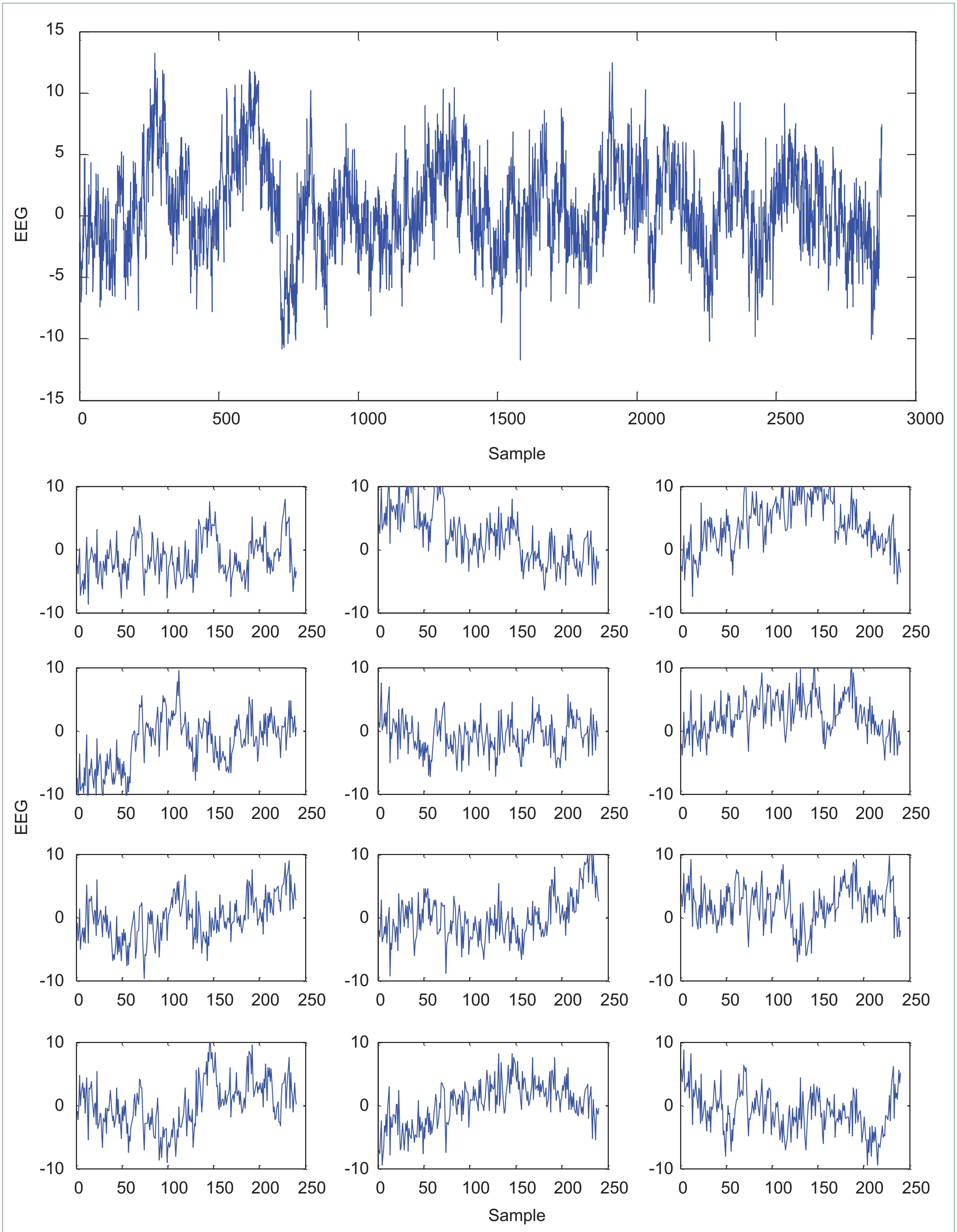

Figure 8: Denoised EEG of Cz channel by ICA (top) and its 12 segments (down)

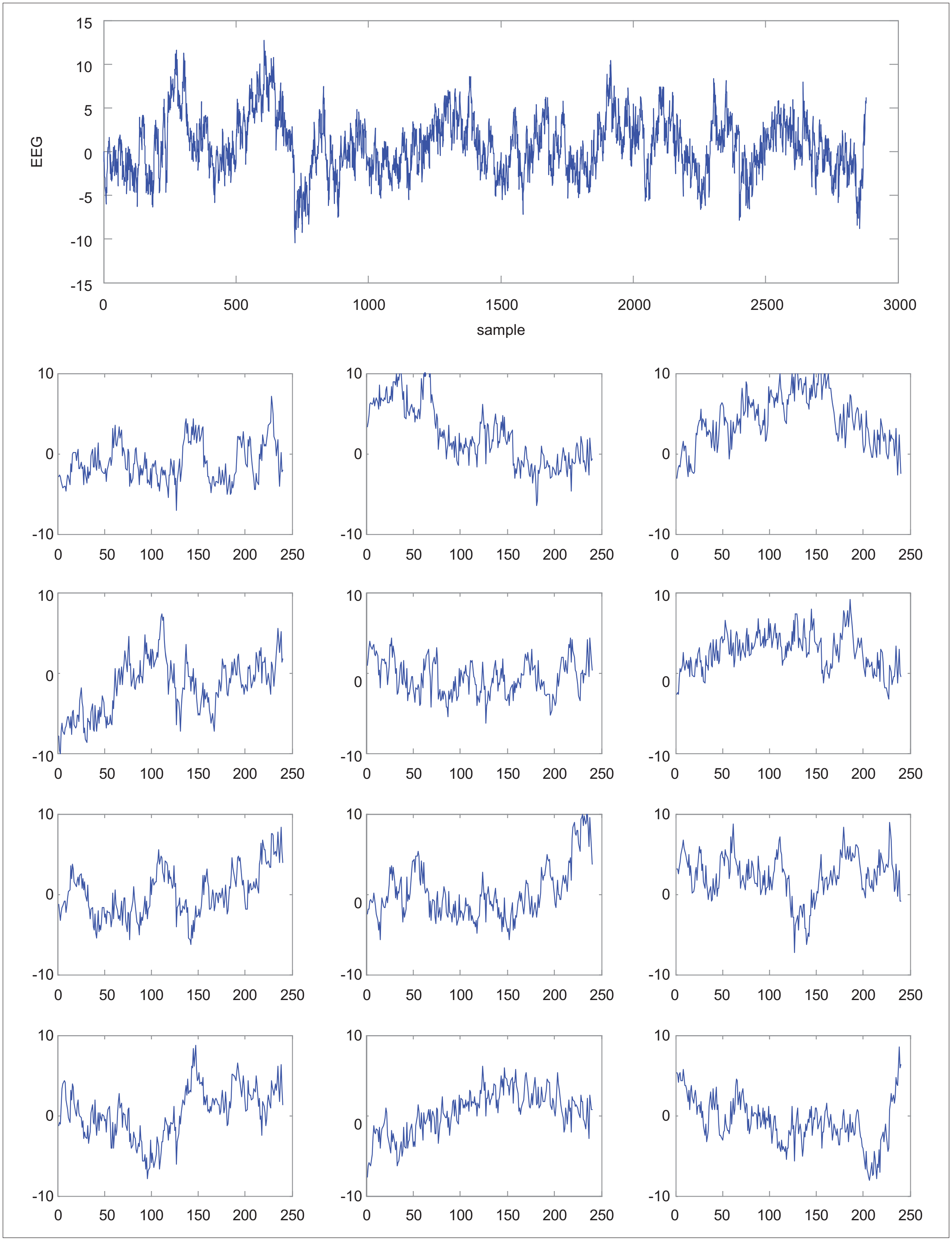

Figure 9: Denoised EEG of Cz channel by adaptive wavelet thesholdings (top) and its 12 segments (down)

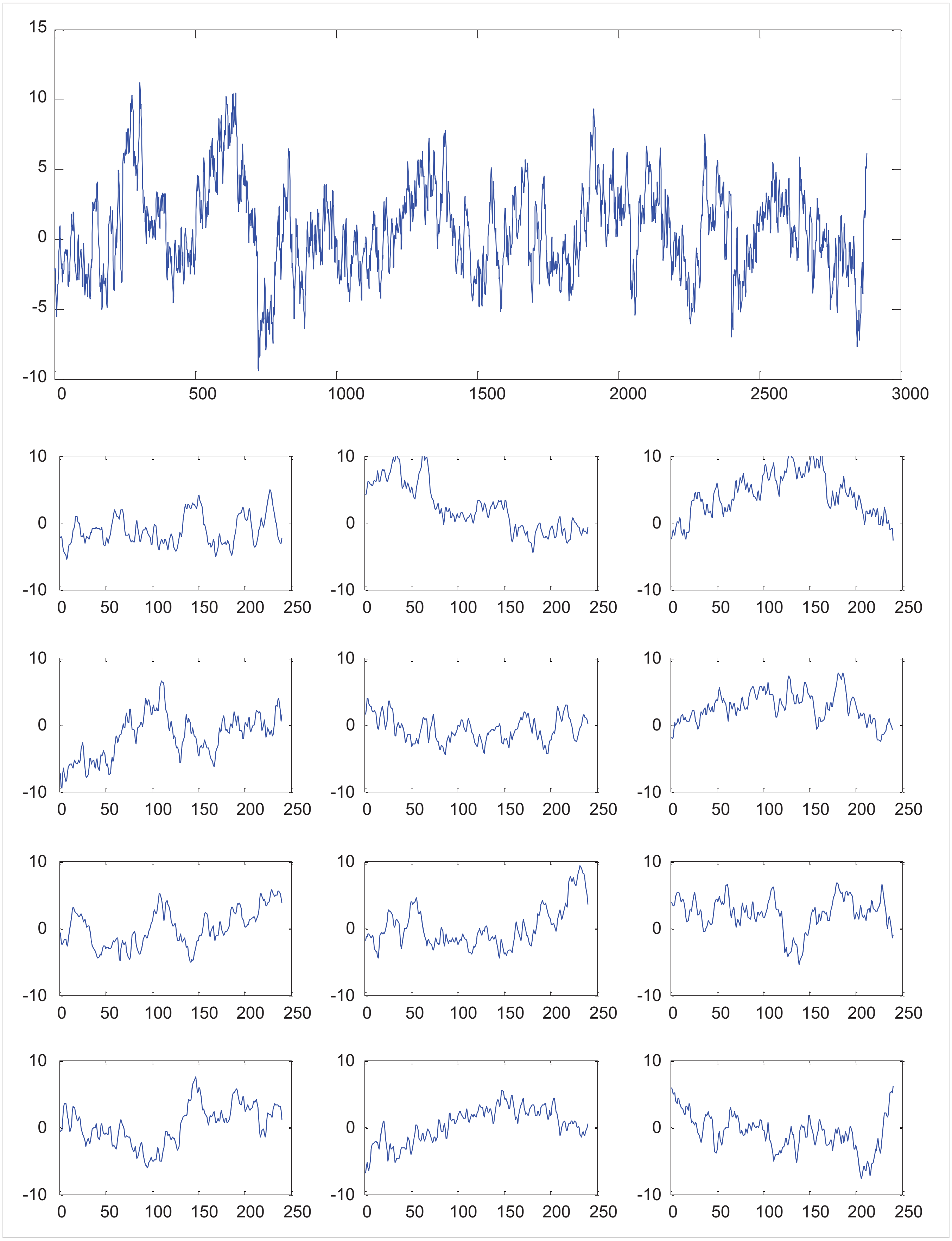

Figure 10: Denoised EEG of Cz channel by proposed algorithm (top) and it's 12 segments (down)

iteration with P300 speller
2. Seriate these 12 parts to one signal
3. Wavelet analysis on the signal 3 level (can be changed).
4. Computing negentropy of all subspaces and sequential differences
5. Determining adaptive threshold for denoising each level
6. Have a soft thresholding in each plane
7. Reconstruct EEG of denoised subspaces.

SIMULATION AND RESULTS

EEG signals in each iteration were averaged, and 12 averaged parts let to determine ERP Eliminating artifacts caused to develop clear EEG signals. This algorithm applied on BCI Competition 2003 database (IIb) and interesting results appeared [Figures 6-10]. In the EEG data with average SNR of 4.4667 dB, we have tested ICA, Wavelet Thresholding (Soft Thresholding with Universal criteria), and Adaptive Wavelet Thresholding via ICA concepts algorithm to compare them for enhancement of ERP detection. Sequentially, averaged SNR=4.5933 dB, SNR=4.6700 dB, and SNR=4.7442 dB obtained. Therefore, the proposed method could be more accurate to EEG enhancement. Some simulation results can be seen as follows in Table 1.

CONCLUSION

This paper has introduced a new algorithm to enhance ERP signals. SNR of EEG is very low. We aim to eliminate artifacts. We proposed a method based on adaptive wavelet thresholding via ICA concepts. Negentropy is the most important feature to diagnose noises. This method was compared with two usual algorithms such as ICA and WT. Algorithms were tested on signals from dataset BCI competition 2003. The proposed method could be more accurate to EEG enhancement.

REFERENCES

1. J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain-computer interfaces for communication and control. Clinical Neurophysiology 113, pp. 767-791, 2002.
2. L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature 442, pp. 164–171, 2006.
3. N. Xu, X. Gao, B. Hong, X. Miao, S. Gao, and F. Yang. BCI Competition 2003–Data Set IIb: Enhancing P300 Wave Detection Using ICA-Based Subspace Projections for BCI Applications. IEEE Trans Biomed Eng 51, pp. 1067-1072, 2004.
4. A. Rakotomamonji, V. Guigue. BCI COMPETITION III: Dataset II-ensemble of SVMs for BCI P300 speller. IEEE Trans. Biomedical Eng 55, pp. 1147-1154, 2008.
5. D. Lemire, C. Pharand, J.C. Rajaonah. Wavelet time entropy T wave morphology and myocardial ischemia. IEEE Trans Biomedical Eng 47, pp. 967-970, 2000.
6. E. Donchin, K.M. Spencer, R. Wijesinghe. The mental prosthesis: Assessing the speed of a P300-based brain computer interface. IEEE Trans Rehabilitation Eng 8, pp. 174-179, 2009.
7. E. Donchin, Y. Arbel. P300 based brain computer interface: A progress report. Lecture Notes in computer Science 5638, pp. 724-731, 2009.
8. L. Farwell, E. Donchin. Talking off the top your head toward a mental prosthesis utilizing event-related brain potentials. Electroencephalogr Clin Neurophysiol 70, pp. 510-523, 1988.
9. B. Blankertz, K.-R. Müller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schlögl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schröder, and N. Birbaumer. The BCI Competition 2003: Progress and Perspectives in Detection and Discrimination of EEG Single Trials. IEEE Trans. Biomed. Eng 51, pp. 1044–1051, 2004.
10. T.-P. Jung, S. Makieg, C. Humphries, T.-W. Lee, M. J. Mckeown, V. Iragui, and T. J. Sejnowski. Removing electroencephalograohic artifacts by blind source separation. *Psycophysiology* 37, pp. 163-178, 2000.
11. E. Donchin, K. Spencer, and R. Wijesinghe. The mental prosthesis: Assessing the speed of a P300-based brain-computer interface. IEEE Trans Rehabil Eng 8, pp. 174-179, 2000.
12. E. Nezhadarya, M. B. Shamsollahi. EOG Artifact removal from EEG using ICA and ARMAX modeling, Proc. of the 1st United Arab Emirates, Int. Conf. of the IEEE BMP, pp. 145-149, Mar 27 -30, 2005.
13. D.L. Donho. De-noising by soft thresholding. IEEE Trans Information theory, Vol.41, pp. 613-627, 1995.
14. A. Hyvarinen, J. Karhunen. Independent Component analysis. Wiley: NewYork; 2001.
15. A. Delorme, S. Makeig. An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134, pp. 9-21, 2004.
16. A. Delorme, S. Makrig. EEG changes accompanying learned regulation of 12 Hz EEG activity. IEEE Trans, Neural Systems and Rehabilitation 11, pp. 133-137, 2003.
17. U. Wiklund, U. Karlsson, Ostlund N, L Berglin, K Lindecrantz, S Karlsson, and L Sandsjö. Adaptive spatio-temporal filtering of disturbed ECG's: A multichannel approach to heart-beat detection in smart clothing. Med Biol Eng Comput 45, pp. 515-523, 2007.
18. M. De Lucia, J. Firtchy, P Dayan, DS Holder. A novel method for automated classification of epileptiform activity in the human electroencephalogram-based on independent component analysis. Med Biomedical Eng Comput 46, pp. 263-272, 2008.
19. M. Milanesi, N. Martini, Vanello N, V Positano, MF Santarelli, and L Landini. Independent component analysis applied to the removal of motion artifacts from electrocardiographic signals. Med Biol Eng Comput 46, pp. 251-261, 2008.
20. RQ. Quiroga, O.A Rosso, Başar E, Schürmann M. Wavelet entropy in event-related potentials- a new method shows ordering of EEG oscillations. Biological Cybernetics 84, pp. 291-299, 2001.
21. C. Gouy-Pailler, M. Congedo, C. Jutten, C. Brunner, and G. Pfurtscheller. Model-based source separation for multi-class motor imagery. In Proceedings of the 16th European Signal Processing Conference (EUSIPCO-2008), EURASIP, Lausanne, Switzerland, Aug 2008.
22. H. A. Al-Nashah, J. S. Paul. Wavelet entropy method for EEG application to global brain injury, proceeding of the 1st international IEEE EMBS, Neural Systems and Rehabilitation: Italy; pp. 348-352, 2003.
23. J. Blaszezuk, Z. Pozorski. Application of the lipschitz exponent and the wavelet transform to function discontinuity estimation. Institute of Mathematics and Computer Science, Vol. 1, No. 6, pp.23-28, 2007.
24. J. Goswami, T. Jaideva. Fundamentals of Waveletes theory, Algorithms and application, (Wiley Series in Microwave and Optical Engineering), 1999.
25. Y. Tang, L. Yang. Characterisation and detection of edges by lipschitz

exponetnts and MASW wavelet transform. Proceeding of 14th international conference on Pattern Recognition: Australia; Vol. 2, pp. 1572-1574, 1998.

26. A. Quinquis. Few practical applications of wavelet packets source. Digital Signal Processing. A review Journal 8, pp. 49-60, 1998.

How to cite this article: ***

Source of Support: Nil, **Conflict of Interest:** None declared

BIOGRAPHIES

Zahra Vahabi received her B.S. and M.S. in Biomedical Engineering. She is currently pursuating Ph.D. in Electrical Engineering. Her main research of interests cover a wide range of topics in digital signal processing, digital image processing and Biomedical signal and image analysis.

Rassoul Amirfattahi was born on 1969. He received his B.S. in Electrical Engineering from Isfahan University of Technology in 1993, M.S. in Biomedical Engineering and Ph.D. in Electrical Engineering both from Amirkabir University of Technology (The Tehran Polytechnic) in 1996 and 2002 respectively. From 2003 he has joined department of Electrical and Computer Engineering at Isfahan University of Technology while he is currently Associate Professor of ECE department and head of Digital Signal Processing research Lab. Dr. Amirfattahi is author and co-author of more than 100 technical papers, Industrial reports and book chapters and has two new academic books under publishing. His main research of interests cover a wide range of topics in digital signal processing, speech and audio processing, digital image processing and Biomedical signal and image analysis.

Abdolreza Mirzaei is an assistant professor in the Electrical and Computer Engineering Department at Isfahan University of Technology, Iran.

His research lies in the broad area of Vision and Pattern Recognition with particular interests on statistical and structural classification methods, multiple classifier systems, and learning methods.

He holds a M.Sc. in artificial intelligence from Iran University of Science and Technology, Tehran, Iran(2003) and a PhD in artificial intelligence from Amirkabir University of Technology, Tehran, Iran (2009).