Intelligent Automated EEG Artifacts Handling Using Wavelet Transform, Independent Component Analysis and Hierarchal Clustering

Shaibal Barua(✉), Shahina Begum, and Mobyen Uddin Ahmed

School of Innovation, Design and Engineering, Mälardalen University, 72123 Västerås, Sweden
{shaibal.barua,shahina.begum,mobyen.ahmed}@mdh.se

Abstract. Billions of interconnected neurons are the building block of the human brain. For each brain activity these neurons produce electrical signals or brain waves that can be obtained by the Electroencephalogram (EEG) recording. Due to the characteristics of EEG signals, recorded signals often contaminate with undesired physiological signals other than the cerebral signal that is referred to as the EEG artifacts such as the ocular or the muscle artifacts. Therefore, identification and handling of artifacts in the EEG signals in a proper way is becoming an important research area. This paper presents an automated EEG artifacts handling approach, combining Wavelet transform, Independent Component Analysis (ICA), and Hierarchical clustering. The effectiveness of the proposed approach has been examined and observed on real EEG recording. According to the result, the proposed approach identified artifacts in the EEG signals effectively and after handling artifacts EEG signals showed acceptable considering visual inspection.

Keywords: Electroencephalogram (EEG) · Ocular artifacts · Muscle artifacts · Hierarchical clustering

1 Introduction

The Electroencephalogram (EEG) signal represents the electrical activity of the brain that is recorded along the scalp. EEG signal analysis has become an important research area in the clinical research and Brain Computer Interfaces (BCI) applications. It has also become useful physiological measurement for sleep study, epilepsy, and cognitive science research. However, the problem with the EEG signal is that, it is non-stationary and non-linear, and contaminates with other biological signals e.g., Electrooculogram (EOG), Electromyography (EMG), Electrocardiogram (ECG) [1, 2]. EEG artifacts are referred to any undesired signals or potential differences due to extra-cerebral source that interfere with the recorded signal [3, 4]. Artifacts make EEG signals uninterruptable and in the EEG signal analysis these can lead to a serious misinterpretation. Moreover, in the EEG signal processing such as power spectral analysis or topographic displays, artifacts can cause false conclusion unless artifactual data are handled or excluded from the EEG data.

P. Perego et al. (Eds.): MobiHealth 2016, LNICST 192, pp. 144–148, 2017.
DOI: 10.1007/978-3-319-58877-3_19

The sources of the EEG artifacts are the ocular and the muscle activities; and both kind of artifacts overlap with the neural brain activity, which increase the difficulty to correctly interpret the EEG signals. The hypothesis on the artifacts is that, they are independent from the brain activity, either collected from normal or pathologic subjects [5]. Over the years, several methods have been proposed for artifacts removal from EEG signal. Daly and his colleagues [6] have developed the FORCe: Fully Online and automated artifact Removal tool for BCI application that combines wavelet transform, ICA and thresholding. Thresholding was applied on the independent components (ICs) to identify artifactual components, and resulting signals were obtained by removing the ICs that are identified as artifacts. LAMIC [7] is a clustering algorithm that has been developed to remove artifacts automatically from EEG signal. Blind source separation algorithm called the Temporal Decorrelation Source Separation (TDSEP) has been applied to decompose EEG signals. Later, clustering has been done on the components, based on the similarity of their lagged Auto-Mutual Information (AMI). ICA and clustering algorithms have been used to classify artifacts in EEG signals in [8, 7].

This paper proposed an approach, combining wavelet transform, Independent component analysis (ICA), and Hierarchical clustering methods to handle artifacts in the EEG signals. Here, recorded EEG signals are decomposed to approximation and details coefficients using wavelet transforms. Then ICA has been applied on the set of approximation coefficients to separate coefficients into ICs. Thereafter, features are extracted from each IC and all the ICs are grouped through Hierarchical clustering to identify clusters containing artifactual ICs, which are used to remove artifacts.

2 Materials and Methods

EEG signals are recorded at 2048 Hz sample rate following the international 10–20-electrode placement system with 19 channels EEG settings. A controlled data collection scenario was used during the data collection and participants were asked to perform different ocular and muscle movement activities to generate artifacts in the recorded EEG data.

A step-by-step process of the proposed system for artifacts handling is presented in Fig. 1. In the pre-processing, 50 Hz channel noise is removed from the recorded EEG signals using notch filter and signals are divided into 1 s segments. After pre-processing, recorded EEG signals are decomposed using wavelet transform that provides a set of approximation and details coefficients. Then approximations coefficients are further decomposed applying ICA to obtain ICs [6, 9].

In order to identify artifactual components i.e., the ocular and the muscle artifacts, several features are extracted from each IC (Hurst exponent, Kutosis, 1/frequency distribution, gamma power spectral density, spectral ratio i.e. the power spectral ratio of two frequency ranges (ratio of 30–60 Hz and 4–30 Hz), energy ratio i.e., $E(55\text{ Hz} \le f \le 100\text{ Hz})/E(f < 20\text{ Hz})$, spectral edge frequency (SEF)) [10].

To identify artifactual components the Hierarchical Clustering algorithm is applied. The Hierarchical algorithm clusters data over a variety of scales by creating a hierarchical structure (tree) or ‘dendrogram’. The tree is not a single set of clusters, but rather a

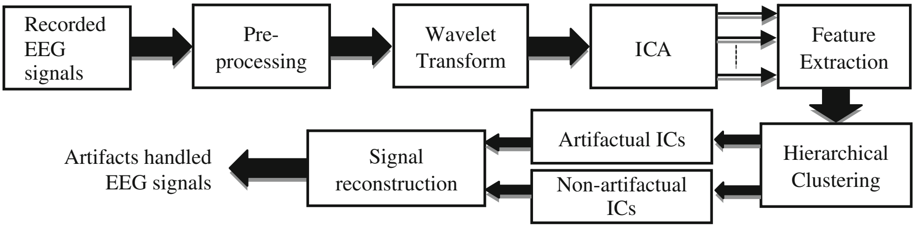

Fig. 1. Steps of the proposed approach in order to handle artifacts in EEG signals

multilevel hierarchy, where clusters at one level are joined as clusters at the next level [11]. In Hierarchical clustering, the distance between pairs of objects is calculated using Euclidean distance as a 'correlation' parameter of the MATLAB function 'pdist'. The linkage function applies 'complete' (i.e. Furthest distance) as a parameter, which determines the objects in the data set that should be grouped into clusters. Finally, cluster function with 'cutoff' value 1.1 is applied to group the data into random number of clusters. In addition, to label each cluster a 1-D feature vector is estimated that consists of 8 features (topography histogram, spectrum fit 1 and 2, frontal and peripheral location scores, average auto-correlation, Lorentz threshold, and Symlets wavelet (sym3) repeatability). Then a distance measure is applied to identify clusters containing artifactual components based Chauvenet's Criterion threshold value of 0.75 [12].

3 Evaluation and Results

EEG recording with artifacts during saccades and body movement activity are shown in Figs. 2(a) and 3(a) respectively. The resulting signals after artifacts handling are depicted in Figs. 2(b) and 3(b). By comparing two signals (the recorded EEG signals and the signals after handling artifacts), it is visible from the visual inspection that the artifacts are handled in the recorded signals.

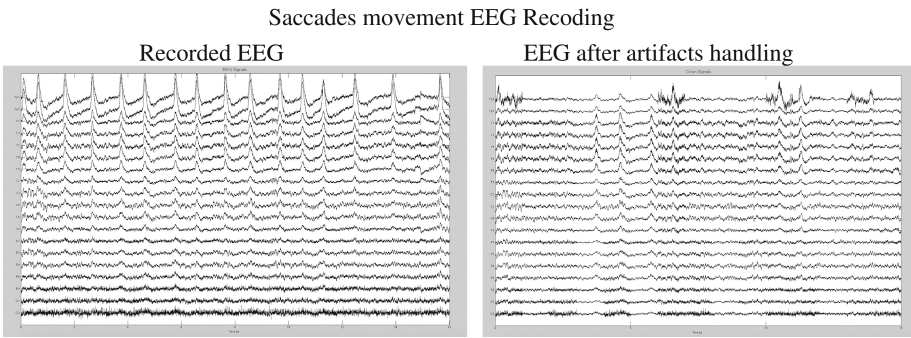

Fig. 2. (a) EEG signal with artifacts (b) EEG signal after artifacts handling

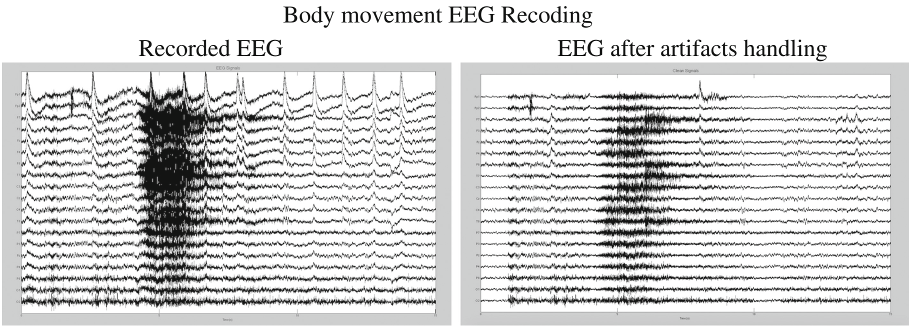

Fig. 3. (a) EEG signal with artifacts (b) EEG signal after artifacts handling

However, the ratio between the recorded signal and the artifacts handled signals is also important issue to observe. Thus, an experiment has performed to observe the difference considering Signal Quality Index (SQI) in different artifacts activities.

Here, SQI in terms of standard deviation of the amplitude values (StdAmp), maximum amplitude values (MaxAmp) and Kurtosis are calculated for saccades, body movement. The calculated values for EEG with and without artifacts are presented in Table 1.

Table 1. Signal quality index of EEG signals with and without artifacts

Activity	Parameter	With artifacts	Without artifacts	Difference in percentage
Saccades	StdAmp	21.03	11.75	44.09%
	MaxAmp	88.9	59.79	32.74%
	Kurtosis	4.3	5.03	15.42%
Body movement	StdAmp	23.54	10.28	56.33%
	MaxAmp	156.82	84.04	46.41%
	Kurtosis	7.4	8.0	08.01%

4 Conclusion

EEG signal measures the electrical activity of the brain and is a valuable physiological measurement in clinical application and research. In the EEG signal processing, it is important to handle artifacts i.e., the interference of physiological signals other than brain activity. This paper proposed an automatic approach for handling ocular and muscle artifacts in the EEG signals. Here, the proposed approach combines wavelet transform, ICA, and Hierarchical clustering. In addition, a distance measure has been applied using Chauvenet's Criterion threshold value of 0.75 to label the artifactual clusters. From the results it is noticeable that for the ocular artifacts handling it can improve signal quality index in the range of 15% to 44% and for the body movement signal quality index the improvement is 8% to 56%. In addition, the proposed approach does not require any reference signal to identify artifactual components from the independent components of ICA.

Acknowledgement. This research work is supported by the Vehicle Driving Monitoring (VDM) project funded by Swedish Governmental Agency for Innovation Systems (VINNOVA) and partially by the ESS-H profile and SafeDriver project funded by the Knowledge Foundation of Sweden.

References

1. Klonowski, W.: Everything you wanted to ask about EEG but were afraid to get the right answer. Nonlinear Biomed. Phys. **3**, 1–5 (2009)
2. Migotina, D., Calapez, A., Rosa, A.: Automatic artifacts detection and classification in sleep EEG signals using descriptive statistics and histogram analysis: comparison of two detectors. In: Spring Congress on Engineering and Technology (S-CET), pp. 1–6 (2012)
3. Chadwick, N.A., McMeekin, D.A., Tan, T.: Classifying eye and head movement artifacts in EEG signals. In: 2011 Proceedings of the 5th IEEE International Conference on Digital Ecosystems and Technologies Conference (DEST), pp. 285–291 (2011)
4. Klass, D.W.: The continuing challenge of artifacts in the EEG. Am. J. EEG Technol. **35**, 239–269 (1995)
5. Romo-Vazquez, R., Ranta, R., Louis-Dorr, V., Maquin, D.: EEG ocular artefacts and noise removal. In: 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2007, pp. 5445–5448 (2007)
6. Daly, I., Scherer, R., Billinger, M., Muller-Putz, G.: FORCe: fully online and automated artifact removal for brain-computer interfacing. IEEE Trans. Neural Syst. Rehab. Eng. **23**(5), 725–736 (2014)
7. Nicolaou, N., Nasuto, S.J.: Automatic artefact removal from event-related potentials via clustering. J VLSI Sign Process. Syst. Sign **48**, 173–183 (2007)
8. Barua, S., Begum, S., Ahmed, M.U., Funk, P.: Classification of ocular artifacts in EEG signals using hierarchical clustering and case-based reasoning. In: Workshop on Synergies Between CBR and Data Mining at 22nd International Conference on Case-Based Reasoning, Cork, Ireland (2014)
9. Ghandeharion, H., Erfanian, A.: A fully automatic ocular artifact suppression from EEG data using higher order statistics: improved performance by wavelet analysis. Med. Eng. Phys. **32**, 720–729 (2010)
10. Barua, S., Begum, S., Ahmed, M.U.: Clustering based approach for automated EEG artifacts handling. In: 13th Scandinavian Conference on Artificial Intelligence (SCAI 15), Halmstad, Sweden (2015)
11. Chen, G., Jaradat, S.A., Banerjee, N., Tanaka, T.S., Ko, M.S.H., Zhang, M.Q.: Evaluation and comparison of clustering algorithms in anglyzing ES cell gene expression data. Statistica Sinica **12**, 241–262 (2002)
12. Hames, E.C.: EEG Artifact Removal and Detection via Clustering. Electrical and Computer Engineering, Ph.D. Texas Tech University (2014)