HAL
open science

Iterative Subspace Decomposition for Ocular Artifact Removal from EEG Recordings

Cedric Gouy-Pailler, Reza Sameni, Marco Congedo, Christian Jutten

To cite this version:

Cedric Gouy-Pailler, Reza Sameni, Marco Congedo, Christian Jutten. Iterative Subspace Decomposition for Ocular Artifact Removal from EEG Recordings. ICA 2009 - 8th International Conference on Independent Component Analysis and Signal Separation, Mar 2009, Paraty, Brazil. pp.419–426, 10.1007/978-3-642-00599-2 . hal-00369922

HAL Id: hal-00369922

https://hal.science/hal-00369922v1

Submitted on 23 Mar 2009

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.

Iterative Subspace Decomposition for Ocular Artifact Removal from EEG Recordings

Cédric Gouy-Pailler[1]*, Reza Sameni[2], Marco Congedo[1], and Christian Jutten[1]

[1] GIPSA-lab, DIS/CNRS/INPG-UJF-Stendhal
Domaine Universitaire, BP 46, 38402 Saint Martin d'Hères Cedex, FRANCE
[2] College of Electrical and Computer Engineering, Shiraz University, Shiraz, IRAN
Email: cedric.gouypailler@gipsa-lab.inpg.fr

Abstract. In this study, we present a method to remove ocular artifacts from electroencephalographic (EEG) recordings. This method is based on the detection of the EOG activation periods from a reference EOG channel, definition of covariance matrices containing the nonstationary information of the EOG, and applying generalized eigenvalue decomposition (GEVD) onto these matrices to rank the components in order of resemblance with the EOG. An iterative procedure is further proposed to remove the EOG components in a deflation fashion.

1 Introduction

Electroencephalography (EEG) is a widely used technique for analyzing and interpreting human cerebral activity. EEG signals are usually interpreted by means of spectral and topographical measures that reflect global activity of the brain network. However, EEG measures are always contaminated by non-cerebral signals, which may disturb the interpretation of the brain activity. This issue has become a recurrent problem, for example in Brain-Computer Interfaces (BCI), where it has been proved to decrease classification error rates [11].

Ocular artifacts generally occur during blinking or saccades of the eye, and are featured by high amplitude transient artifacts that defect the EEG. They are best recorded by an Electrooculogram (EOG) electrode or a pair of electrodes located close to the eyes. The high amplitude peaks are not seen on all channels, but mainly (and almost exclusively) on the fronto-paretal channels in combination with the occipital electrodes. These peaks are considered as one of the most considerable artifacts in EEG studies [7].

A common way of removing these artifacts is to apply independent component analysis (ICA) on multichannel EEG recordings and to remove the components which show a maximal correlation with a reference EOG channel [6, 5]. However, it is not always possible to associate the components extracted by ICA

* This work was supported by Égide, the French Ministry of Defense (Délégation Générale pour l'Armement – DGA), the Open-ViBE project, and the European program of Cooperation in the Field of Scientific and Technical Research.

to the EOG in an automatic and unsupervised manner. Moreover, EEG recordings can be rather noisy, and since ICA is based on a measure of independence (and not a measure of signal "cleanness"), the noise in the input channels can be amplified in the output, which again makes the detection of the true EOG component rather difficult. Lastly, most ICA methods are blind to Gaussian noise and as a consequence spread the Gaussian noise among the components.

Another common way of removing EOG artifacts from the EEG is to use a subtraction-based approach [8]. Here, the idea is to use the clean EOG recordings to remove ocular signals from EEG by a simple subtraction of a scaled EOG. Yet there is no evidence that EOG recordings are free of EEG. Thus by subtracting EOG we can also remove EEG signals of interest.

In recent works we have shown the applicability of Generalized Eigenvalue Decomposition (GEVD) for separating pseudo-periodic maternal ECG from fetal ECG signals recorded from the maternal abdomen [9]. In that work, one of the advantages of GEVD over other source separation techniques was the ability of ranking the extracted components in order of periodicity, which provided a means of automatic and unsupervised ECG decomposition and filtering. In this work, we develop a similar idea based on GEVD for the automatic detection and removal of EOG artifacts from multichannel EEG recordings.

The remainder of this paper is organized as follows: in section 2, we present a general nonlinear framework to decompose signals into independent subspaces using GEVD; the results of this method are presented in section 3 over simulated and real signals. The last section is devoted to conclusion and perspectives.

2 Method

2.1 Linear Transform

Suppose that we have an array of N EEG channels $\mathbf{x}(t)$, and a reference EOG channel denoted by $\mathrm{EOG}(t)$. Due to the spiky nature of the EOG, it is possible to detect the onset and offset times of the EOG artifact from the reference EOG channel. To do so, we define $\mathrm{E}(t)$, the averaged power of the EOG signal (or, alternatively the variance) within a sliding window of length w around t, as

$$\mathrm{E}(t) \doteq \frac{1}{w} \sum_{\tau=-w/2}^{w/2} \mathrm{EOG}(t-\tau)^2 \ . \tag{1}$$

Using this definition, an EOG is detected whenever $\mathrm{E}(t)$ passes some predefined threshold th. The *active periods* of the EOG may therefore be defined as

$$t_a \doteq \{t | \mathrm{E}(t) > th\} \ . \tag{2}$$

As we will note later, the procedure of finding the offsets and onsets of the EOG does not need to be perfect, and the results can be further improved in a recursive procedure.

We now seek linear transforms of the multichannel recordings $\mathbf{x}(t)$, that maximally resemble the EOG (in the sense that the power of the extracted signals are concentrated during the active time periods). Therefore, if we denote the linear mixture as $y(t) = \mathbf{b}^T\mathbf{x}(t)$, we can try to maximize the cost function

$$\zeta(\mathbf{b}) = \frac{\mathbb{E}_{t_a}\left\{y^2(t_a)\right\}}{\mathbb{E}_t\left\{y^2(t)\right\}} \ , \tag{3}$$

where $\mathbb{E}_t\{\cdot\}$ represents averaging over t. Here, the idea is to find linear mixtures of the input signals, with a maximal energy during the EOG activation time t_a, while minimizing the global component energy. Equation (3) may be rewritten as

$$\zeta(\mathbf{b}) = \frac{\mathbf{b}^T\mathbb{E}_{t_a}\left\{\mathbf{x}(t_a)\mathbf{x}(t_a)^T\right\}\mathbf{b}}{\mathbf{b}^T\mathbb{E}_t\left\{\mathbf{x}(t)\mathbf{x}(t)^T\right\}\mathbf{b}} \ , \tag{4}$$

which is in the form of the Rayleigh Quotient [10, 3], and may be solved by the joint diagonalization of two covariance matrices: the covariance matrix of the EEG channels over the whole dataset, and the covariance matrix of the data during the active periods of the EOG, respectively defined as

$$C_x \doteq \mathbb{E}_t\{\mathbf{x}(t)\mathbf{x}(t)^T\} \ , \tag{5}$$

$$A_x \doteq \mathbb{E}_{t_a}\{\mathbf{x}(t_a)\mathbf{x}(t_a)^T\} \ . \tag{6}$$

The intuition behind this method is to achieve decorrelated components that are at the same time *globally* and *locally* decorrelated. We already have a sense about the global decorrelation, which is achieved by sphering C_x. In addition, the diagonalization of A_x assures that the achieved components are also locally decorrelated over the active EOG epochs, too. This assures that the later extracted components have no redundancy up to second order statistics.

The matrix that jointly diagonalizes the matrix pair (A_x, C_x) is in fact the solution to the GEVD problem written as

$$\begin{cases} UA_xU^T = \Lambda \\ UC_xU^T = I \end{cases} \ , \tag{7}$$

where Λ is a diagonal matrix containing the generalized eigenvalues on its diagonal in descending order, and U is the matrix containing the generalized eigenvectors on its columns[3].

The decomposed signals may now be found by

$$\mathbf{y}(t) = U^T\mathbf{x}(t) \ , \tag{8}$$

where the elements of $\mathbf{y}(t)$ correspond to a linear transformation of the original data $\mathbf{x}(t)$ that are ranked according to their resemblance with the EOG activation epochs. This means that $y_1(t)$ contains the most information regarding

[3] Note that, contrary to the eigenvalues of symmetric matrices that are mutually orthogonal, generalized eigenvectors, i.e., the columns of U, are not generally orthogonal to each other; but following (7) they are "C_x-orthogonal" [10], p. 344. It means that each column of U is orthogonal to every column of C_x.

the EOG while $y_N(t)$ is the least resembling the EOG. Another interpretation for the method is that $y_1(t)$ is the most dominant component of the EOG while having the least contribution in the overall energy of the signals (corresponding to the smallest eigenvalue of C_x). On the other hand $y_N(t)$ is the main non-EOG component, with the least EOG contamination (corresponding to the smallest eigenvalue of A_x, i.e., having the least contribution of the active EOG epochs).

2.2 Signal/Noise Separation

Up to now, the components have been ranked according to their resemblance with the EOG through a linear transformation. The next step is to remove the ocular artifacts from the most contaminated components using a linear or nonlinear transformation. The result of this noise removal, denoted by $\mathbf{z} \doteq (z_1, \cdots, z_N)^T$, can be expressed as

$$\forall i \in [1, ..., N] \quad z_i(t) = y_i(t) - f_i[y_i(t)] \ , \tag{9}$$

The denoising transform must be carefully chosen to remove ocular artifacts from the most contaminated signals, while preserving the non-EOG components. A simple but efficient choice is

$$f : \begin{cases} f_i(u) = u & , \quad i \in [1, ..., M] \\ f_i(u) = 0 & , \quad i \in [M+1, ..., N] \end{cases} \ . \tag{10}$$

This transform removes the first M ($M \ll N$) components contaminated by ocular artifacts and keeps the remaining $N - M$ components unchanged. Therefore, in order to find the EOG free components, we can eliminate the first few components of $\mathbf{y}(t)$ and transform back the rest of the components using the inverse of the matrix U. The number of eliminated components M depends on the number of expected dimensions of the EOG subspace. Note that due to the elimination of the first M components, the rank of the multichannel signals are reduced to $N - M$.

2.3 Iterative Improvements

So far, we have removed the most dominant EOG components through a combination of a linear transformation, denoising, and back-projection. The results may be further improved by repeating the upper mentioned method in a recursive procedure. To do so, we can use $y_1(t)$, the component which most resembles the EOG, to re-estimate the onset and offsets of the EOG and its activation epochs using a smaller energy threshold th, and to recalculate C_x, A_x, and the other steps of the algorithm in each iteration. By repeating this procedure in several iterations, a better estimate of the EOG will be achieved. This iterative extension of the algorithm is of special interest for the cases in which a good EOG reference is not available. Therefore, we start with a coarse EOG onset and offset estimate; but we improve this estimate in the next iterations.

3 Experiments

Simulated Data. In order to test the method, we generate artificial EEG contaminated by ocular artifacts. The signals are generated as follows:

$$\mathbf{x}(t) = \mathbf{EEG}(t) + \beta\,(k_1 \cdots k_N)^T\,\mathrm{EOG}(t)\ , \tag{11}$$

where (1) $\mathbf{EEG}(t)$ is generated using a multivariate autoregressive filter (Yule-Walker order $p = 8$, cf. [1] for details), trained using $N = 4$ channels of artifact-free EEG data, (2) $\mathrm{EOG}(t)$ results from the convolution of a typical blink segment, extracted from real EOG data, and an ensemble of Dirac impulses having a Poisson distribution with parameter $\lambda = 0.2\,Hz$, (3) β is a scale parameter to adjust the signal-to-noise ratio (SNR) and (4) $(k_1, \cdots, k_N)^T$ is a vector of gain under a linear hypothesis between EOG and EEG, for all i, $k_i \sim \mathcal{N}(1, 0.3)$. A multivariate autoregressive filter has been preferred over a multi-univariate autoregressive filter, because the multivariate filter naturally takes into account spatial correlations of the data and learns those correlations from the real data.

Our approach is compared with a classical ICA approach for EOG removal, based on FastICA [4]. The input of the algorithm consists of the N EEG channels $\mathbf{x}(t)$ and the EOG channel $\mathrm{EOG}(t)$. FastICA seeks the most independent $N + 1$ components. We then remove the component which is the most correlated to the EOG, by setting it to zero. Components are then back-projected onto the sensor space.

The method proposed in this paper and FastICA, abbreviated GEVD and ICA, respectively, are evaluated by comparing the first N components of the back-projected signals and the EEG generated by the multivariate autoregressive filter. The idea behind this evaluation procedure is that a perfect denoising would lead to perfect recovery of the multivariate autoregressive process $\mathrm{EEG}(t)$, free of any EOG contamination. We therefore define the error signals as a function of the denoised processes resulting from ICA and GEVD ($\mathbf{x}_{\mathrm{ICA}}$ and $\mathbf{x}_{\mathrm{GEVD}}$, respectively) as

$$\epsilon_{\mathrm{ICA}}(t) = \mathbf{x}_{\mathrm{ICA}} - \mathrm{EEG}(t) \tag{12}$$

$$\epsilon_{\mathrm{GEVD}}(t) = \mathbf{x}_{\mathrm{GEVD}} - \mathrm{EEG}(t)\ , \tag{13}$$

from which we can define the following performance index (in decibels):

$$\mathcal{Q} = 10 \log \left(\frac{1}{N} \sum_{i=1}^{N} \frac{\mathbb{V}(\epsilon_{\mathrm{GEVD}_i}(t))}{\mathbb{V}(\epsilon_{\mathrm{ICA}_i}(t))} \right)$$

where $\mathbb{V}(\cdot)$ denotes the variance operator. This index is built such that a positive value indicates that ICA outperforms GEVD, while a negative index indicates that GEVD outperforms ICA.

Figure 1 shows the results obtained for simulated data. Two signal lengths are considered (50000 and 100000 points). Two hundred simulations are done for each box-plot. The SNR tuned by β, is varying between -10dB and 30dB. These

values are chosen by considering that a blinking artifact can have an amplitude of up to 100 times stronger than the EEG. Such a condition would yield an SNR of between -10dB to 0dB, depending on the blinking rate.

In the GEVD method, only the first component given by GEVD is removed. This choice is due to the rather simple method used to generate the EOG and EEG mixture. Moreover, in this case the iterative EOG improvement method is not used, since the EOG onsets and offsets are already known. We highlight the fact that in our simulation scenario, the simulation condition for the ICA method is better than it is for real signals, since we provide it with the clean EOG channel, directly.

Our method based on GEVD is shown to clearly outperform a classical ICA approach for a large set of SNRs (t-test highly significant). This effect is much more significant for low SNRs, showing that our method is of particular interest for bad conditionings of the signals.

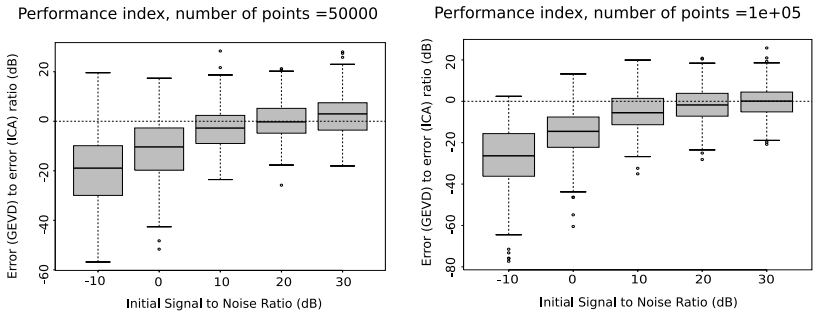

Fig. 1. Method comparison: for different values of initial signal to noise ratios between EEG and EOG, we performed 200 simulations. The performance index is positive when ICA performs better than GEVD whereas it is negative otherwise.

Real Data. We also evaluate our algorithm using a BCI experiment dataset[4] [2]. The data consists of 22 EEG channels and three EOG channels. The inter-electrode distance is about 3.5 cm. The signals are sampled with 250 Hz and bandpass filtered between 0.5 and 100 Hz. An additional 50 Hz notch filter is used to suppress power-line noise.

The back-projected signals after removing the first EOG component are depicted in Fig. 2. In this example, due to the quality of one of the EOG channel, no improvements are achieved by iterating the method. This figure only shows two of the 22 denoised EEG channels. The first one, Fz, is known to be highly contaminated by EOG artifacts because it is close to the eyes. Fig. 2 shows that the EOG component is removed from this channel and also that the EEG signal

[4] BCI competition IV, dataset 2a, subject A05 [2]

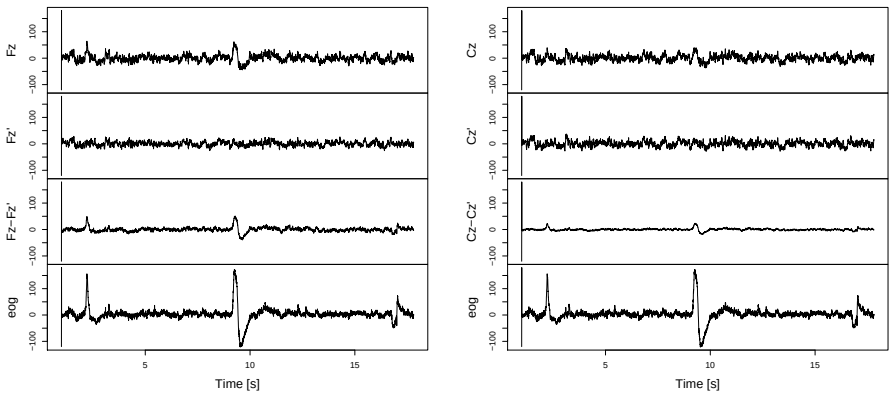

Fig. 2. Denoising of two specific channels, Fz and Cz electrodes (left and right, respectively). From top to bottom, the signals are (1) initial EEG, (2) denoised EEG, (3) residuals, and (4) EOG recordings.

outside blink contamination is kept perfectly unchanged. Channel Cz, known to be less contaminated than Fz is also drawn (Fig. 2 middle). While not as clear as in the first presented channel, we can see that the most important part of the EOG contamination is removed from the channel.

To illustrate the interest of the iterative method, we then show that due to the high amplitude features of the EOG contamination, we can even use some EEG channels to evaluate the active periods of EOG. This is of particular interest when no EOG channel is provided. Such a situation is illustrated in Fig. 3 where we have started with a noisy EEG reference and improved this reference after two iterations of GEVD.

4 Conclusion

We presented an automatic method to remove ocular artifacts from EEG measurements. Our algorithm outperformed a classical approach based on independent component analysis of EEG data. This method is a special case of a more general framework designed to generally decompose multivariate signals into independent subspaces. This flexible framework allows to choose different criteria for determining the optimal linear transform. In this paper, the nonstationarity of the EOG was used; but in other applications other prior information such as the periodicity or spectral contrast may be used.

In future works, further improvements can be achieved by replacing our simple component nulling function by some wavelet-based denoising method applied on the first few channels. In that case, it is expected that we separate the EOG from the EEG, without reducing the rank of the EEG signals.

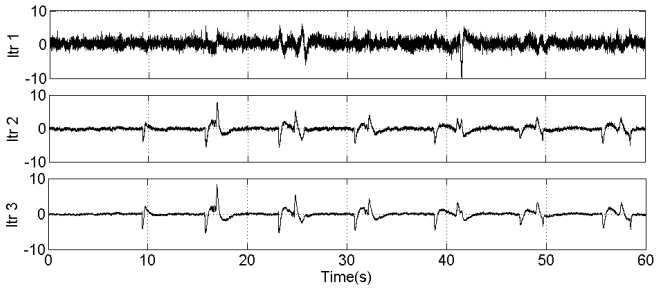

Fig. 3. Iterative improvements of the estimation of EOG. We start from a noisy EEG channel (up). The active periods is then determined according to the energy of the noisy EEG channel. We then select the first component to iterate the method: the active periods of the $(i+1)^{th}$ iteration are determined using the energy of the first component extracted in the ith iteration.

References

1. C. W. Anderson, E. A. Stolz, and S. Shamsunder. Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks. *IEEE Trans. Biomed. Eng.*, 45(3):277–286, Mar 1998.
2. C. Brunner, R. Leeb, G. R. Müller-Putz, A. Schlögl, and G. Pfurtscheller. BCI competion 2008 – Graz dataset A, http://ida.first.fraunhofer.de/projects/bci/competition_iv, Jul 2008.
3. G. Golub and C. van Loan. *Matrix Computations.* The Johns Hopkins University Press, third edition, 1996.
4. A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis.* Wiley-Interscience, 2001.
5. N. Ille, P. Berg, and M. Scherg. Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *J. Clin. Neurophysiol.*, 19(2):113–124, Apr 2002.
6. T. Jung, C. Humphries, T. Lee, M. McKeown, V. Iragui, S. Makeig, and T. Sejnowski. Removing electroencephalographic artifacts by blind source separation. *J. Psychophysiol.*, 37:163–178, 2000.
7. Z. J. Koles. The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG. *Electroencephalogr. Clin. Neurophysiol.*, 79:440–447, 1991.
8. S. Romero, M. A. Mañanas, and M. J. Barbanoj. A comparative study of automatic techniques for ocular artifact reduction in spontaneous EEG signals based on clinical target variables: A simulation case. *Comput. Biol. Med.*, Jan 2008.
9. R. Sameni, C. Jutten, and M. B. Shamsollahi. Multichannel electrocardiogram decomposition using periodic component analysis. *IEEE Trans. Biomed. Eng.*, 55(8):1935–1940, Aug. 2008.
10. G. Strang. *Linear Algebra and Its Applications.* Brooks/Cole, 3rd edition, 1988.
11. M. Thulasidas, C. Guan, S. Ranganatha, J. Wu, X. Zhu, and W. Xu. Effect of ocular artifact removal in brain computer interface accuracy. *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 6:4385–4388, 2004.