EYE BLINKS REMOVAL IN SINGLE-CHANNEL EEG USING SAVITZKY-GOLAY REFERENCED ADAPTIVE FILTERING: A COMPARISON WITH INDEPENDENT COMPONENT ANALYSIS METHOD

Faridah Abd Rahman¹,² and Mohd Fauzi Othman¹
¹Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
²Kulliyyah of Engineering, International Islamic University of Malaysia, Kuala Lumpur, Malaysia
E-Mail: faridahrman@gmail.com

ABSTRACT
Eye blink artifact is one of the major problems in electroencephalograph (EEG) signals which mainly affected a frontal channel. A frontal channel often involved in recent applications of portable EEG devices which require a real time processing including for artifact removal. In this paper, we proposed a new referencing method in adaptive filtering for eye blinks removal of a single-channel EEG. The proposed method adopts Savitzky-Golay (SG) filter to extract the blink components from noisy EEG signals. The extracted component is then employed in adaptive filter as a reference input. We implemented adaptive neural fuzzy inference system (ANFIS) algorithm in adaptive filtering for the blink removal process. The reliability of the proposed method is demonstrated on real EEG dataset. By using the signal to noise ratio (SNR), mean squared error (MSE) and correlation coefficient as performance indicators, the proposed method is compared to independent component analysis (ICA), one of the widely accepted methods for artifact removal. The results show a low correlation between a corrected signal and a measured electrooculograph (EOG) signal, which indicates its efficiency in estimating and removing the blinks from the measured EEG signals. The results also demonstrate an improved performance compared to conventional ICA method.

Keywords: eye blinks, EEG, Savitzky-Golay filter, ANFIS, adaptive filtering.

INTRODUCTION
Electroencephalograph (EEG) is a medical imaging technique used to measure scalp electrical activity. EEG is attached to a head surface when measuring the brain waves, thus this technique is totally non-invasive, painless and safe for human. An array of sensors from metal electrodes and conductive media that cover the scalp surface, read the voltage fluctuation which is generated by the activated neurons in the cerebral cortex (S. Sanei & Chambers, 2007). A system in EEG device is actually measures the potential difference between two sensors, which is either between two active sensors, or between the active sensor and the reference sensor. One of the widely recognized methods for sensor or electrode placement is a 10-20 system established by the International Federation of Electroencephalography and Clinical Neurophysiology in 1958. The number of ‘10’ and ‘20’ refer to the percentage of actual distance between the adjacent electrodes to the total front-back or right-left distance of the skull. The front-back distance is pointed from nasion (point between the forehead and nose) to inion (the lowest point of the skull behind neck), while right-left distance is measured between both preauricular points. Figure-1 shows the example of electrode placement using 10-20 international standard placement system (Fisch, 1999). Brain can be divided into four main lobes which are frontal lobe, parietal lobe, temporal lobe and occipital lobe. Each lobe is named as Fx, Tx, Px and Ox respectively. X is a number to represent the exact location.

Figure-1. 10-20 international standard for electrode placement system.

EEG SIGNALS CHARACTERISTIC
The amplitude of EEG signals in normal adult typically ranging from 1 uV to 100 uV. Meanwhile, the frequencies from various regions of the brain could lie between 0.5 Hz to 100 Hz. Generally, the type of brain waves can be divided into five groups according to their frequency band. There are gamma (31 Hz -100 Hz), beta (13 Hz – 30 Hz), alpha (8 Hz – 13 Hz), theta (4 Hz – 7 Hz) and delta (0.5 Hz – 4 Hz) (Sanei & Chambers, 2007; Schomer, 2007). The waves rhythms correlated with the attention level of the brain. Typically, higher frequencies band representing the focused activity while lower frequencies indicate strongly synchronized activity such as during relaxing, and sleeping. Lower frequency in delta band sometimes could represent the injury or seizure occurred in that particular brain region.
EEG has been widely used in a brain study due to its high temporal resolution property and relatively low cost compared to other brain mapping techniques. The ongoing development in signal processing brought the use of EEG to a broader applications, from clinical studies (Striano et al., 2003) to real time applications such as neurofeedback (Alhaddad et al., 2012) and brain-computer interfaces (BCI) (Tomita et al., 2014). Despite its broad utilization, the EEG signals suffer from the contamination of noises from various sources. The noises, which are also known as artifacts, can be classified into two categories; environmental artifacts and biological artifacts. The biological artifacts include electrooculogram (EOG) artifact, caused by eye movement and blinking, electromyogram (EMG) artifact by muscle activity, electrocardiogram (ECG) artifact by heart beats, and many others. Among those, the interference of eye blinks artifact becomes major problem that cause a serious distortion to EEG signals which could bias the analysis and the interpretation of brain signal (Saeid Sanei & Chambers, 2007). In fact, the magnitude of eye blink can goes up to hundreds times larger than the EEG signals (Croft & Barry, 2002). The blink artifacts particularly affect the frontal channel severely because it is located near to the eyes. The example of raw EEG signals with eye blinks artifact affected the frontal channel is shown in Figure-2. The graph in the figure is plotted based on a publicly available dataset (Oliver & Suendermann, 2013). The recording was performed using Emotiv EPOC headset. A distortion in EEG signals appears in frontal channel, F7 and F8, around between sample 150 and sample 250 represent the eye blink artifacts. Note that there is no distortion in EEG signals for T7, O1 and P7 at the same sample number indicate that those sensors are less affected by eye blinks.

![Figure-2. Example of eye blink artifact in the frontal channel.](image)

The frontal channel of EEG carries important information of frontal lobe functions such as thought, voluntary movement, and emotions, which is often involved in problem solving based applications (Srinivasan, 2007). Thus, automatic blinks removal has become a vital step in preprocessing considering the recent focus on portable single-channel EEG devices which merely consist of frontal channel.

RELATED WORKS

In recent years, numerous techniques have been developed to remove the artifacts in EEG signals. Independent component analysis (ICA) is a widely accepted component-based method for artifact removal especially for EOG artifacts (Jung et al., 2000). ICA separates multi-channel observed signals into separated independent sources. ICA performs the separation based on the assumption that the recorded EEG signals are linear mixture of unknown independent component (Jung et al., 2000). ICA shows a great performance in separating the artifact from EEG signal into different components for a large-size data; however it is inapplicable for single-channel data. In addition ICA requires data storing and visual inspection on selecting the independent component, hence, a complete automatic artifact removal in real time is almost impractical. Recently, the hybrid approaches, which is the combination of several methods, have been proposed for automatic artifact removal using ICA (Junfeng Gao, Yang, Lin, & Wang, 2010; Hamaneh, Chitravas, Kaiboriboon, Lhatoor, & Loparo, 2014). The proposed methods successfully remove the artifact without human intervention but there is still a problem with the computational complexity.

One of the best options for automatic and real time processing is using adaptive filtering method. Adaptive noise cancellation (ANC) system in adaptive filtering (Haykin, 1996) provides a noise removal technique which can be operated in real-time and available for a single-channel EEG. However, adaptive filter requires a reference signal which correlated with the artifact source in the noisy signal. Recently, the approaches of using a camera or eye-tracker based references to eliminate the artifacts have been proposed (Kierkels, Riani, Bergmans, & Van Boxtel, 2007; Noureddin, Lawrence, Member, & Birch, 2012). Nevertheless, those additional sensors still bring both the hardware size and cost problem, which is unfavorable for portable applications. Other methods to tackle the problem of reference channel have also been proposed. Shahabi et. al developed a model of eye blink to be used as reference in Kalman filter (Shahabi, Moghimi, & Zamiri-Jafarian, 2012). The proposed method used Output-error (OE) model (Morbidi et al., 2007) to model the eye blink artifact. The reference EEG channel is only needed at the first time to build a basic eye blink artifact before the modeling process. However, an additional unnecessary complexity is one of the issue need to be considered when introducing a new model into the system. Meanwhile, a method of using discrete wavelet transform (DWT) to construct a reference signal of ocular artifacts (eye blinking and eye movement) has been proposed (Peng et al., 2013). The proposed method was able to suppress the artifact without implementing a reference sensor, but it is found to be difficult to select the best threshold value.
In this paper, we present a method to construct a reference signal in time domain, which is more straightforward and does not require any transformation to frequency domain signal. We demonstrate the use of Savitzky-Golay filter to estimate an eye blink reference signal for adaptive filter. The proposed method is advantageous not only of its availability for real time processing, but also does not require additional reference sensor for calibration and during recording, thus available for single channel application. The artifact removal is performed using ANFIS algorithm. The structure of this paper is composed as follows: In the next section, we provide an introduction on adaptive filtering and the layout of proposed model. The experiment setup and dataset preparation are briefly described in the same section, followed by the methods used for performance analysis. Next, the performance of the proposed method is evaluated and we compare the proposed method with ICA. Finally, the overall findings and the limitations on the current method are discussed.

**METHODOLOGY**

**Savitzky-Golay Filter**

The basic concept of smoothing filter is to let the low frequency components pass through while attenuate the high frequency components. Eye blinks in EEG signals are mainly concentrated on low frequency band. Here, the idea of using smoothing filter for blink extraction so that the target eye blink which is in low frequency band is filtered out while leaving the higher frequency of original brain signals. The Savitzky-Golay filter is first proposed by Savitzky and Golay for data smoothing based on local least squares polynomial approximation (Savitzky & Golay, 1964). This filter was originally developed for noise removal in analytical chemistry, but it has been adopted in various areas later on due to its ability to preserve the peak shape, unlike many other smoothing techniques (Joensson & Eklundh, 2004; Schafer, 2011). The output of SG filter is calculated as follows:

\[
Y_j = \sum_{n=-L}^{n} C_i y_{j+i}
\]  

where C represents the coefficients value from a set of weighting factors. The coefficients perform k degree polynomial which fit to window size, N. The window size \( N = nR + nL + 1 \), where \( nR \) and \( nL \) are the number of points used to the right and the number of points used to the left from the current data point, respectively. The recorded EEG signals contaminated with eye blinks artifacts will be smoothed using above equation, and the output (smoothed values) will be set as a reference in adaptive noise cancellation system.

**Artifact Removal using Adaptive filtering**

The use of conventional filter such as low-pass filter and high-pass filter to remove the artifact directly from the signal could possibly remove the relevant EEG signals as well, because there might be overlapping in frequency spectra between the artifacts and the EEG signals (Jianbo Gao, Sultan, Hu, & Tung, 2010). One of the advantages of ANC is it could automatically adjusts its parameters following the changing of statistical properties of the input in order to achieve optimal filtering. The filter suppresses the noise by subtracting the references from the measured signal based on linear regression process. The schematic diagram for ANC system mainly consists of two input which are, a primary input, \( d(n) \) and one reference input, \( x(n) \) (Widrow, Lehr, Beaufays, Wan, & Bilello, 1993). In ANC adaptive filter, interference source is used as reference input in estimating the clean signal while adjusting the filter coefficients to achieve the optimal filtering (Haykin, 1996). In this study, the primary input is the measured EEG signal while the reference input is the signals representing the blink components. The proposed method of combining the Savitzky-Golay filter and ANC is illustrated in Figure-3. The implementation of Savitzky-Golay filter is to extract blink component from the noisy EEG signal and input to ANC filter. The filter will estimate the clean EEG signal \( e(n) \) by predicting the artifact component, \( y(n) \) from the primary input \( d(n) \). This can be modeled as

\[
d(n) = s(n) + x_0(n)
\]  
\[
x(n) = q(n) + u(n)
\]  
\[
e(n) = d(n) - y(n)
\]

where \( s(n) \) is the signal of interest (clean EEG signal) and \( x_0(n) \) is the artifact source which contaminates the primary input sensor. \( q(n) \) is the actual source of the artifact which is correlated to \( x_0(n) \), while \( u(n) \) is the measurement noise. The output \( y(n) \) will be estimated as close as to \( x_0(n) \), by filtering the reference input \( x(n) \). The estimated \( y(n) \) is to be subtract from \( d(n) \) to produce an estimation of clean signal, \( e(n) \). In this study, we implemented ANFIS in the adaptive algorithm.

**Dataset Preparation**

The algorithm was tested using real EEG dataset which was obtained from Keio University (Kanoga & Mitsukura, 2015). The recording was performed using a
cap type multi channel EEG device, g.tec. The EEG signals were recorded from 14 points (Fp1, Fp2, F3, Fz, F4, T3, C3, C4, T4, P3, Pz, P4, O1, and O2) followed the international standard of 10-20 positioning system. A vertical EOG signal was measured at upper and lower right eye by using two disposable electrodes. The device measured 14 locations of EEG signals and 1 vertical EOG signal. The sampling frequency was set to 256 [Hz] while reference and ground electrode were placed at A1 and Fz, respectively. In this study, we used dataset from 15 participants. The participants were asked to sit still and relax during the experiment. They were instructed to blink according to metronome sound that will be played every 5 seconds. The blinks were measured simultaneously using EOG electrode. The total length of the experiment is 55 seconds including 5 seconds of margin.

We used the measured EEG signals at Fp1 channel for adaptive filtering dataset. The reason of using Fp1 channel from frontal location is because the position of this point is the nearest to the eyes and severely affected by eye blinks artifact. The Savitzky-Golay filter is applied only on the detected blinks locations. Figure-4 is the example of recorded EEG signal at Fp1 and simultaneously recorded EOG signal. We can see that EOG signal caused by eye blink return to baseline about 2.5 second after the first peak. For this, we apply Savitzky-Golay filter to the measured EEG signal from 30 points before the first peak to 470 point after the first peak. The total length of window is 500 samples. This blink onset can easily be determined by applying peak detection algorithm with a certain threshold value. The filter parameters were chosen based on trial and error. We found that using a combination of k=1 and N=47 produced the highest correlation to the measured EOG signal. Thus, these values were set in the parameters of the Savitzky-Golay filter. The example of extracted blink form Fp1 signal is shown in Figure-5. The output values from SG filter is then input to adaptive filter as reference. We then perform the filtering on the blink segments using ANFIS algorithm.

In this study, we compare the result from the proposed method with the result obtained using ICA. All the data from 14 EEG channels and 1 vertical EOG signals were used for eye blinking artifact rejection in ICA. The recorded signals were decomposed into intrinsic EEG signals and estimated eye blink artifact signals. The signals were extracted in regard to Fp1 point as target. This channel is chosen in order to obtain equivalent comparison with the proposed adaptive filter.

Performance Analysis

In order to provide a quantitative measure of the performance for the proposed method, first, we used signal to noise ratio (SNR) and mean squared error (MSE) by comparing the value before and after the filtering. SNR and MSE values are calculated as follows:

\[
\text{SNR}_{dB} = 10 \log \left( \frac{\sum x^2(n)}{\sum (x(n) - \tau(n))^2} \right)
\]

\[
\text{MSE} = \frac{1}{N} \sum (x(n) - \tau(n))^2
\]

where \( x(n) \) is the target clean EEG signal (corrected Fp1 signal by ICA) and \( \tau(n) \) is the corrected EEG signal by proposed method (for SNR before filtering, original measured Fp1 signal is used). N represents the length of the signal.

Next, to measure the degree of similarity between EOG signals to the corrected EEG signals, we calculate the cross-correlation between both signals using Pearson correlation coefficient \( \rho \). We compared the value of \( \rho \) between Fp1 signals before and after corrected. From the value of \( \rho \), we calculate the performance index for both methods. The performance index is defined as (Klados, Papadelis, Lithari, & Bamidis, 2009):

\[
P = \rho_{\text{EEG, EOG}} - \rho_{\text{EEG, EOG}}
\]

where \( \rho_{\text{EEG, EOG}} \) and \( \rho_{\text{EEG, EOG}} \) is the correlation coefficient between measured EEG and EOG signal. Meanwhile \( \rho_{\text{EEG, EOG}} \) is the correlation coefficient between corrected EEG signal and EOG signal.
RESULTS

The result of estimated clean signal after filtering is shown in Figure 6. Figure 6(a) represents the example of measured Fp1 signal with blink artifact, estimated Fp1 by ICA and estimated Fp1 by the proposed method. The figure clearly shown that blink segment has been successfully removed by both methods. Figure 6(b) is a close section on estimated Fp1. It is clearly from the figure that the estimated signal by the proposed method overlapped with the estimated signal by ICA in most of the time. Then, we compare the filter robustness in estimating eye blink component from the measured EEG signal. The figure for comparison of estimated blink with ICA method is shown in Figure-7. Based on visual inspection, there is no significant difference in terms of shape and magnitude, between estimated blink by adaptive filter and ICA. From here we could say that the performance of the proposed method is comparable with ICA method.

Next, for quantitative evaluation, we analyzed the effect of the proposed method on SNR and MSE value before and after artifact correction. SNR here is defined as the power ratio between the reference EEG signals i.e. constructed Fp1 signal by ICA method, to the power of tested signal. The change in SNR value before and after artifact correction by proposed method is illustrated in Figure-8. We observed that the SNR for the raw signal was below than -10[dB] for all of the subjects giving the average at -15.34 [dB]. On the other hand, the average of SNR for corrected Fp1 is 0.92 [dB], shows an approximately 105% of increase in power ratio. Meanwhile, the MSE values are plotted in Figure-9. The average MSE for raw and corrected Fp1 signal are 9.93 and 1.50 respectively. The average of 85% decreases in MSE value after corrected proved the reliability of the proposed method in removing the artifact component from measured signal.
Finally, we compute the performance index using cross correlation coefficient. The mean value for cross correlation between Fp1 signal and EOG signal for both ICA and the proposed method are listed in Table-1. The filter successfully reduced the coefficient value to almost zero indicates that the blink artifacts were almost completely been removed. The value of P is illustrated in Figure-10. The performance of adaptive filter exceeds the performance of ICA in all subjects except for subject 7 with a slight difference. Note the value of P for the proposed method reach near 1 show that there is almost no similarity between the corrected EEG signal and the EOG signal. The obtained result is consistent with the assumption that EOG signal and clean EEG are uncorrelated (Shahabi et al., 2012). Based on the performance index, our result outperforms the method proposed by (Szibbo, Luo, & Sullivan, 2012), which performed the noise removal by direct subtracting the values from SG filter without going through an adaptive filter.

Nevertheless, there are some limitations in the studies need to be considered. First limitation is on the dataset preparation. In this paper, we used the EEG dataset with voluntary eye blinks. The magnitudes of voluntary eye blinks are much larger compared to clean EEG signals. The duration is also longer than involuntary blinks. Thus, there are very easy to be detected from their magnitudes and shapes. However, for involuntary blinks, the signals are more random and no specific shape or duration. It is basically comes with sharper spike and shorter period and could be mixed up with eye movements artifacts. Hence, further study of the proposed method on involuntary eye blinks should be conducted to verify its reliability on real environment. Second limitation is all of the quantitative evaluation is done based on estimation. The real value for clean EEG is unknown thus the percentage of EEG signal retained in reconstructed signal is unknown. A systematic artificial model of contaminated EEG signal and target signal is needed to further demonstrate the availability of the filter with quantitative analysis. The third limitation is the possibility of filtering out relevant EEG data in lower frequency band. The Savitzky-Golay filter is a type of low-pass filter. A small fluctuation in extracted blink signal by the filter may result in subtracting the EEG signal in lower frequency band by the adaptive filter.

In this paper, we proposed an eye blinking artifact removal method for single channel EEG using Savitky-Golay filter referencing in adaptive filtering. We performed the proposed method on Fp1 channel of real EEG dataset. ANFIS algorithm is then adopted in ANC system to perform the blink removal by applying the SG filtered values as a reference input. The filter performance using the proposed referencing method is then compared with the ICA method. The result demonstrated a better performance of adaptive filter over conventional ICA method.

ACKNOWLEDGEMENTS

This research is supported by Universiti Teknologi Malaysia (UTM) and Malaysia-Japan International Institute of Technology (MIJIT) under grant vot. no. R.K430000.7743.4J011. Special thanks to Associate Professor Dr. Yasue Mitsukura from Graduate School of Science and Technology, Keio University and all her laboratory members for the provision of EEG data.

REFERENCES


