



**HAL**  
open science

## Unsupervised Eye Blink Artifact Identification in Electroencephalogram

Ashvaany Egambaram, Nasreen Badruddin, Vijanth S Asirvadam, Eric Fauvet, Christophe Stolz, Tahamina Begum

► **To cite this version:**

Ashvaany Egambaram, Nasreen Badruddin, Vijanth S Asirvadam, Eric Fauvet, Christophe Stolz, et al.. Unsupervised Eye Blink Artifact Identification in Electroencephalogram. TENCON 2018, Oct 2018, Jeju, South Korea. pp.2148-2152, 10.1109/TENCON.2018.8650467 . hal-02140696

**HAL Id: hal-02140696**

**<https://hal.science/hal-02140696v1>**

Submitted on 27 May 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Unsupervised Eye Blink Artifact Identification in Electroencephalogram

Ashvaany Egambaram, Nasreen Badruddin,  
Vijanth S Asirvadam

Centre for Intelligent Signal and Imaging Research  
Department of Electrical and Electronic Engineering  
Universiti Teknologi PETRONAS, Malaysia  
Email:ashvaany\_g03337@utp.edu.my

Eric Fauvet, Christophe Stolz  
Laboratoire (Le2i)  
ERL VIBOT CNRS 6000  
Universite de Bourgogne  
France

Tahamina Begum  
Department of Neuroscience  
Universiti Sains Malaysia  
Kelantan, Malaysia

**Abstract**—The most prominent type of artifact contaminating electroencephalogram (EEG) signals is the eye blink (EB) artifact. Hence, EB artifact detection is one of the most crucial pre-processing step in EEG signal processing before this artifact can be removed. In this work, an approach that identifies EB artifacts without human supervision and automated varying threshold setting is proposed and evaluated. The algorithm functions on the basis of correlation between two EEG electrodes, Fp1 and Fp2, followed by EB artifact threshold determination utilizing the amplitude displacement from the mean. The proposed approach is validated and evaluated in terms of accuracy and error rate in detecting events of EB artifacts in EEG signals. Analysis has revealed that the proposed approach achieved an average of 96.6% accuracy compared to a conventional method of identifying EB artifacts with a fixed constant threshold.

**Keywords**—Electroencephalogram, EB Artifacts, Automated Threshold.

## I. INTRODUCTION

The Electroencephalogram (EEG) signal has been in use to interpret cognitive processes and physiological activity of the brain for medical purposes such as diagnosing epilepsy, sleep disorders, coma, encephalopathies and brain death. The EEG signal is also being extensively used for various research purposes, for example in neuroscience, cognitive science, cognitive psychology, neurolinguistics and psychophysiological research [1]. In addition, EEG signal is being progressively investigated for real time applications such as continual epilepsy monitoring and brain computer interface (BCI). Hence, it is important to extract meaningful neurological information from EEG signal to facilitate its interpretation. However, EEG signals are often contaminated with undesired non-neurological information which causes deviations in the signal of interest. These undesired signals are termed artifacts. Superimposition of these artifacts with EEG signal could potentially mystify EEG's interpretation. This is particularly relevant in medical field where EEG signals are being used as a sole diagnostic source, thus failing to recognize artifacts may severely affect clinical decisions. Therefore, artifact identification in EEG signal processing is the first and most crucial step prior to artifact

removal. The most common types of artifacts contaminating EEG signals are the cardiac artifact, the muscle artifact and the eye blink artifact. The muscle artifact is induced by muscle movement and contraction during EEG recording which takes place when the subject talks and swallows. The pattern of the artifact purely relies on the degree of muscle movement and contraction. Cardiac artifact on the other hand is an effect of electrical activity of the heart. The influence of cardiac artifact on the scalp is typically low, with low amplitude, repetitive and are of regular pattern. The eye blink (EB) artifact is the strongest and most prominent type of artifact recorded along with EEG, hence this research will focus on identifying it. EB artifacts appear as spikes with amplitudes of around 10 times greater than the actual brain signals, noticeable in the delta wave range and can last up to 200ms to 400ms [2], [3].

A lot of methods have been developed in the past for automatic identification or detection of EB artifacts in EEG signal. One of the easiest and preferred ways for EB artifact recognition is by simply using an amplitude threshold [4]. This method determines if an EB artifact present in an EEG segment if the amplitude in the segment exceeds a predetermined amplitude threshold. However, the amplitude of eye blinks may vary depending on the blinking strength of an individual, hence this method may not identify EB artifacts from an individual who exhibits gentle blinks which are lower than the threshold value. The other common way of detecting EB artifacts is through feature based identification which determines the presence of EB artifact in an EEG segment after extracting certain features. Some of the common features used are the kurtosis, maximum absolute value, entropy-based features and second-order difference in [4], [5], [6], [7]. As a general rule, these artifact detection features require a certain threshold value to classify or make a binary decision whether or not an EEG segment is contaminated by EB artifact. As elaborated earlier for the amplitude threshold, applying a fix threshold value for the features discussed may lead to detection errors due to the individual variance in blinking pattern and blinking strength. As a result, the threshold values may need

to be tailored for every individual which will be impractical in real time applications. This study proposes a new approach to automatically identify EB artifacts with varying threshold without any supervision on the EEG signal. The performance of the proposed approach is measured by validating if the proposed approach is accurate in identifying EB artifacts in comparison with the use of a fix threshold.

The paper is organized as follows: Section II illustrates the methodology, Section III elaborates the proposed algorithm, results and discussions are included in Section IV and V. Finally, Section VI provides the conclusion.

## II. METHODOLOGY

### A. Evaluation on Real EEG Signal

10 sets of EEG signals collected at Universiti Teknologi PETRONAS, (UTP) are used to evaluate the proposed algorithm. EB artifacts can be clearly captured in the frontal channels, Fp1-Fp2 electrodes of the EEG recordings. Hence, the proposed algorithm is evaluated on the frontal channel, Fp1, of these EEG signals. All recorded signals are of different durations, with a sampling frequency of 256 Hz. The evaluation was performed using MATLAB R2016b in Windows 7 Professional (64 bit OS) with a 4GB RAM.

## III. PROPOSED APPROACH

### A. Unsupervised EB Artifact Region Detection

In an EEG signal recording, the EB artifacts are primarily captured in the frontal electrodes, Fp1 and Fp2. This is because the frontal electrodes are positioned close to the eyes. Another logical point to note here is, both eyes of any individual blinks simultaneously, hence the Fp1 and Fp2 electrodes should theoretically unveil high correlation whenever eyes blink. To validate this theory, the correlation between Fp1 and Fp2 channels of an EEG signal is computed in segments of 500 samples (1.95 seconds) per segment. This window size is chosen so that at least one EB artifact can be captured in this window. Figure 1 shows an example of Fp1's and Fp2's recordings plotted out and their corresponding correlation coefficient values in each segment.

The test has revealed that segments of Fp1 and Fp2 without EB artifact produces correlation below than 0.7, whereas segments containing EB artifact results in higher correlation, usually more than 0.9. Thus, existence of EB artifact in a particular segment or window is supported by the high correlation coefficient value, which proves the theory. Following this, segments of the EEG signal contaminated with EB artifacts are identified. However, EB artifact components should be identified for subsequent analysis or artifact removal, which requires a threshold. Next section will discuss on the automated threshold level determination on every window that was identified containing an EB artifact.

### B. Automatic Thresholding to Identify EB Artifacts

Displacement or deviation of amplitude from the mean is chosen as a threshold criteria to classify the EB artifact's onset point. The displacement of amplitude is chosen as EB

artifacts are higher in amplitude in nature relative to that of the EEG or brain signal, hence the EB artifacts are expected to yield higher displacement from mean amplitude in comparison with uncontaminated EEG potentials. A threshold value is required in the case of recognizing the starting event of an EB artifact. First, the windows of Fp1 channel that were confirmed containing EB artifacts identified through steps illustrated in section III-A are subjected to attain displacement distribution from the mean amplitude using:

$$\text{Displacement}[t] = |[X[t] - \mu]| \quad (1)$$

where,  $X[t]$  is the EEG signal's amplitude at sample  $t = n$  until  $t = n+500$ ,  $n$  is any sample point of the signal where the window starts, and  $\mu$  is the mean of that window or segment.

An EB artifact event detection algorithm within a window or segment confirmed containing an EB artifact is then designed. The onset of EB artifact is assumed 100 samples (0.39 seconds) before the first sample of amplitude displacement more than the threshold, within the window. The reason setting the onset point in advanced of 100 samples before the threshold is to provide a buffer for any subsequent analysis. The end point of the EB artifact is found by aligning a 1 second (256 samples) window apart from the first sample of amplitude displacement more than the threshold. A 1 second window is chosen, because an eye blink can last up to 0.8 seconds (205 samples) in duration which can fit well into this frame (100+256=356 samples).

An experiment was conducted to define classification criteria to classify EEG potentials and EB artifacts. The threshold value for EB artifact for windows containing EB artifacts are checked and determined by this experiment. From the displacement distribution, a mean ( $\mu$ ) and a standard deviation ( $\sigma$ ) are acquired. Figure 2 shows an example of displacement distribution by setting the threshold value to be any displacement values beyond  $1\sigma$  from the mean, while figure 3 shows EB artifacts components plotted in red when threshold is set beyond  $1\sigma$ .

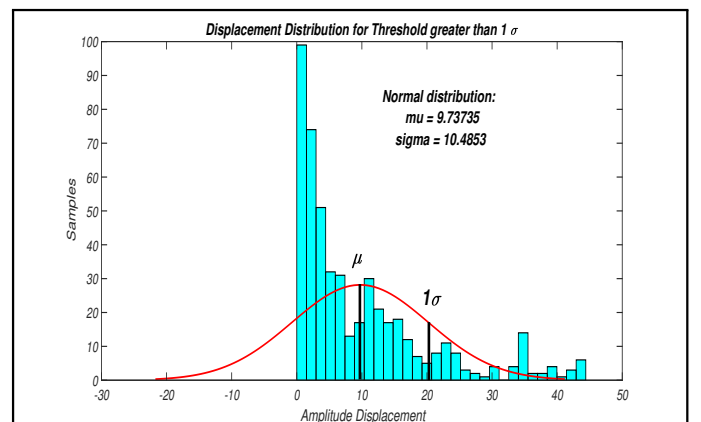


Fig. 2. Displacement Distribution for Threshold greater than  $1\sigma$

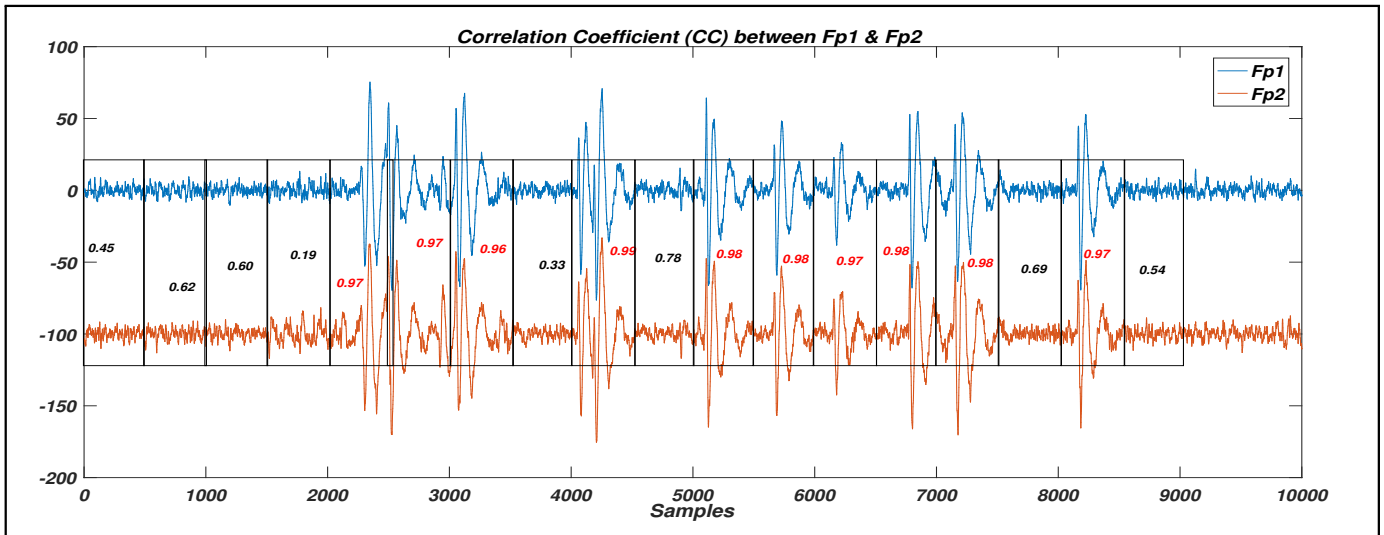


Fig. 1. Correlation Coefficient, CC, between Fp1 and Fp2 Electrodes

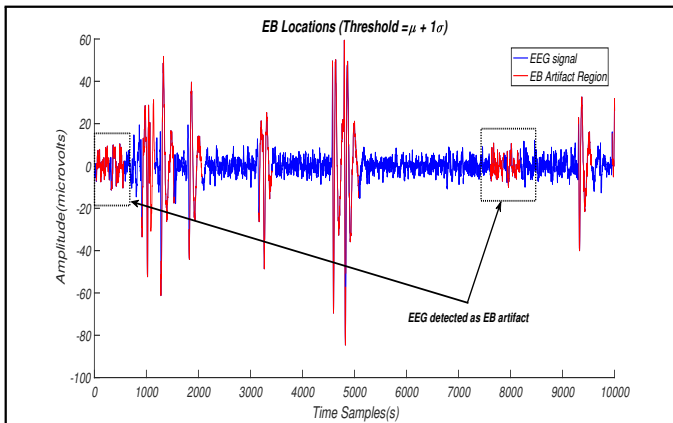


Fig. 3. EB Artifacts for Threshold greater than  $1\sigma$

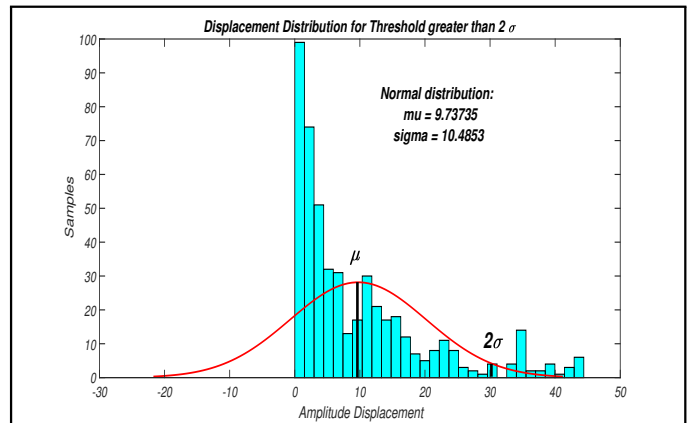


Fig. 4. Displacement Distribution for Threshold greater than  $2\sigma$

From Figure 3, it can be seen that some of the EEG potentials are identified as EB artifacts when the threshold is set beyond  $1\sigma$  from the mean.

Figure 4 shows the displacement distribution of the same segment by setting the threshold value to be any displacement values beyond  $2\sigma$  from the mean, while figure 5 shows EB artifacts components plotted in red when threshold is set beyond  $2\sigma$ .

From Figure 5, it can be seen that all EB artifacts are identified as EB artifacts when the threshold is set beyond  $2\sigma$  from the mean.

Figure 6 shows the same displacement distribution by setting the threshold value to be any displacement values beyond  $3\sigma$  from the mean, while figure 7 shows EB artifacts components plotted when threshold is set beyond  $3\sigma$ .

From Figure 7, it can be seen that the EB artifacts are not identified as EB artifacts when the threshold is set beyond  $3\sigma$  from the mean.

Therefore the threshold for EB artifact dominating the

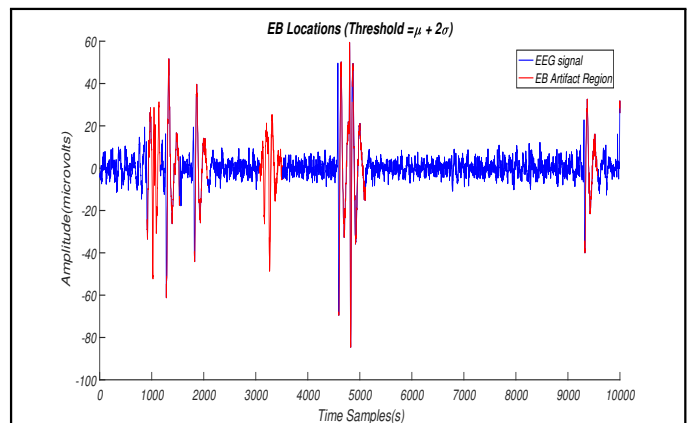


Fig. 5. EB Artifacts for Threshold greater than  $2\sigma$

EEG window in question can be correctly determined by taking two standard deviation,  $2\sigma$  width from mean of the

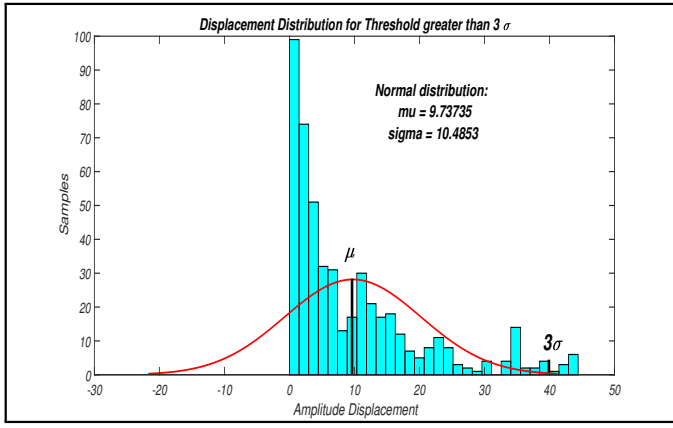


Fig. 6. Displacement Distribution for Threshold greater than  $3\sigma$

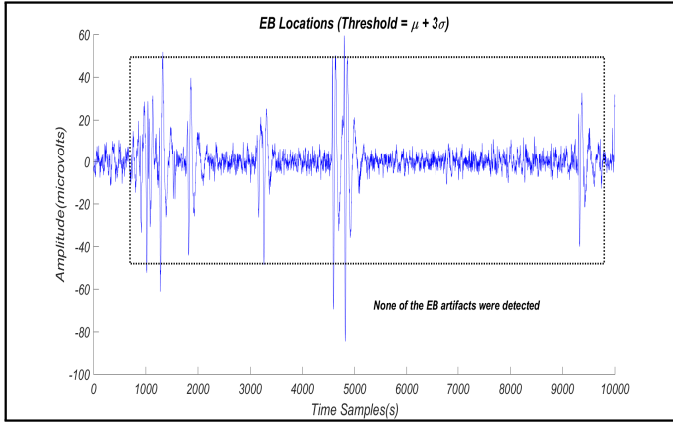


Fig. 7. EB Artifacts for Threshold greater than  $3\sigma$

displacement distribution acquired.

$$\text{threshold} = \text{mean} + 2\sigma \quad (2)$$

Figure 8 summarizes the EB artifact detection algorithm in a flowchart.

#### IV. RESULTS

This section aims to measure the accuracy of the proposed approach in detecting EB artifacts compared to the use of a constant or common threshold. The constant threshold is evaluated by fixing the threshold values of more than 10uV, 20uV, 30uV, 40uV and 50uV, whichever amplitude displacement that exceeds these thresholds are considered to be EB artifacts.

The binary prediction is used to test the detection of EB artifacts, which produces below outcomes and the confusion matrix as in Table I:

- True positive (TP): correct EB artifact detection
- False positive (FP): clean EEG identified as EB artifact
- True negative (TN): correct clean EEG identification
- False negative (FN): EB artifact identified as clean EEG

The efficiency of the proposed algorithm compared to fix a constant threshold, is validated by manually inspecting

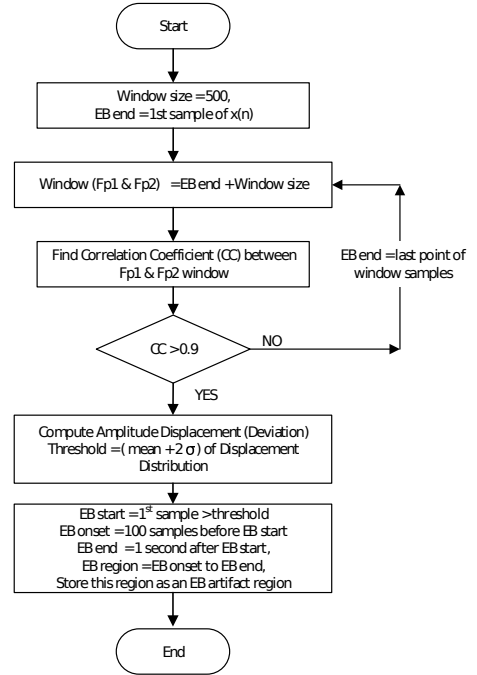


Fig. 8. Flowchart of the Automated EB Artifact Identification Algorithm

TABLE I  
CONFUSION MATRIX - EB ARTIFACT DETECTION

		Detected	
		EB Artifact	Clean EEG
Observed	EB Artifact	TP	FN
	Clean EEG	FP	TN

the EEG signals with measures derived from the confusion matrix, such as the accuracy and error rate.

*Error:* The error rate, ERR is computed with mis-detection of EB artifact and EEG divided by total number of detections. Values approaching zero denotes better error rate.

$$\text{ERR} = \frac{FP + FN}{TP + TN + FN + FP} \quad (3)$$

*Accuracy:* ACC is the opposite of error rate, where it is the ratio of correct EB artifact and EEG detections by total number of detections. The best accuracy is 1.

$$\text{ACC} = \frac{TP + TN}{TP + TN + FN + FP} \quad (4)$$

The results obtained from the proposed method in comparison with constant thresholds in detecting EB artifacts correctly, applied on the Fp1 channel of 10 EEG signals are tabulated in Table II.

#### V. DISCUSSIONS

As elaborated in Section III-A, an automated EB artifact detection algorithm is designed. The proposed approach detects EB artifacts without any human supervision. First EEG segments containing EB artifacts are recognized with the concept of correlation between EEG electrodes, Fp1 and Fp2. Secondly, to address the issue of fixing a constant threshold,

TABLE II  
COMPARISON OF ACCURACY

Signal	Accuracy					
	Proposed (Automated Varying Threshold)	Fix Thresholds				
		10uV	20uV	30uV	40uV	50uV
EEG 1	94.4%	<b>97.2%</b>	91.7%	77.8%	50.0%	50.0%
EEG 2	<b>98.9%</b>	<b>98.9%</b>	97.8%	89.9%	57.3%	42.7%
EEG 3	<b>97.9%</b>	92.5%	98.9%	96.3%	49.7%	21.4%
EEG 4	<b>97.5%</b>	94.9%	94.9%	75.9%	27.8%	26.6%
EEG 5	96.9%	96.1%	97.7%	<b>98.4%</b>	97.7%	70.3%
EEG 6	94.3%	94.3%	94.3%	<b>100.0%</b>	98.7%	98.7%
EEG 7	<b>97.9%</b>	94.7%	<b>97.9%</b>	96.8%	<b>97.9%</b>	<b>97.9%</b>
EEG 8	<b>97.0%</b>	96.3%	94.0%	94.8%	92.5%	65.7%
EEG 9	94.8%	89.6%	95.7%	<b>96.5%</b>	87.0%	73.0%
EEG 10	<b>96.2%</b>	93.7%	94.9%	94.9%	94.3%	81.0%
AVERAGE	<b>96.6%</b>	94.8%	95.8%	92.1%	75.3%	62.7%

an automated and varying threshold is determined for every EEG segment containing EB artifact using the amplitude displacement. From Table II, the accuracy achieved in detecting EB artifacts by the proposed technique is higher in average, 96.6% compared to 94.8%, 95.8%, 92.1%, 75.3% and 62.7% achieved through fixing thresholds between 10uV to 50uV of amplitude displacement. The individual accuracy value for each EEG signal of the proposed automated threshold reveals the accuracy has not fluctuated much, between 94% to 99% compared to the constant thresholds, between 89% to 99% for 10uV, between 91% to 99% for 20uV, between 75% to 100% for 30uV, between 27% to 99% for 40uV, and between 21% to 99% for 50uV. The constant threshold has achieved 100% accuracy for EEG set 6, which means the constant threshold of more than 30uV for EB artifact detection suits well for this data set. On the other hand, the 30uV threshold has only achieved 75.9% of accuracy for EEG set 4. This indicates the performance of fixing a constant threshold is purely dependent on the nature of the EEG signal and is not consistent across all EEG signals. While the automated threshold changes for every window, whereby the threshold will correspond to the varying nature of the EB artifacts. Hence, threshold is automatically determined for every window without setting any specific value.

## VI. CONCLUSION

Conventional EB artifact detection algorithms depends on constant thresholds or constant features to make a binary decision to recognize if an EEG segment contains EB artifact or not. In this paper, an unsupervised EB artifact detection algorithm in EEG signal is proposed and evaluated. The algorithm relies on the concept of correlation between two EEG electrodes and the amplitude displacement range when there is an event of EB activity. Based on the results and discussions, it is apparent that the automated and unsupervised

EB artifact detection algorithm proposed in this paper is accurate in identifying EB artifact events in an EEG signal. On the other hand, the algorithm is also consistent in detecting EB artifacts across different EEG signals compared to a conventional algorithm which is fixed with a constant threshold to detect EB artifacts. Thus, it can be concluded that the proposed algorithm is a reliable solution in detecting EB artifacts across all types of EEG signals which may have individual variance due to blinking pattern and strength.

## ACKNOWLEDGMENT

The authors would like to thank the Ministry of Education, Malaysia for supporting this research through the Fundamental Research Grant Scheme, FRGS (FRGS/2/2014/TK03/UTP/02/1) and the Higher Institution Centre of Excellence (HiCoE) Scheme.

## REFERENCES

- [1] J. A. Urigüen and B. Garcia-Zapirain, "Eeg artifact removal: State-of-the-art and guidelines," *Journal of neural engineering*, vol. 12, no. 3, p. 031001, 2015.
- [2] G. H. Klem, H. O. Lüders, H. Jasper, C. Elger *et al.*, "The ten-twenty electrode system of the international federation," *Electroencephalogr Clin Neurophysiol*, vol. 52, no. 3, pp. 3–6, 1999.
- [3] S. Sanei and J. A. Chambers, *EEG signal processing*. John Wiley & Sons, 2013.
- [4] H. Nolan, R. Whelan, and R. Reilly, "Faster: fully automated statistical thresholding for eeg artifact rejection," *Journal of neuroscience methods*, vol. 192, no. 1, pp. 152–162, 2010.
- [5] G. Barbati, C. Porcaro, F. Zappasodi, P. M. Rossini, and F. Tecchio, "Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals," *Clinical Neurophysiology*, vol. 115, no. 5, pp. 1220–1232, 2004.
- [6] A. Mognon, J. Jovicich, L. Bruzzone, and M. Buiatti, "Adjust: An automatic eeg artifact detector based on the joint use of spatial and temporal features," *Psychophysiology*, vol. 48, no. 2, pp. 229–240, 2011.
- [7] A. Klein and W. Skrandies, "A reliable statistical method to detect eyeblink-artefacts from electroencephalogram data only," *Brain topography*, vol. 26, no. 4, pp. 558–568, 2013.