

Artifacts Removal of EEG Signals using Adaptive Principal Component Analysis

Arjon Turnip*, Dwi Esti Kusumandari

Technical Implementation Unit for Instrumentation Development, Indonesian Institute of Sciences, Bandung, Indonesia.

Received: 18 September 2014 / Accepted: 30 November 2014

Abstract:

Analysis of EEG activity usually raises the problem of differentiating between genuine EEG activity which is introduced through a variety of external influence. These artifacts may affect the outcome of the EEG recording. In this paper, wavelet denoising and band pass filter for preprocessing and an adaptive principal component analysis based recursive least squares algorithm for extraction are proposed to remove the artifacts. The algorithm is designed to adaptively derive a relatively small number of decorrelated linear combinations of a set of random zero-mean variables while retaining as much of the information from the original variables as possible. The proposed method was tested in real EEG records acquired from eight subjects. The experimental result show that the proposed method can effectively remove the artifacts from all subjects.

Key words: *Artifacts, EEG, noise, principal component analysis*

Introduction

When brain cells (neurons) are activated, the synaptic currents are produced within the dendrites. This current generates a magnetic field measurable by electromyogram (EMG) machines and a secondary electrical field over the scalp measurable by Electroencephalogram (EEG) systems [1]. EEG is the non-invasive measurement of the electrical activity on the scalp over multiple areas of the brain. The measured of currents that flow during synaptic excitations of the dendrites of many pyramidal neurons in the cerebral cortex is called EEG signal. In clinical application, EEG signal used for diagnosing, monitoring, and managing neurological disorders and in research field used for brain computer interface (BCI) application. In recording EEG Signals, subjects normally will blink, squirm, and glance about, as they would be expected to sit in a chair for a long time and engage in what is often a repetitive task. Unfortunately, through these movements a period of electrical noise or artifacts that may be difficult to discriminate from neural activity may be introduced. Hence, a necessary stage in EEG processing is artifact removal. Artifacts can dramatically alter the signal recorded at all scalp sites, especially those closest to the source of the noise [2-7]. The experimenter may also wish to remove (reduce) portions of the EEG record for other reasons, such as the subject becoming drowsy or failing to perform a task properly. Hence, it is often necessary to view the EEG data, on either a monitor or printout, and manually determine which regions of the data contain excessive artifacts. This is where signal processing becomes relatively complicated. Determining what is considered artifact, how much artifact is excessive, and removing artifacts from real data require a skilled human or a sophisticated artificial system [7].

In a BCI application that removed all data containing artifact might be left with too little clean data to be of practical use. Therefore, the use of noise cancelation method to the EEG records and leaving clean data would be of tremendous value (there has been an ample amount of research toward this goal) [8-10]. Many of these newer approaches involve techniques including independent component analysis, neural networks, Kohonen maps, principal component analysis (PCA), and other methods which were either unavailable or much less well known during the early days of EEG signal processing. Techniques for online cancelation of EEG noise without rejecting other data are remains difficult in EEG processing (APCA) software but necessary one for BCI designers. In this paper, an adaptive principal component analysis algorithm is proposed to remove the artifacts. The algorithm which can significantly enhance prominence of the spike in the clean EEG signals is designed to adaptively derive a relatively small number of decorrelated linear combinations of a set of random zero-mean variables while retaining as much of the information from the original variables as possible. The proposed method was tested in real EEG records acquired from eight subjects. The experimental result show that the proposed method can effectively remove the artifacts from all subjects.

Experimental

In the experiment, eight healthy adult subjects (all men, ranging in age between 20-22 old years) participated in this study, which are 5 subjects have slight hair and the rest are thick hair.

*Corresponding author: Arjon Turnip,
E-mail: arjon.turnip@lipi.go.id

All of the subjects do same experiment with three stimulus which are baseline, closed eye, and blink eye condition. It has been made sure that every participant was not in a stress mental condition while the experiment was being conducted. All participant picture show in Figure. 1. During the experiment, all participants were sitting in a comfortable chair in front of 14" monitor at a distance of about 1 m. To build the experiment setting used in the EEG sample collecting process, we utilize the OpenVibe software to perform the data acquisition, stimulus visualization, and EEG recording with the function block that already built in. This experiment consists of two sessions, first session had a period of 130 seconds with 2 stimulus (i.e., baseline (eye normal) and closed eye condition). Second sessions had a period of 66 seconds with 2 others stimulus (i.e., baseline (eye normal) and blink eye condition). The EEG signals are recorded continuously using six electrodes (channels) at F7, F8, T7, T8, O1 and O2 that represent the visual of human brain and digitized at a 128 Hz sampling rate.



Figure 1. Eight subjects participated in the experiment.

Let m be the number of measured EEG signals and n be the number of unknown input sources. Then, the measured signals at channel i , $x_i(k)$, can be represented as a linear combination of n unknown mutually statistically independent source signals $s_j(k)$, $j=1, 2, \dots, n$, as follows independent source signals

$$y_i(k) = w_i^T x_i(k) = \sum_{p=1}^m w_{ip} x_p(k) \quad (1)$$

where w_i are the separation matrix, whose values are updated through the APCA learning rule. The optimal weights correspond to the statistical independence of the output signals $y_j(k)$. The two random variables $y_i(k)$ and $y_j(k)$ are statistically independent if the values of $y_i(k)$ has no information about the values of $y_j(k)$. Mathematically, the independence of $y_i(k)$ and $y_j(k)$ can be expressed by the relationship $p(y_i(k), y_j(k)) = p(y_i(k))p(y_j(k))$, where $p(y)$ indicates the probability density function of the random variable y . Through equation (1), the i th principal component of the measured signals are extracted and next the source

signals is reconstructed from $y_i(k)$ by using the transformation $\hat{x}_i(k) = w_i y_i(k)$. The weight vector $w_i = [w_{i1}, w_{i2}, \dots, w_{im}]^T$ should be adaptively updated in such a way so that the estimated vector $\hat{x}_i(k)$ will reconstruct the input measured signals $x_i(k)$ as correctly as possible. The general learning algorithms are optimal only for a Gaussian distribution of the input data and they are rather sensitive to impulsive noise or outliers. Many approaches can be taken to increase the robustness of PCA with respect to noise and outliers. Firstly, outlying measurements can be eliminated from the data; secondly outliers can be suppressed or modified by replacing them with more appropriate values; and finally, more robust criteria can be applied.

If the independent signals are zero-mean, the generalized covariance matrix of $f(y_i(k))$ and $g(y_j(k))$, where $f(y)$ and $g(y)$ are different and odd nonlinear activation functions such that $f(y) = y^3$ and $g(y) = \tanh(y)$ is a non-singular diagonal matrix $R_{fg} = E\{f(y)g^T(y)\} - E\{f(y)\}E\{g^T(y)\}$. Therefore, the covariance $E\{f(y_i)g^T(y_j)\} - E\{f(y_i)\}E\{g^T(y_j)\}$ are all zero and all variances of $E\{f(y_i)g^T(y_i)\} - E\{f(y_i)\}E\{g^T(y_i)\}$ are all non-zero. In order to derive an adaptive algorithm, we can formulate a cost function as [11]:

$$J_p(w_i) \triangleq \rho(e_i) = \sum_{p=1}^m \rho_p(e_{ip}), (i=1, 2, \dots, n) \quad (2)$$

where $e_i = e_{i-1} - w_i f(y_i)$, with $e_0(t) \triangleq x(t)$. The minimization of these cost functions by the gradient descent technique leads to an adaptive learning algorithm

$$\frac{dw_i}{dt} = \mu_i [f(y_i) g^T(y_i) + e_{i-1} w_i^T g^T(y_i)] w_i \quad (3)$$

for any $w_i(0)=0$, ($i=1, 2, \dots, n$), where $\mu_i(t) > 0$. Usually, the second term in the right hand side of Equation (3) is relatively small and can be neglected. Thus, yielding a simplified version of the learning algorithm for the extraction of the first m PCs:

$$\frac{dw_i(t)}{dt} = \mu_i(t) f(y_i) g^T(y_i) w_i \quad (4)$$

or in a discrete-time version as

$$W(k+1) = W(k) - \mu(k) f(y(k)) g^T(y(k)) W(k) \quad (5)$$

where $g^T y = (f^T(y(k)) - y^T(k))$ with $w_i(0) \neq 0, \mu_i(k) > 0$. Since the separation matrix $W(k)$ is

assumed to be orthogonal (i.e., $W^T(k)W(k)=I$), the real-time adaptation rule can be rewritten as

$$W(k+1)=W(k)+\mu(k)f(y(k))\left[u^T(k)-f^T(y(k))\right]W(k), \quad (6)$$

where $y(k)$ is the separated signal and the output of the second step, $\mu(k)$ is the learning parameter (it is adaptively adjusted according to $\mu(k)=1/\left\{\gamma/(\mu(k-1))+\|y(k)\|_2^2\right\}$ with the forgetting factor $0 < \gamma < 1$), and $f(\cdot)$ is a suitably chosen nonlinear function that is usually selected to be odd in order to ensure both stability and signal separations. These nonlinear functions require the use of high-order statistics. In the present study, $f(\cdot)$ was chosen as $f(k)=\tanh(k)$. Finally, since $f(k)=dg(k)/dk$, $g(k)=\ln[\cosh(k)]$.

Results and Discussion

In this simulation, actual signals were recorded in a six-channel (F7, F8, T7, T8, O1 and O2) configuration. The raw data (see Figure. 2 which each block indicated: normal, closed, normal, and blinking eyes, respectively) were first pre-processed using a wavelet denoising. Wavelet method is performed to obtain signal finer and enhance noise removal of artifacts by eliminating high frequency in order to obtain the final result of clean EEG signals with amplitude 2-6 μV range of values where the value is smaller than the signal Raw EEG signal. So the value of the amplitude of the wavelet denoising result is closer to the EEG signal amplitude range. The preprocessed result is shown in Figure. 3. The denoised EEG signals, therefore, were filtered using a sixth-order band-pass filter with cut-off frequencies of 3 Hz and 13 Hz, respectively (see Figure. 4). Since the EEG signal power to the noise power ratio is very small, a method of extraction the event-related potential (ERP) component from the EEG is desirable. One way of gaining further insights into EEG signals is by applying adaptive PCA techniques. The extracted signal is given in Figure. 5.

Since the artifacts can randomly occur and are unexpected, they are difficult to identify. Thus, instead of detecting and removing artifacts, our approach is to extract the event-related components based on a global pattern that encapsulates models for signals of interest. Neuroscientists have found that, for the inhibition task, most event-related components project strongly around Fz channel and have two fixed significant ERP patterns [12]. The signal before and after our artifact removal techniques are shown in Figures. 2 and 5 from the 1st subject. As shown in the diagram, after artifact separation, a view of ERP is clearly recognizable in the each channels except channels P7 and P8 (i.e., contaminated by noise or artifacts). It demonstrates that the proposed algorithm can effectively extract the ERP

from even when there are no clearly visible ERPs in the original signals.

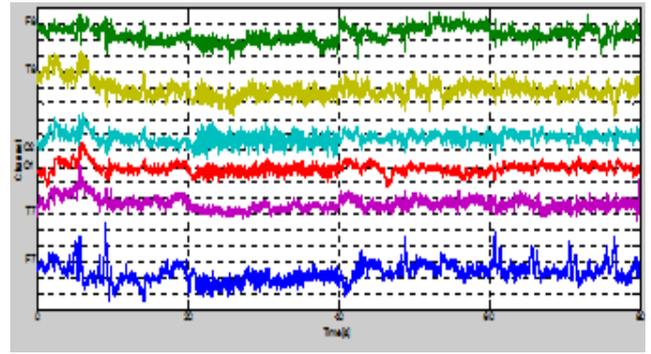


Figure 2. Recorded EEG signals from 1st subject.

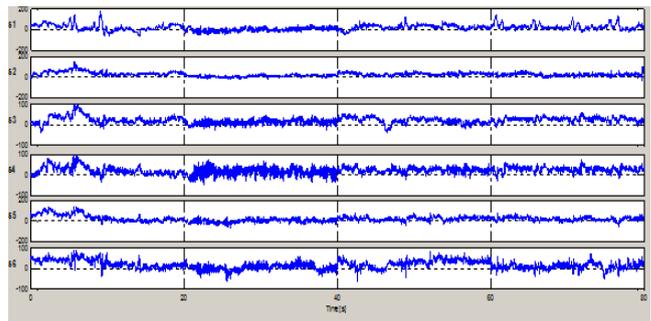


Figure 3. EEG signals preprocessed using the Wavelet denoised.

In order to show the visually performance of the extracted EEG signals, the brain mapping process is applied. The brain map using proposed methods according to the extracted signals in Figure. 5 is given in Figure. 6. The active brain regions that have been separated on any channels are indicated with a yellow to red. The red color indicated the higher brain activity and the blue color indicate the lower brain activity. The brain maps using PCA method indicates that the ocular artifacts are not perfectly separated (especially on channels F8 and O1).

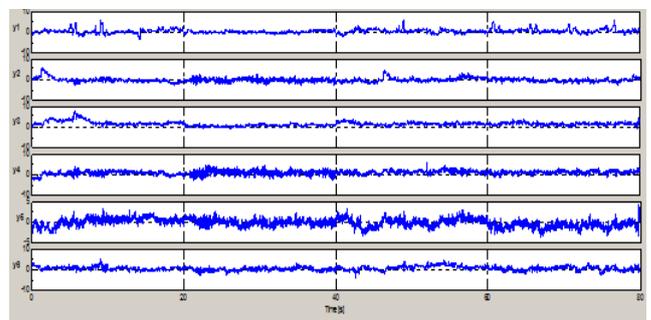


Figure 4. Filtered EEG signals using BPF.

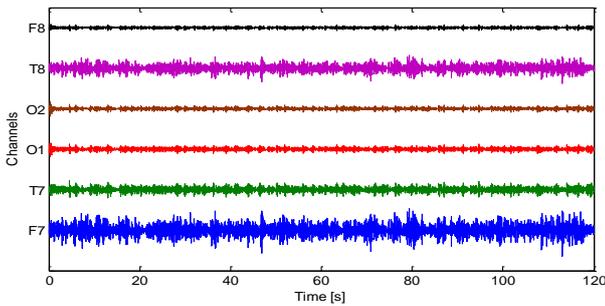


Figure 5. Clean EEG signals using APCA based moving average.

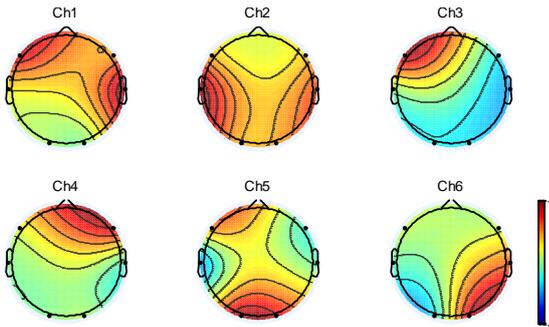


Figure 6. Corresponding spike of waveforms of extracted EEG signals.

Conclusions

In this paper, an adaptive PCA algorithm to automatically extract event-related components and refine raw EEG signals has been addressed. The superiority of this method is adaptive, identifies the artifacts in a sufficiently small number of iterations and tracks successfully changes in the source signals. Finally, the results using the proposed illustrate the effectiveness of the proposed algorithm removing the artifacts and other non-event-related sources, and increasing the visibility of the ERPs on all subjects.

Acknowledgement

This research was supported by the thematic program through the Bandung Technical Management Unit for Instrumentation Development, Deputy for Scientific Services, funded by Indonesian Institute of Sciences, Indonesia.

References

- [1] S. Sanei and J.A. Chambers, EEG Signal Processing, USA: John Wiley and Sons, 2007.
- [2] A. Jafarifarmand and M.A. Badamchizadeh, Artifacts removal in EEG signal using a new neural network enhanced adaptive filter, *Neurocomputing*, **103**, 2013, 222–23.
- [3] G. Geetha and S.N. Geethalakshmi, Artifact removal from eeg using spatially constrained independent component analysis and wavelet denoising with Otsu's thresholding technique, *Procedia Engineering*, **30**, 2012, 1064 – 1071.
- [4] S. O'Regan, S. Faul, and W. Marnane, Automatic detection of EEG artefacts arising from head movements using EEG and gyroscope signals, *Medical Engineering & Physics*, **35**, 2013, 867–874.
- [5] K.J. Mullinger, J. Havenhand, and R. Bowtell, Identifying the sources of the pulse artefact in EEG recordings made inside an MR scanner, *NeuroImage*, **71**, 2013, 75–83.
- [6] M.E.H. Chowdhury, K. J. Mullinger, P. Glover, and R. Bowtell, Reference layer artefact subtraction (RLAS): A novel method of minimizing EEG artefacts during simultaneous fMRI, *NeuroImage*, **84**, 2014, 307–319.
- [7] B. Allison, P3 or not P3: Toward a better P300 BCI. University of California, San Diego, 2003.
- [8] A. Turnip and K.-S. Hong, Classifying mental activities from EEG-P300 signals using adaptive neural network, *Int. J. Innov. Comp. Inf. Control*, **8**(7), 2012.
- [9] A. Turnip, K.-S. Hong, and M.-Y. Jeong, Real-time feature extraction of P300 component using adaptive nonlinear principal component analysis, *BioMedical Engineering OnLine*, **10**(83), 2011.
- [10] K.I. Diamantaras and S.Y. Kung, Principal Component Neural Networks. Theory and Applications. Adaptive and Learning Systems for Signal Processing, Communications and Control. John Wiley & Sons Inc., New York, 1996.
- [11] A. Cichocki and R. Unbehauen, Robust estimation of principal components in real time, *Electronics Letters*, **29**(21), 1993, 1869–1870.
- [12] M. Falkenstein, J. Hoormann, and J. Hohnsbein, ERP components in Go/Nogo tasks and their relation to inhibition, *Acta Psychologica*, **101**(2-3), 1999, 267-291.