Automatic Removal of Ocular Artifacts in the EEG without an EOG Reference Channel

Germán Gómez-Herrero[1†], *Wim De Clercq*[2*], *Haroon Anwar*[1], *Olga Kara*[1], *Karen Egiazarian*[1], *Sabine Van Huffel*[2*], *and Wim Van Paesschen*[2⋆]

[1]Tampere University of Technology
Institute of Signal Processing
P.O. Box 553, 33101 Tampere
FINLAND
[†]E-mail: german.gomezherrero@tut.fi

[2]Katholieke Universiteit Leuven
[*]Dept. of Electrical Engineering (ESAT/SCD) and
[⋆]Dept. of Neurology, University Hospital Gasthuisberg
BELGIUM

ABSTRACT

This paper presents a method for removing electroocular (EOG) artifacts in the electroencephalogram (EEG). The procedure is based on blind source separation (BSS) and, in contrast to methods already available in the literature, it is completely automated and does not require the availability of peri-ocular EOG electrodes. The proposed approach removed most EOG artifacts in 6 long-term EEG recordings containing epilectic seizures without distorting the recorded ictal activity.

1. INTRODUCTION

Eye movements and blinks produce electrical potentials that propagate over the scalp creating significant electrooculographic (EOG) artifacts in the recorded electroencephalogram (EEG).These artifacts often complicate the interpretation of the EEG. EOG artifacts obscuring the EEG at the time of seizure onset may be problematic in the setting of a preoperative evaluation of patients with refractory epilepsy, since the ictal recordings are crucial for the localization of the epileptogenic zone. Furthermore, EOG artifacts are a source of false positives in automatic seizure detection systems. Some epileptic seizures appear very infrequently and very long EEG recordings are necessary to capture them. Therefore, manual analysis of such huge amount of data is very time consuming and fully-automated processing methods are highly desirable.

Traditionally, EOG artifacts have been corrected using regression-based methods which assume that signals recorded from electrodes placed near the eye collect clean EOG free of EEG contamination. The corrected EEG is obtained by regressing out these reference EOG signals from the signals recorded at the scalp electrodes [1–3]. However, in an Epilepsy Monitoring Unit (EMU) the EOG is not always recorded simultaneously with the EEG since the EOG electrodes are cumbersome for the patient who is monitored for typically one week. As a consequence, processing methods should not rely on EOG recording.

Other group of methods are based on decomposing the recorded EEG signals into spatial components that isolate the artifacts and the cerebral activity. Classical approaches for performing such spatial decompositions are Principal Component Analysis (PCA) [4] and Singular Value Decomposition (SVD) [5]. More recently, Blind Source Separation (BSS) techniques and specially Independent Component Analysis (ICA) have become very popular for separating the EEG and EOG components [6, 7]. Apart from an accurate spatial decomposition, fully-automated component-based methods require of robust criteria for identifying EOG-related components. Several criteria have been proposed in the literature that require the availability of one or more EOG channels (e.g. [2, 8]). In [9], a criteria based on Singular Value Fraction (SVF) that did not require an EOG reference was used to select interesting components for epileptic seizure detection. However, the latter approach requires the user to manually select an SVF threshold which critically determines the accuracy of the selection. Furthermore, the authors of [9] recognize the difficulties of finding a robust threshold. In contrast, our method is truly automatic since the user is not required to select any critical analysis parameter.

2. METHODS

2.1. The EEG inverse problem

Typically, EEG observations are obtained at the output of an array of scalp electrodes, where each sensor receives a different combination of source signals of ocular (the "EOG" sources) and neural origin (the "EEG" sources).

Let $\mathbf{s}(t) = [s_1(t), ..., s_n(t)]^T$ be the n original unobserved sources, and let $\mathbf{x}(t) = [x_1(t), ..., x_m(t)]^T$ be the m mixtures observed at the electrodes. Here, as in the following, ^{T} denotes transposition and t denotes the sample index $t = 1, ..., L$. Let us call Γ_{EEG} the set of indexes such that $s_i(t)\ \forall i \in \Gamma_{EEG}$ are the source potentials of neural origin. Similarly, let Γ_{EOG} be the set of indexes such that $s_i(t)\ \forall i \in \Gamma_{EOG}$ are the sources of ocular origin. Most EOG correction methods assume that the signal recorded at the jth electrode can be modeled as the following instantaneous mixture:

$$\begin{aligned} x_j(t) &= x_{j,EEG}(t) + x_{j,EOG}(t) \\ &= \sum_{i\in\Gamma_{EEG}} a_{ji}s_i(t) + \\ &\quad + \sum_{i\in\Gamma_{EOG}} a_{ji}s_i(t) \end{aligned} \quad (1)$$

where a_{ji} is the transfer coefficient from the ith source to the jth scalp electrode. This model can be compactly expressed using matrix notation as:

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{x}_{EEG}(t) + \mathbf{x}_{EOG}(t) = \mathbf{A}(t)\mathbf{s}(t) \\ &= \mathbf{A}_{EEG}(t)\mathbf{s}_{EEG}(t) + \mathbf{A}_{EOG}(t)\mathbf{s}_{EOG}(t) \end{aligned} \quad (2)$$

Using Eq. (2) we can formally express the EEG inverse problem as the problem of estimating $\mathbf{x}_{EEG}(t)$ from $\mathbf{x}(t)$.

2.2. EOG correction through spatial filtering

Spatial filtering methods address the EOG inverse problem in three steps:

1. Estimate the mixing matrix $\mathbf{A}$ using a finite set of observed data $\mathbf{x}(t)$, $t = 1, ..., L$.

2. Identify the columns of $\mathbf{A}$ corresponding to artifactual and neural EEG components, i.e. identify the sub-matrices $\mathbf{A}_{EEG}$ and $\mathbf{A}_{EOG}$ from Eq. (2).

3. Recover the clean neural EEG activity by means of the following spatial filter:

$$\hat{\mathbf{x}}_{EEG}(t) = \mathbf{A}_{EEG}\mathbf{A}^{\#}_{EEG}\mathbf{x}(t) \quad (3)$$

   where # denotes the Moore-Penrose pseudoinverse.

Provided that the time courses of the ocular and neural electrical activity are mutually independent, matrix $\mathbf{A}$ can be estimated using spatial ICA, which is a BSS approach that considers the source signals as mutually independent random variables. An alternative BSS method is Second Order Blind Identification (SOBI) [10], which assumes that the source signals are temporally uncorrelated to each other but they have non-zero time-delayed auto-correlations. Under these assumptions, SOBI computes the mixing matrix as the matrix that jointly diagonalizes a set of p cross-correlation matrices $\mathbf{R}(\tau_i) = E\left[\mathbf{x}(t)\mathbf{x}(t-\tau_i)^T\right]$, where $i = 1, ..., p$, and $E[\ ]$ is the expectation operator. In this paper we use $p = \lfloor L/3 \rfloor$ where L is the number of EEG data samples.

In general, BSS makes very weak assumptions about the mixing matrix, namely that it is of full column-rank. By contrary, PCA requires the columns of $\mathbf{A}$ to be mutually orthogonal which is very restrictive since the scalp activity maps of EOG and frontal EEG sources are likely to be non-orthogonal to each other. A major disadvantage of ICA-based methods is that they require the estimation of complex statistical measures of independence which makes them rather sensible to model inaccuracies. By contrary, SOBI is based on simple second-order statistical quantities which are much easier to estimate and more robust to modeling errors. Furthermore, several neuroscientific studies support the use of SOBI for separating EEG and EOG components [7, 8]. Because of this, the first step of our EOG correction approach consists on performing a SOBI decomposition of correlative EEG frames (or data windows). By using this moving window analysis scheme we aim to partially cope with the non-stationary nature of EEG signals. The length of the frames should be enough for allowing an accurate estimation of the mixing matrix coefficients but otherwise does not critically determine the accuracy of the correction. As a simple rule of thumb, we recommend using as frame length at least $0.25 \times m^2$ seconds, where m is the number of scalp electrodes. In the correction results shown below we analyzed the data in correlative non-overlapping frames of 200 seconds.

2.3. Identifying ocular components

The spectrum of ocular electrical activity is typically characterized by few predominant low-frequency components. Such type of signals can be easily identified by a low fractal dimension (FD) which is a measure of signal complexity. By contrary, neural EEG traces are typically characterized by a flatter and more spread spectrum which accounts for higher FDs [11]. The FD can be calculated using several methods. We decided to use Sevcik's algorithm [12] because it allows a fast computation and is quite robust to the presence of noise.

If a waveform has coordinates (x_i, y_i), Sevcik's algorithm first maps it into a unit square by $x_i^* = \frac{x_i}{x_{max}}$, $y_i^* = \frac{(y_i - y_{max})}{(y_{max} - y_{min})}$, where x_{max} is the maximum x_i, and y_{max} and y_{min} are the maximum and minimum y_i. Then, the FD is computed as $FD = 1 + \frac{\ln(l)}{\ln(2(n-1))}$, where l is the total length of the waveform and n is the number of points of the waveform.

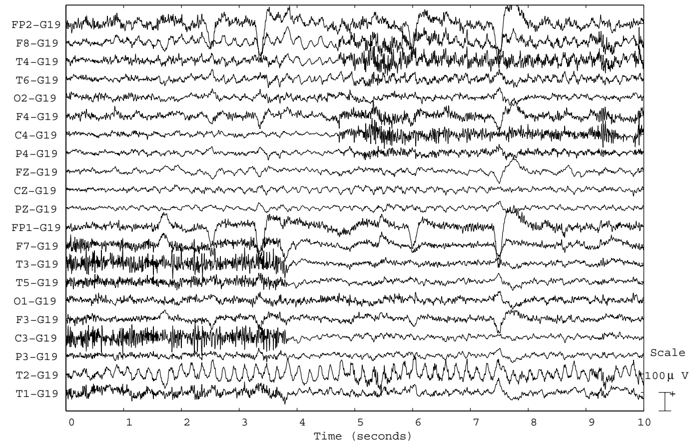
In order to obtain a more robust estimate from non-stationary and noisy signals like the EEG we propose to compute the FD of an EEG signal as the mean value of the FDs obtained in correlative frames of length equal to 10% of the total length of the signal. We will refer to this estimated FD as *mean FD* (mFD). After analyzing various real EEG datasets we have confirmed that the mFDs of EOG-like waveforms are consistently lower than those of EEG-like waveforms. However, the actual values of the mFDs depend on experimental factors such as the sampling rate and pre-processing filter used to acquire the EEG data. Therefore, it is not possible to select a universal mFD threshold below which a component is to be considered of EOG nature. To overcome this problem we propose the following approach for selecting the EOG-components among the components estimated by SOBI in each analyzed EEG frame:

1. Sort the components according to increasing values of their corresponding mFD. Let us denote the sorted components by $s^{(1)}(t), ..., s^{(N)}(t)$ and their corresponding mFDs by $\phi^{(1)}, ..., \phi^{(N)}$. Note that the number of components estimated by SOBI is $N = \min(n, m)$ being m the number of data channels and n the number of hidden sources. Note also that for $\mathbf{A}$ to be of full column-rank it is necessary that $m \geq n$.

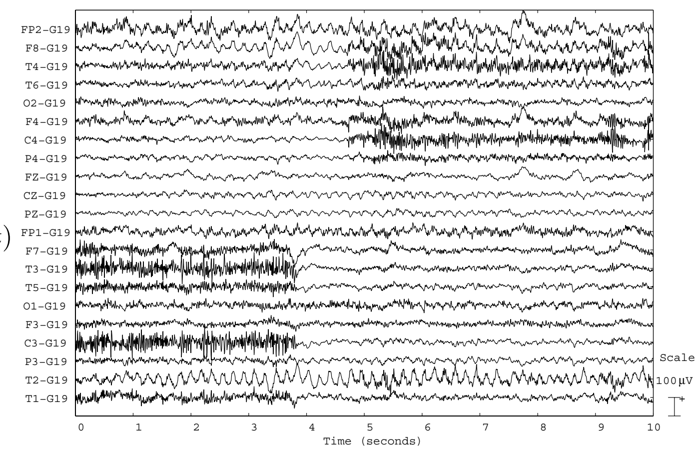
2. Identify as EOG components $s^{(1)}(t), s^{(2)}(t), ..., s^{(k)}(t)$ where k is the smallest integer in the range $\lfloor N/2 \rfloor \geq k > 1$ such that $\left(\phi^{(k+1)} - \phi^{(k)}\right) < \left(\phi^{(k)} - \phi^{(k-1)}\right)$. If there is not any k in the specified range satisfying the required condition then $k = 1$.

3. RESULTS

Six long-term EEG recordings, recorded at an EMU, were used in this study. The data was collected from 21 scalp electrodes placed according to the international 10-20 System with additional electrodes *T1* and *T2* on the temporal region. The sampling frequency was 250 Hz and an average reference montage was used. The EEG recordings contain ictal activity from patients with Mesial Temporal Lobe Epilepsy (MTLE). Fig. 1 (a) shows a 10 seconds EEG epoch of one of the recordings used in this study. This EEG epoch contains the activity of the seizure onset and this epileptic theta activity is mainly present on the *T2*, *F8* and *T4* electrodes. The seizure EEG is contaminated with EOG and muscle artifacts. Fig. 1 (b) shows the same ictal EEG recording after EOG artifact correction by the proposed method. Notice that the EOG artifact was removed, while preserving the sharp quality ictal theta activity on the *T2*, *F8* and *T4* electrodes. A detailed analysis of the corrected recordings showed that virtually all blinking artifacts and most slow eye movements were removed. Fig. 2 shows another correction example from one of the analyzed files. More snapshots from every corrected recording as well as the Matlab code used to perform the correction are available in the internet [13].



(a)



(b)

Fig. 1. (a) Original 10 seconds EEG epoch; (b) EEG recording after EOG artifact correction. Notice that the EOG artifact was removed completely, while preserving the sharp quality ictal theta activity in the right temporal lobe (*T2* , *F8*, and lesser degree *T4*).

4. CONCLUSIONS

The proposed automatic EOG correction technique was applied on 6 long-term EEG recordings containing ictal activity. In all cases, most EOG artifacts were removed without altering the recorded ictal activity. Therefore, the presented approach is expected to be highly beneficial for increasing the accuracy of seizure zone localization and for reducing the false positive rates in automatic epileptic seizure detection systems. Although the method does not require any EOG reference channel, we have observed that the accuracy of the separation produced by SOBI, and therefore of the final correction, is clearly improved when EOG channels are included in the estimation of the hidden source signals. Thus, a topic

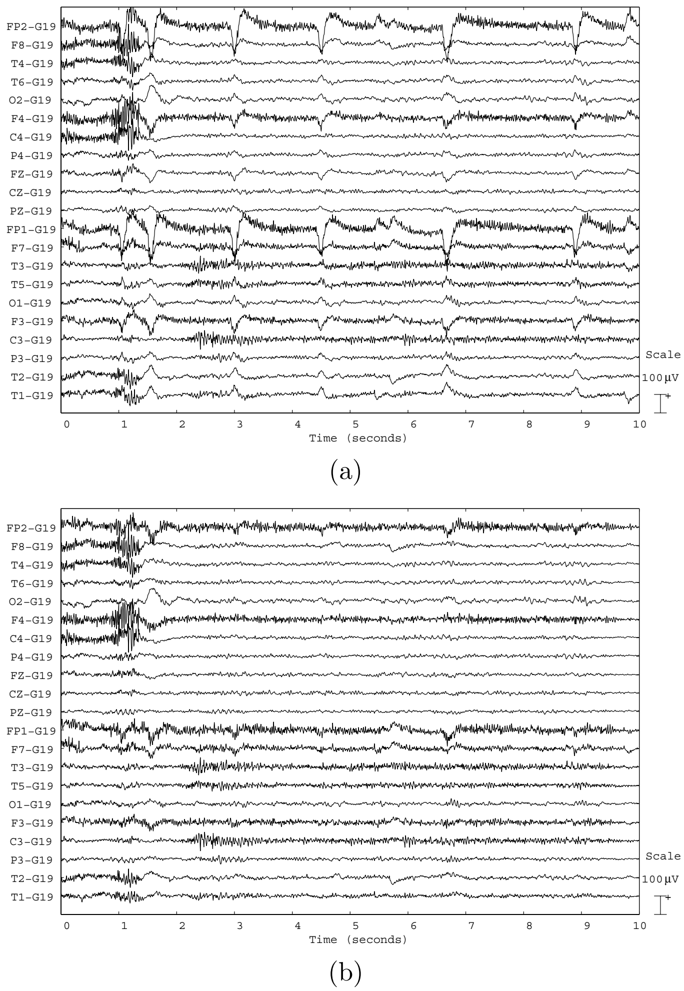


Fig. 2. (a) Original 10 s EEG epoch; (b) corrected EEG epoch.

for further research is to compare the performance of the proposed approach with classical automatic correction techniques on a dataset with peri-ocular EOG channels.

5. ACKNOWLEDGEMENT

This study was financially supported by the EU project No 508803, Biopattern Network of Excellence, by the Academy of Finland, project No 44876 (Finnish Centre of Excellence Program, 2000-2005), by the FWO through project G.0360.05 (EEG signal analysis for epilepsy monitoring) and by the Research Council of the Katholieke Universiteit Leuven (GOA-AMBioRICS).

6. REFERENCES

[1] R. J. Croft and R. J. Barry, "Removal of ocular artifact from the EEG: a review," *Neurophysiol. Clin.*, vol. 49, pp. 5–19, 2000.

[2] Garrick L. Wallstrom, Robert E. Kass, Anita Miller, Jeffrey F. Cohn, and Nathan A. Fox, "Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods," *Int. J. of Psychophysiol.*, vol. 53, pp. 105–119, 2004.

[3] S. Puthusserypady and T. Ratnarajah, "H^∞ adaptive filters for eye blink artifact minimization from electroencephalogram," *IEEE Sig. Proc. Lett.*, vol. 12, no. 12, pp. 816–819, 2005.

[4] T. D. Lagerlund, F. W. Sharbrough, and N. E. Busacker, "Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition," *J. Clin. Neurophysiol.*, vol. 14, pp. 73–82, 1997.

[5] P. K. Sadasivan and D. N. Dutt, "SVD based technique for noise reduction in electroencephalographi signals," *Signal Processing*, vol. 55, pp. 179–189, 1996.

[6] T-P. Jung, S. Makeig, C. Humphries, T-W. Lee, M. J. McKeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, pp. 163–178, 2000.

[7] A. C. Tang, M. T. Sutherland, and C. J. McKinney, "Validation of SOBI components from high-density EEG," *NeuroImage*, vol. 25, pp. 539–553, 2004.

[8] C. A. Joyce, I. F. Gorodnitsky, and M. Kutas, "Automatic removal of eye movement and blink artifacts from eeg data using blind component separation," *Psychophysiology*, vol. 41, pp. 313–325, 2004.

[9] S. Faul, L. Marname, G. Lightbody, G. Boylan, and S. Connolly, "A method for the blind separation of sources for use as the first stage of a neonatal seizure detection system," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2005, vol. 5, pp. 409–412.

[10] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Transactions on Signal Processing*, vol. 45, pp. 434–444, 1997.

[11] A. Accardo, M. Affinito, M. Carrozzi, and F. Bouquet, "Use of the fractal dimension for the analysis of electroencephalographic time series," *Biol. Cybern.*, vol. 77, pp. 339–350, 1997.

[12] C. Sevcik, "A procedure to estimate the fractal dimension of waveforms," *Complexity International*, vol. 5, 1998, Online: http://journal-ci.csse.monash.edu.au/ci/vol05/sevcik/.

[13] "Automatic artifact removal toolbox for Matlab," 2006, http://www.cs.tut.fi/~gomezher/software/.