Kmeans-ICA based automatic method for ocular artifacts removal in a motor imagery classification

Elie BOU ASSI, Sandy RIHANA, IEEE-EMBS, and Mohamad SAWAN, Fellow, IEEE

Abstract— **Electroencephalogram (EEG) recordings aroused as inputs of a motor imagery based BCI system. Eye blinks contaminate the spectral frequency of the EEG signals. Independent Component Analysis (ICA) has been already proved for removing these artifacts whose frequency band overlap with the EEG of interest. However, already ICA developed methods, use a reference lead such as the ElectroOculoGram (EOG) to identify the ocular artifact components. In this study, artifactual components were identified using an adaptive thresholding by means of Kmeans clustering. The denoised EEG signals have been fed into a feature extraction algorithm extracting the band power, the coherence and the phase locking value and inserted into a linear discriminant analysis classifier for a motor imagery classification.**

I. INTRODUCTION

Electroencephalogram (EEG) is a measurement of electrical currents that flow during synaptic excitations of the dendrites of many pyramidal neurons in the cerebral cortex [1]. Therefore, EEG is used to record the electrical activity of the brain via electrodes placed non-invasively on the scalp. This electrical signal is always contaminated by internal and external noises and should be processed for further analysis and feature extraction. Ocular artifacts, especially eye blinks, are one of the main interferences in EEG and consist of low frequency high amplitude signals and show propagation over the anterior scalp regions [2].While some artifacts can be minimized by training the subject to relax and to avoid facial expressions, it is almost impossible to ask a subject to control involuntary eye blinking. Furthermore, ocular artifacts overlap with the EEG frequency of interest and cannot be filtered by known traditional methods. Two main approaches are used in order to deal with this type of artifacts; the first consists of a manual rejection of the epochs on the EEG channels that represent an ocular artifact which is not accurate [3]. The second consists of conventional ocular artifacts denoising methods such as linear filtering, regression methods, Principal Component Analysis (PCA) and Independent Component Analysis (ICA) [4]. ICA has shown to be very effective in eliminating the activity of a wide variety of artifactual sources from EEG recordings with results better than those obtained with other methods [2]. However, when implementing an ICA-based method, the artifactual components should be identified. While some researchers added the acquisition of an EOG additional channel and studied the relations between the EOG and the independent components (ICs) [5], others identified ocular artifacts ICs manually by interpreting scalp topographies and frequency distribution characteristics in terms of visual inspection [6].This study implements an automatic ocular artifacts rejection method based on ICA without the use of an additional channel such as an EOG channel. Therefore, the ocular artifact components will be classified according to different features based on the proper characteristics of these ICs. The artifactual ICs are identified using an adaptive thresholding by means of Kmeans clustering. After EOG correction, the EEG signals are used as inputs for a right hand motor imagery based classification. Combining three features and using an LDA classifier, the recognition rate of the BCI hugely increased after ocular artifacts removal.

Reseach supported by Higher Center Of Research-USEK

E. BouAssi is with the Biomedical Engineering Department, Holy Spirit University, Jounieh, Lebanon; email: elie.g.bouassi@net.usek.edu.lb

S.Rihana is with the Biomedical Engineering Department, Holy Spirit University,USEK,Jounieh,Lebanon(email:sandyrihana@usek.edu.lb)

M.Sawan is with the Polystim Neurotechnologies Lab, dept of Elec. Eng., Ecole Polytechnique Montreal, Canada(email: mohamad.sawan@polymtl.ca)

II. MATERIALS AND METHODS

A. Signal acquisition

The EEG signals were downloaded from “PhysioNet.org” which is an international database that includes a large amount of datasets regarding different physiological states. “EEG Motor Movement/Imagery” signals were adopted; they were recorded using the BCI2000 acquisition system. Each subject performed 14 different trials while 64 EEG channels were recorded and sampled at 160Hz [7].

B. Temporal and spatial filtering

Since this broad frequency band contains all frequency components necessary for a motor imagery classification mainly the mu rhythm (8-12 Hz), all EEG channels were band pass filtered between 1 and 30 Hz using a finite impulse response filter (FIR). Furthermore, a spatial filtering was performed since it increases the signal to noise ratio what will increase the classification accuracy. This is done by subtracting the common activity of the brain to a certain position of interest as depicted in equation 1.

$$\widehat{\mathrm{V}}_{\mathrm{i}} = \mathrm{V}_{\mathrm{i}} - \frac{1}{N}\sum_{\mathrm{i}\in\omega_i}\mathrm{V}_{\mathrm{j}} \quad (1)$$

Vi and $\widehat{\mathrm{V}}_{\mathrm{i}}$ are the considered electrode potentials respectively before and after filtering, ω_i is the set of neighboring electrodes, and N is the total number of neighboring electrodes.

C. Independent Component Analysis

Independent Component Analysis separates a multi-channels signal into additive activations in a way that these components are statistically independent. All ICA algorithms start by whitening the data to remove any existing correlation. Since the EEG signals consist of a linear mixture of real brain activity with different noise components, when we whiten these mixtures, the variance on both axes is equal and the correlation of the projection of the data on all axes is null. After the whitening phase, the algorithm consists of rotating the resulting axis of the matrix in a way that itminimizes the Gaussianity of the projection on all axes. The full transformation from the original space is known as the weight matrix [8].

$$S= W*X \quad (2)$$

where the matrix X is the data in the original space, W is the weight matrix and S represents the sources activity. As seen in equation 2, the objective of the ICA algorithm is to find the weight matrix W that will decompose the EEG signals into ICs assuming temporal and spatial independency. In equation 2, X contains the EEG signals recorded all over the scalp. In this study, an Extended-Infomax ICA was adopted due to its ability to separate sub and sup Gaussian signals simultaneously [8]. The ICs representing ocular artifacts should be omitted during the signal reconstruction. The implemented method automates the process of ocular artifacts identification.

D. Automatic artifact identification

Three features were used for the automation of the ocular artifacts ICs identification:

Average correlation: this first feature adopts the average of correlation between each IC and the two prefrontal electrodes (FP1 and FP2) as depicted in equation 3. The cross correlation,R_{xy} ,is widely used to estimate the degree of similarity of two time series [1]. R_{xy} is the cross correlation of tow time series x (n) and y (n) as cited in equation 3.

$$R_{xy} = \sum_{n=-\infty}^{+\infty} x[n] * y[n+m] \quad (3)$$

Where, m denotes the number of samples by which y[n]is delayed..

$$Avrg(i) = (R_{fp1,IC(i)} + R_{fp2,IC(i)})/2 \quad (4)$$

Distribution ratio: this is a second classifying vector and is equal to the ratio between the peak amplitude and the variance of each IC as depicted in equation 5.

$$Dist_{ratio} = \frac{\max(IC(i))}{\delta^2(\mathrm{IC(i)})} \quad (5)$$

For i= 1, 2… N. Where max(IC (i)) is the maximum value or amplitude of the i^{th} independent component and $\delta^2(\mathrm{IC(i)})$ is its variance.

Maximum value: this is a third feature vector and consists of calculating the maximum value of each independent component. Kmeans-ICA method for EOG denoising in multichannel EEG recordings has been detailed in [10].

E. Adaptive thresholding: Kmeans clustering

Independent components classification was done by applying Kmeans algorithm that starts by partitioning the set of ICs into two groups. Then, it calculates the centroids of each group and assigns each component to the group that has the closest centroids based on the squared Euclidian distance. Once all the components are assigned, it recalculates each centroid and iterates until both remain the same. Actually, it is an iterative partitioning in a way that the considered threshold will maximize the inter class variance and thus minimize the intra class variance [8].

F. Motor imagery based BCI

After EOG denoising, a right hand motor imagery classification was performed. It starts by extracting the features used as inputs of a statistical test before an LDA classification.

- Feature extraction :

This phase consists of extracting features that will display relevant information about brain dynamics and that are unique to each mental task. Three features are adopted in this study: the first is a traditional one and consists of calculating the band power in the mu rhythm (8-12Hz) band as depicted in equation 6. Since this BCI deal with the right hand motor imagery, the band power feature is calculated over C3 during relax and motor imagery.

$$X(t) = \langle x^2(t) \rangle_t \quad (6)$$

The second and the third features start by decomposing the signal using the discrete wavelets transform which is a multi-resolution analysis in which the resolutions compromise is overtaken by decomposing the signal into a basis of functions. When an EEG signal with a sampling frequency F_s is decomposed using the wavelet transform, the frequency bounds of the approximation and the details levels are computed as shown in equation 7.

$$[0, \frac{F_S}{2^{L+1}}], [\frac{F_S}{2^{L+1}}, \frac{F_S}{2^L}], \ldots\ldots\ldots,[\frac{F_S}{2^2}, \frac{F_S}{2}] \quad (7)$$

Considering a sampling frequency of 160 Hz and a Debauchee mother wavelet, the mu rhythm is depicted in the fourth details coefficient level (D4). The second feature consists of calculating the coherence between the electrodes C3 and C4 after DWT and will be known as DWT-coherence. Considering x and y as the Fourier coefficients of the two signals and w1 and w2 the bounds of frequency, the coherence is computed as depicted in equation 8 [12].

$$C_{xy} = \frac{0.5.\sum_{w1}^{w2} \bar{x}^* \bar{y} + \bar{x}\bar{y}^*}{\sqrt{\sum_{w1}^{w2} \bar{x}\, \bar{x}^* . \sum_{w1}^{w2} \bar{y}\, \bar{y}^*}} \quad (8)$$

The third feature uses a large scale Phase locking value between the electrodes C3 and Fz after DWT and will be marked as DWT-PLV. The PLV is used as a reliable measurement of phase synchrony and coupling between two signals as depicted in equation 9[13].

$$PLV_{Large} = \left|\langle e^{j\Delta\emptyset(t)} \rangle\right| \quad (9)$$

Where $\Delta\emptyset(t)$is the phase difference between the two signals computed using the Hilbert transform.

- Statistical test: t-test

As it name implies, it is a statistical testing used to deduce if there is a significant difference regarding a feature or if the apparent difference is a sampling artifact. A student test is used in this study in which the statistic is based on the mean of the two sets as depicted in equation 10[13].

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{10}$$

Where x_i is the mean, s_i is the standard deviation and n_i is the number of samples of the i^{th} set. In this way, the null hypothesis is that the mean difference between the paired observations is zero. If the test has rejected the null hypothesis this means there is a significance difference.

- Linear Discriminant Analysis classifier

This classifier separates the data that represents different classes by means of a linear hyper plane. In fact, the separation is performed by finding the projection that will maximize the difference between the classes' means and therefore minimizes the interclass variance. The discriminant function requires the projection of the n-dimensional data into a line as shown in equation 11[13].

$$y = W^T \tag{11}$$

The slope or the orientation of the line is modified in order to separate the data into the given categories.

III. RESULTS AND DISCUSSION

All mentioned methods were implemented on the EEG dataset from the international database on Matlab© as well as in EEGLab[11].

A. *Manual ocular artifact identification*

Once the ICA decomposition of the 64 input channels is done, the 2D scalp map of each component was plotted, the power spectrum as well as the ICs time series. The artifactual ICs were identified by a field expert. Components 1, 3, and 6in the considered dataset were assigned as ocular artifacts since there scalp maps show a far frontal projection typical of an eye artifact; furthermore, they occupy a low frequency range what was verified when plotting their power spectrum.

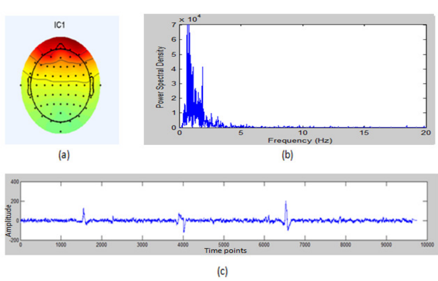

Figure 1 Example of an ocular artifact IC. (a) Scalp map; (b) Power spectrum; (c) signal time series

B. *Automatic ocular artifact identification*

- Average Correlation:

As expected, the results suggest that the components 1, 3 and 6 have the highest average correlation values. Fig. 2 shows the results of the classification based on Kmeans clustering; components 1, 3 and 6 were grouped in the first class which is the artifactual class.

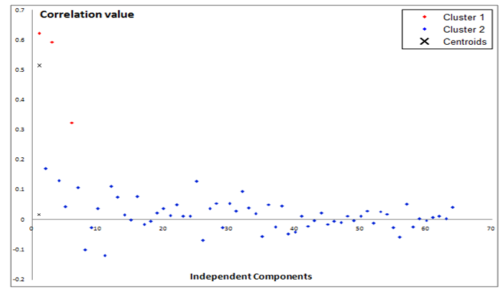

Figure 2 Kmeans clustering based on the correlation feature.

- Distribution ratio:

While components 1, 3 and 6 have high values and are close to the artifact cluster centroid (10.41), some non-ocular artifact ICs also show high values.

- Maximum Value:

Components 1, 3 and 6 have the highest values regarding this feature. Kmeans clustering was performed on 3 classes. All components of class 1 are considered as artifactual.

C. *Motor Imagery based BCI*

- Feature extraction

The three considered features show a decrease during right hand motor imagery. Since these features are not correlated and each gives specific information about the brain states, they will be used together for the classification

- Statistical test : t-test

The t statistic test is chosen to evaluate the significant and non-significant differences between the features during relax and motor imagery. This significance difference testing is used in order to delimit the use of features used in the classification and in order to increase the true positivity and the accuracy of the classifier. Generally, the t-test is used by computing the p-value using the t-distribution. The threshold for the t statistic is chosen in order to minimize the number of false positives and to omit some features and maintain the crucial ones in the classification. The threshold for this test is set to 0.5. Table 1 shows the results of the statistical test computed for the mentioned features after ocular artifact removal. As shown in table 1, the main hypothesis has been rejected for the above mentioned features conducting thus a significance difference for the Band Power, DWT-coherence and DWT-PLV between the relax and the motor imagery states.

Table 1 t-test results after ocular artifact removal

T-test	Feature		
	Band Power	***DWT-coherence***	***DWT-PLV***
P	6.1795 e^{-07}	0.0068	4.1535 e-04
Result	Hypothesis rejected	Hypothesis rejected	Hypothesis rejected

.

- Linear Discriminant Analysis classifier

Leave One Out cross validation (LOOV) was implemented using an LDA classifier on Matlab©. It was performed for all the mentioned feature vectors as well as combination of vectors. The best accuracy was obtained while mixing the three feature vectors together leading to an average accuracy of 88.10% computed on 50 signals from three different subjects. On the other hand, for evaluating the classification accuracy improvement after ocular artifact removal, the LOOV was performed on the contaminated EEG signals. The maximal classification accuracy reached 66% while combining the three features together. Therefore, the classification accuracy has increased with the ocular artifacts removal.

IV. Conclusion and Perspectives

In this work, a denoising and feature extraction algorithm has been proposed for the classification of the right hand motor imagery. It uses an automatic ocular artifact rejection method based on ICA without the use of an additional reference channel such as an EOG channel. Ocular artifacts identification is done using an adaptive thresholding by means of Kmeans clustering also known to be an unsupervised machine learning technique. After ocular artifact removal, the denoised EEG signals are used for a motor imagery classification based on three physiologically related features depicting the band power, the coherence and the phase locking value. The results show that the ocular artifacts removal increases the classification accuracy from 66 to 88.10% when using an LDA classifier considering the inter-session and inter-subject variability. As future work, we aim to improve the accuracy of the classifier by adding prominent features or testing other classifiers and finally we aim to test this algorithm on real data acquired in our laboratory.

References

[1] J. Chambers, S. Sanei, *EEG Signal Processing. (*Chichester, England : John Wiley & Sons, Ltd., 2007).

[2] Eleni Kroupi, Ashkan Yazdani, Touradj Ebrahimi, Jean-Marc Vesin**,** Ocular artifact removal from EEG: a comparison subspace projection and adaptive filtering method*. Eusipco, Proceedings of 19th European Signal Processing Conference*, Barcelona, Spain, 2011, EPFL-CONF-167489.

[3]. Garrick Wallstroma, Robert Kassb, Anita Millerc, Jeffrey Cohn, Nathan Fox, Automatic correction of ocular artifacts in the EEG: a comparision of regression-based and component-based methods, *International Journal of Psychophysiology*,2004,53,105-119.

[4].Mehrdad Fatourechi, Ali Bashashati, Rabab K Ward, and Gary E Birch**,** EMG and EOG Artifacts in Brain Computer Interface Systems: A Survey,*Clinical Neurophysiology*, 118, 2007.

[5]. Joyce Carrie, Irina Gorodnitsky, Marta Kutas*,* Automatic removal of eye movement and eye blink artifacts from EEG data using blind component seperation,*Psychophysiology, 41*(2), 2004, 315-325.

[6]. S. Romero , MA. Mananas, J. Riba, A. Morte, S. Gimenez, S. Clos, MJ. Barbanoj,Evaluation of an automatic ocular filtering method for awake spontaneous EEG signals based on independent component analysis. *Proc. 26th Annual International Conference of the IEEE,*San Francisco, CA,2004, 925-928.

[7]. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE**.***"Components of a New Research Resource for Complex Physiologic Signals."* Circulation 101(23):e215-e220. *PhysioBank, PhysioToolkit, and PhysioNet.* [Online] June 13, 2000 . http://circ.ahajournals.org/cgi/content/full/101/23/e215.

[8]. Clemens Brunner, Arnaud Delorme,Scott Makeig*,* Eeglab – an Open Source Matlab Toolbox for Electrophysiological Research, *Biomedical Engineering / Biomedizinische Technik,* 2013.

[8]. Rangaraj M. Rangayyan, Biomedical Signal Analysis.(Wiley-IEEE Press,Canada, 2002).

[9]. J. Kelly, D. Siewiorek, A. Smailagic, J. Collinger, D. Weber and W. Wang, Fully Automated Reduction of Ocular Artifacts in High-Dimensional Neural Data,*IEEE Trans. on Biomed. Eng., 58,* 2011,*598 - 606.*

[10] E. Bou Assi, S. Rihana, Kmeans-ICA based automatic method for EOG denoising in multi-channel EEG recordings, *Proc. 11th IASTED International Coneference on Biomedical Engineering,* Zurich, Switzerland, 2014.

[11]. M. Gyanaprava, B.Kumar and M. Asit Kumar**,** Denoising of heart sound signal using wavelet transform,*International Journal of Research in Engineering and Technology*, 2(4), 2013, 2319 - 1163.

[12]. A.Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independentcomponent analysis,*Journal of Neuroscience Methods*,*134*(1) 2004, 9-21.

[13]. Anthony Bell, Terrence Sejnowski**,** An Information-Maximization Approach to Blind Separation and Blind Deconvolution,*Neural Computation, 7*(6),1995,1129-1159.

[14].Yijun Wang, Bo Hong, Xiaorong Gao, and Shangkai Gao, Phase Synchrony Measurement in Motor Cortex for classifying single-trial EEG during Motor Imagery, *Proceedings of the 28th IEEE EMBS International conference, New York,2006*

[15].D. Hudson, M. Cohen, *Neural Networks and Artificial Intelligence for Biomedical Engineering* (IEEE Press,New York City,2000)