Multimodal Approach to Remove Ocular Artifacts from EEG Signals Using Multiple Measurement Vectors*

Victor Maurandi[1], Bertrand Rivet[1], Ronald Phlypo[1], Anne Guérin–Dugué[1], and Christian Jutten[1,2]

[1] Univ. Grenoble Alpes, GIPSA-Lab, F-38000 Grenoble France.
CNRS, UMR 5216, F-38000 Grenoble France.
[2] Institut Universitaire de France, 75231 Paris France.
{victor.maurandi,bertrand.rivet,ronald.phlypo,
anne.guerin,christian.jutten}@gipsa-lab.grenoble-inp.fr

Abstract. *This paper deals with the extraction of eye-movement artifacts from EEG data using a multimodal approach. The gaze signals, recorded by an eye-tracker, share a similar temporal structure with the artifacts induced in EEG recordings by ocular movements. The proposed approach consists in estimating this specific common structure using Multiple Measurement Vectors which is then used to denoise the EEG data. This method can be used on single trial data and can be extended to multitrial data subject to some additional preprocessing. Finally, the proposed method is applied to gaze and EEG experimental data and is compared with some popular algorithms for eye movement artifact correction from the literature.*

Keywords: Ocular artifact extraction, EEG, Gaze, Multiple Measurement Vectors, Multimodality.

1 Introduction

Electroencephalography (EEG) is a popular non-invasive method to monitor cerebral activity. It allows to measure the effect of electrical brain activity on the potential field at the scalp using surface electrodes. However, interpreting the recordings is challenging, in part due to different kinds of noise [1]. Among them, one finds the ocular artifacts that are induced by blinks or eye-movements, see, e.g., Iwasaki *et al.* (2005) [2] for an in-depth study on the topic. The most straightforward method to avoid these artifacts is to restrict subjects to move their eyes during the experimental recordings. However, this excludes experimental protocols where visual scene exploration or reading is a key aspect of the cognitive study.

For the last thirty years, a number of numerical methods to remove ocular artifacts have been considered in the literature. Among those, one finds the regression approach [3, 4] and Independent Component Analysis (ICA) [1, 5].

* This work has been partly funded by ERC-AdG-2012-320864 CHESS project.

The regression approach requires a reference for the ocular artifact. Usually, this method uses the electrooculogram (EOG) that provides a measurement of the electric field associated with the ocular activity which is recorded by electrodes localized in the vicinity of the eyes. In the regression approach, it is assumed that the EOG matches with the ocular artifacts contained in the EEG observations up to scaling factors. The goal is then to identify these factors and substract weighted reference channels from the EEG. Despite its simplicity, this method presents some major drawbacks. First, it needs additional electrodes for the reference channels to be available. Second, since EOG electrodes are also placed on the skin, volume conductivity of the latter results in cross-talk of cerebral activity and ocular artefact, even on the EOG electrodes. This implies a bias in the regression toward closeby electrodes.

Under the ICA model, a linear, instantaneous mixing model is estimated, since this is in line with the linearized Maxwell equations at the frequencies of interest. The latent sources constituting the EEG observations are assumed to be statistically independent. The goal of this method is to estimate the linear mixing operator and the latent sources through maximization of the source independence. Once the sources and the linear operator estimated, one can identify the ocular artifact components among these sources and remove their contribution from the EEG [1]. ICA has shown its efficiency and it is still widely used in the EEG community. Nevertheless, ICA also suffers from some drawbacks. First, it needs a large number of observations (large with respect to the number of electrodes) to accurately estimate the probability density functions or their approximations used in the computation of the independence criterion. In addition, since the sources are not truly independent, removing identified ocular artifact source components may result in the removal of cerebral activity, thus losing information of interest. Finally, since we consider only linear operators, suppression of the contribution of a source results in a decrease of the dimension of the signal subspace.

In this paper, we propose a novel method for the denoising of EEG data contaminated by eye-movement artifacts based on the multimodal nature of the gaze and EEG [6]. We focus on saccades, which are the eye-movements related to the action of moving from one fixation point to another. During a saccade, the EEG observations can be decomposed as a linear superposition of the electrical brain activity and a potential induced by the eye-movement (ocular artifact) [1]. In the meantime, an eye-tracker provides a measurement of the eye-movement (gaze direction relative to a screen). These gaze signals present a main advantage compared to EOG as they contain no brain activity. Since the gaze signals share a similar temporal structure with the ocular artifacts in the EEG observations [2], we consider the eye-tracker observations as reference signals for saccade denoising of EEG data. Motivated by the temporal similarity of eye-movement (artifacts) signals recorded from both modalities, we propose to use a Multiple Measurement Vectors method (MMV) [7] (also called Collaborative Lasso [8] or Multichannel Sparse Recovery [9]). MMV aims at exploiting the structure shared by gaze and EEG recordings, sparsely representing them in a single well-chosen dictionary. Our hypothesis is that only the part of the EEG observations related to saccades will be estimated and can then be substracted from EEG channels to recover a clean brain activity. Although a linear superposition of the temporal signals is considered, this method will not suffer from data dimension reduction as is the case for regression or ICA.

This paper is organized as follows: in Section 2, the proposed method and data preprocessings are described. In Section 3, numerical processings on gaze and EEG experimental data are presented and comparison with some other methods are provided. Finally, the conclusion and some perspectives are detailed in Section 4.

2 Proposed Method

In this section, we present the MMV approach and we describe how to preprocess gaze and EEG data for using the proposed method.

2.1 Multiple Measurement Vectors

The purpose of MMV is to obtain a sparse representation of multiple observed signals in a single, well-chosen dictionary, exploiting redundancy in these signals. The considered model is the following

$$\mathbf{Y} = \mathbf{\Phi X} + \mathbf{R}, \tag{1}$$

where $\mathbf{Y} = [\mathbf{y}_1 \dots \mathbf{y}_N]$ is a data matrix containing N signals $\mathbf{y}_n$ $(n \in \{1, \dots, N\})$ stored in N_s-dimensional column vectors (where N_s is the number of samples). $\mathbf{\Phi} \in \mathbb{R}^{N_s \times M}$ is a (finite) dictionary of M atoms chosen to extract the particular structure shared among the $\mathbf{y}_i$, maximaly capturing its redundancy. There is no assumption about orthogonality among the atoms. $\mathbf{X} \in \mathbb{R}^{M \times N}$ is a row sparse code matrix in which the nonzero coefficients model the particular signal shape in the dictionary. $\mathbf{R} \in \mathbb{R}^{N_s \times N}$ is the residual, *i.e.*, all $\mathbf{Y}$ components that do not present the particular shape we are looking for. In this work, the goal is to estimate $\mathbf{X}$ optimizing the following proxy cost function

$$\Psi(\mathbf{X}) = \frac{1}{2} \parallel \mathbf{Y} - \mathbf{\Phi X} \parallel_F^2 + \lambda \parallel \mathbf{X} \parallel_{2,1}, \tag{2}$$

with $\parallel \cdot \parallel_F$ the Frobenius norm and $\parallel \cdot \parallel_{2,1}$ the 2,1–mixed norm [10] defined as

$$\parallel \mathbf{X} \parallel_{2,1} = \sum_{m=1}^{M} \left(\sum_{n=1}^{N} \mid \mathbf{X}_{m,n} \mid^2 \right)^{1/2} . \tag{3}$$

This mixed norm is used to keep or discard entire rows of coefficients from the matrix $\mathbf{X}$ in order to represent each signal from $\mathbf{Y}$ with the same atoms. Thus, we extract a common structure shared among all $\mathbf{y}_n$. Finally, λ (2) is a regularization parameter inducing row sparsity on $\mathbf{X}$. In this paper, λ is arbitrarly fixed, however it is important to notice that this parameter can be chosen by using for example cross-validation.

For this estimation issue, we consider a variable splitting and an augmented Lagrangian as follows

$$\overline{\Psi}(\mathbf{X}, \mathbf{Z}, \mathbf{U}) = \frac{1}{2} \parallel \mathbf{Y} - \mathbf{\Phi X} \parallel_F^2 + \lambda \parallel \mathbf{Z} \parallel_{2,1} + \mathbf{U}^t (\mathbf{X} - \mathbf{Z}) + \frac{\rho}{2} \parallel \mathbf{X} - \mathbf{Z} \parallel_F^2, \tag{4}$$

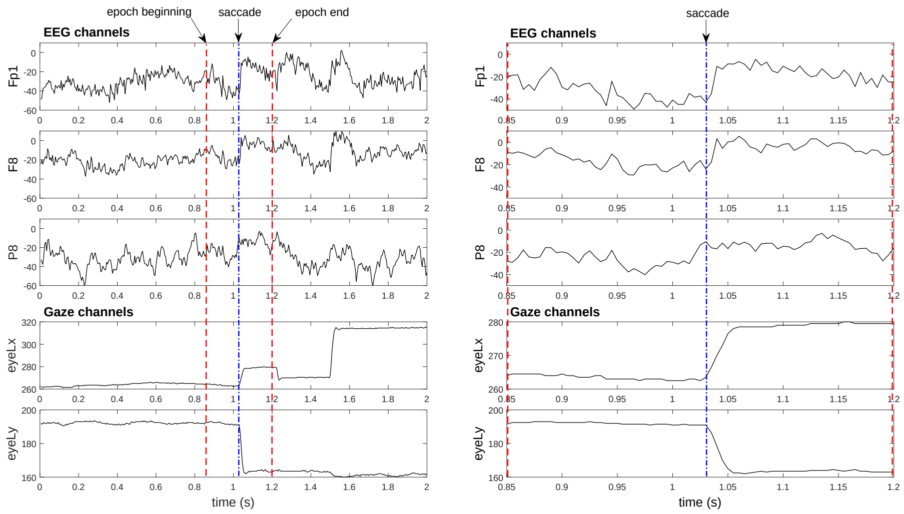

Fig. 1: Selection and epochage of one saccade from gaze and EEG data. On the left side: observed signals (from three EEG channels and two gaze ones) containing at least one interesting saccade to extract. On the right side: the epoched saccade on the same five channels during the previously selected time segment.

where $\mathbf{Z}$ is the split variable, $\mathbf{U}$ is the matrix of Lagrangian multipliers, and ρ is a regularization constant linked to the converge speed [8]. The optimization problem reads

$$\widehat{\mathbf{X}} = \arg\min_{\mathbf{X}} \min_{\mathbf{Z}} \max_{\mathbf{U}} \overline{\Psi}(\mathbf{X}, \mathbf{Z}, \mathbf{U}) \approx \arg\min_{\mathbf{X}} \Psi(\mathbf{X}), \tag{5}$$

and we use the Alternating Direction Method of Multipliers algorithm (ADMM) [8] as a solver.

The main question remains how to build the data matrix $\mathbf{Y}$ from the available observations and the choice of a dictionary $\mathbf{\Phi}$ adapted to the problem.

2.2 Data matrix building

This MMV method can be used to denoise either only one saccade or several saccades (respectively, single trial and multitrial processing). The first preprocessing step is to epoch the recordings in order to keep only the interesting parts of the signals (see Fig. 1). To do so, we localize the saccades on gaze channels (at the dash-dot line in Fig. 1). Then, we extract a predefined time window containing only one saccade (in dash lines in Fig. 1). These constitute the trials. Each trial is made of N_s samples and contains a first fixation, then the saccade and finally a second fixation.

We consider K trials. For each trial, we have P recordings from the eye-tracker (the number of gaze channels) and Q recordings from the EEG sensors (the number of electrodes). For the kth trial, $k \in \{1, \ldots, K\}$, we can build, respectively (resp.), a gaze matrix $\mathbf{G}^{(k)} \in \mathbb{R}^{N_s \times P}$ and an EEG matrix $\mathbf{E}^{(k)} \in \mathbb{R}^{N_s \times Q}$ defined for all $n \in \{1, \ldots, N_s\}$ by

$$\mathbf{G}^{(k)} = [\mathbf{g}_1^{(k)}(n), \ldots, \mathbf{g}_P^{(k)}(n)] \quad \text{and} \quad \mathbf{E}^{(k)} = [\mathbf{e}_1^{(k)}(n), \ldots, \mathbf{e}_Q^{(k)}(n)], \tag{6}$$

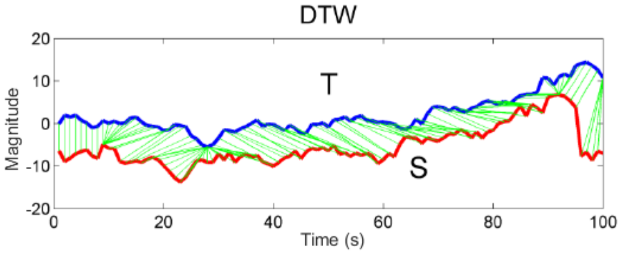

Fig. 2: Dynamic Time Warping method associating to each sample of signals S and T the point of, resp., T and S that presents the smallest distance (green lines are mapping between points of time series T and S). (provisional figure extracted from [13]).

where $\mathbf{g}_p^{(k)}(n) \in \mathbb{R}^{N_s \times 1}$, $p \in \{1, \dots, P\}$ and $\mathbf{e}_q^{(k)}(n) \in \mathbb{R}^{N_s \times 1}$, $q \in \{1, \dots, Q\}$ are column vectors representing the signals from, resp., the pth gaze channel and the qth EEG channel. An optional preprocessing is the downsampling. Indeed, the number of samples directly impacts the computational time. Thus, if the epoched signals are made of too many samples, then one can downsample them respecting the Nyquist-Shannon sampling theorem. As, in the same trial, the gaze signals and the ocular artifacts contained in EEG signals share the same structure, one may directly use the MMV on data matrix $\mathbf{Y} = [\mathbf{G}^{(k)}, \mathbf{E}^{(k)}]$.

The case of multitrial processing raises a main issue. Indeed, in each trial, data present a common temporal structure linked to the ocular artifacts that is important for the proposed method efficiency. Among trials, one can find similar structures (or shape) but this time with some temporal distortions due to the difference among saccades or among subjets moving their eyes. These distorsion may impact the MMV performance in the considered application. In order to fix this issue, we propose to align the different trials using an extension of the Dynamic Time Warping (DTW) [11] called the Generalized Time Warping (GTW) [12]. DTW is an algorithm for measuring similarity between two time series. This method can be used to compute an optimal match between two temporal signals. As it is shown in Fig. 2, DTW calculates nonlinear functions for each time serie, such that the sum of the distances between their points is smallest and so the correlation between both signals is maximum. The considered distance depends on which algorithm is used. GTW generalizes DTW method for more than two sets of time series. Whatever the saccade orientation, GTW aims at matching the shape of signals from different trials. For that, GTW computes, for each set, a nonlinear bijective function that warps time and allows to minimize the shape difference among the set of time series. Due to their very similar shape (see Fig. 3), gaze signals seems simpler to align. Thus, we directly apply the GTW on all matrices $\mathbf{G}^{(k)}, k \in \{1, \dots, K\}$. Once the nonlinear functions are computed, they are applied to gaze and EEG matrices of the corresponding trial.

After this shape matching step, we can build two new matrices. They contain the concatenation of each channel of each trial, resp., for the gaze and for the EEG observations. They are defined as follows

$$\mathbf{G} = [\mathsf{GTW}(\mathbf{G}^{(1)}), \dots, \mathsf{GTW}(\mathbf{G}^{(K)})] \quad \text{and} \quad \mathbf{E} = [\mathsf{GTW}(\mathbf{E}^{(1)}), \dots, \mathsf{GTW}(\mathbf{E}^{(K)})], \tag{7}$$

where $\mathsf{GTW}(\cdot)$ is the generalized time warping operator, $\mathbf{G}$ and $\mathbf{E}$ are, resp., made of KP and KQ channels. These dimensions may be very large and have a major impact on the computational complexity. We propose to reduce the size of these matrices. As we have induced a similar shape among the gaze trials using the GTW, we can expect that the contribution of the ocular artifact components is also similar in EEG observations for all the trials. Hence, we propose to do a Principal Component Analysis (PCA) on $\mathbf{G}$ and an other one on $\mathbf{E}$. For the gaze matrix $\mathbf{G}$, all the channels share a smooth step shape and thus the first principal component should explain almost entirely the signal power. We propose to threshold the principal components keeping the most powerful ones and dropping the others when we have enough to explain 99% of the original signal power. For the EEG, we only want to extract the saccade components. As saccades induce an electrical potential much larger in magnitude than the brain activity, we propose to keep the first principal components explaining 95% of the entire power. Both these empirical thresholds highly reduce the size of $\mathbf{G}$ and $\mathbf{E}$. Finally, we use MMV on the new data matrix $\mathbf{Y} = [\mathsf{PC}_{99\%}(\mathbf{G}), \mathsf{PC}_{95\%}(\mathbf{E})]$, where $\mathsf{PC}_{99\%}(G)$ and $\mathsf{PC}_{95\%}(E)$ are the operators extracting the principal components.

It remains to explain how to select the dictionary $\mathbf{\Phi}$.

2.3 Dictionary selection

Since we aim at decomposing only the ocular artifact components, we consider a dictionary containing atoms that match with the gaze signals. As these signals look like smooth steps, our choice is to use the following sigmoidal function

$$f_{\alpha,\beta}(t) = \frac{1}{1 + e^{-\alpha(t-\beta)}}, \tag{8}$$

where α and β are, resp., the scale and the translation parameters. In order to take into account the gaze signals overshoots and the side effects due to the signals finite support, we include the derivative of (8) in the dictionary

$$g_{\alpha,\beta}(t) = \frac{\partial f_{\alpha,\beta}(t)}{\partial t} = \alpha f_{\alpha,\beta}(t) f_{-\alpha,\beta}(t). \tag{9}$$

We also add the constant function which acts as an offset and we normalize all the atoms. Finally, all the atoms are seen as column vectors and we concatenate them in the dictionary $\mathbf{\Phi}$ for all considered scales α and translations β.

Hereafter, we summarize the outline of the proposed novel MMV method for gaze and EEG multimodal approach (called MMV-G&E).

1. Preprocessing from gaze and EEG observations
   - Epoch $\rightarrow$ build $\mathbf{G}^{(k)}$ and $\mathbf{E}^{(k)}, k \in \{1, \ldots, K\}$ (6)
   - Downsample (optional)
   - If $K > 1$: perform GTW $\rightarrow$ build $\mathbf{G}$ and $\mathbf{E}$ (7)
   - Build $\mathbf{Y} = [\mathsf{PC}_{99\%}(\mathbf{G}), \mathsf{PC}_{95\%}(\mathbf{E})]$
2. MMV method: optimize the cost function $\Psi(\mathbf{X})$ (4)
   - Choose the dictionary $\mathbf{\Phi}$ with respect to the data
   - Fix the regularization parameters λ and ρ
   - Solve (5) for $\widehat{\mathbf{X}}$, e.g., using ADMM

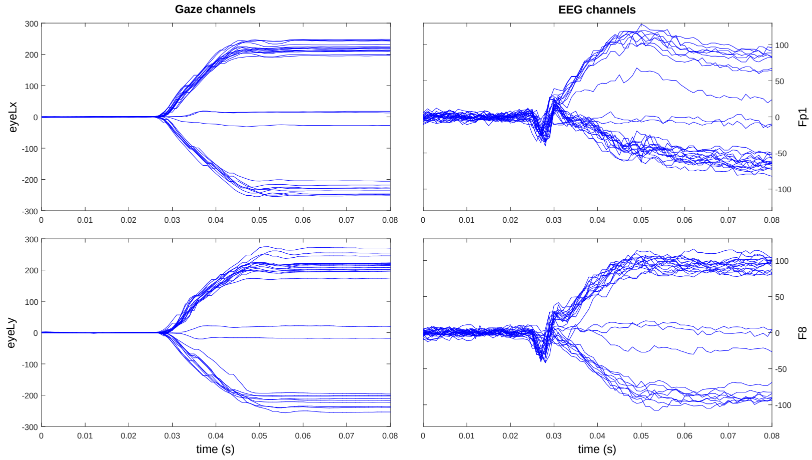

Fig. 3: All trials for some gaze and EEG channels after downsampling and preprocessing by GTW method.

3 Experiments

In this section, we assess the MMV-G&E performance on gaze and EEG real-data. These come from an experiment in visual exploration where participants had to search a target from a set of distractors [14]. Sixty four active electrodes (BrainProductsGmbH) were mounted on an EEG cap (BrainCap^{TM}) placed on the scalp in compliance with the international 10-20 system. To be compatible with the EEG acquisition, eye-movements were recorded by a remote binocular infrared eye-tracker EyeLink 1000 (SR Research) to track the gaze direction of the left eye while the observer was looking at the stimuli. The EyeLink system was used in the Pupil-Corneal Reflection tracking mode. For both acquisition devices, the sampling frequency was 1000Hz. Off-line, EEG signals and gaze samples were synchronized using hardware triggers signals sent in parallel to the EEG recorder and the eye-tracker, along the experiment. Let note that the EEG electrode F3 was defective during the experiment and has been removed from the data ($Q = 63$). Concerning the gaze information, we take into account the vertical and the horizontal channels ($P = 2$). We consider $K = 26$ epoched trials. Each signal, downsampled at 333Hz, is composed of $N_s = 75$ samples and lasts about 225 ms. After the GTW preprocessing, we obtain the data displayed in Fig. 3 for both gaze channels and two EEG channels. In Section 3.1, we describe the selected parameters for using the proposed method and we show some qualitative results obtained on real-data. Finally, a validation method is proposed to assess the performance of MMV-G&E and comparisons with standard methods from the literature are provided in Section 3.2.

3.1 MMV-G&E parameters and qualitative results

For this experiment, the MMV-G&E regularization parameters have been heuristically fixed: $\lambda = 42$ and $\rho = 1$. Future work will consist in optimizing λ. Concerning the dictionary, the atoms defined, for $t = -10, \ldots, 10$, with a step of

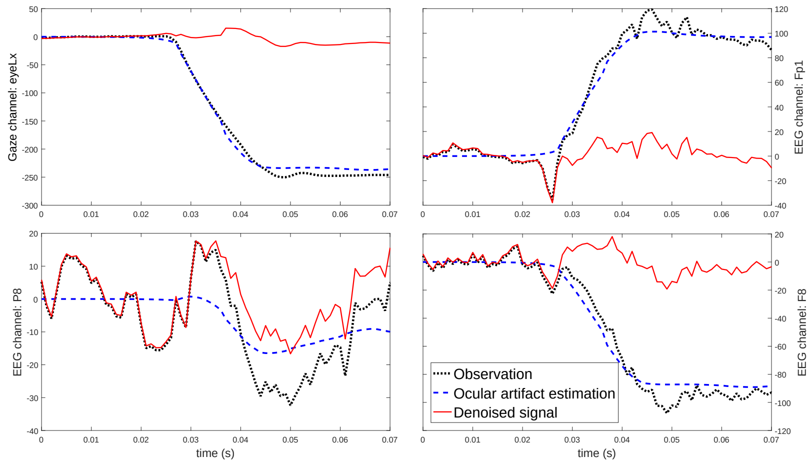

Fig. 4: MMV-G&E denoising effect for one trial and some gaze and EEG channels.

$20/(N_s - 1)$, by (8) and (9) are concatenated with the constant function as explained in Section 2.3. The scale and translation parameters are empirically chosen: $\alpha \in \{1, \dots, 10\}$ and $\beta = -10, \dots, 10$, with a step of $20/(N_s - 1)$. Fig. 4 shows the denoising by the proposed method on the considered experimental data. After removing the saccade contribution estimates (in dashed lines) from the observations (in dotted curves), we obtain the denoised signals (in solid lines) which seem to conserve the pre-saccadic behavior that corresponds to pure brain activity. We can observe that, as expected, the saccade contribution estimates depend on the considered electrodes. Thus, MMV-G&E method derives high magnitude saccades for Fp1 and F8 and very low magnitude ones for P8.

3.2 Comparisons and validation

Here, we compare MMV-G&E to some algorithms from the state of the art:

- the regression method [3, 4] with the gaze taken as reference,
- Infomax algorithm (ICA) [1] applied to a matrix in which gaze and EEG channels are concatenated for each trial and then all trials are stacked,
- CCA [15] that finds projections on a common space, maximizing the correlation between gaze and EEG,
- the coupled tensor factorization method RACMTF [6].

In order to assess the efficiency of these methods, we propose the following validation. From each EEG trial $\mathbf{E}^{(k)}$, we extract three windows of 20 samples representing, resp., the pre-saccadic fixation, the saccade and the post-saccadic fixation, stored in three matrices, resp., $\mathbf{E}_{pr}^{(k)}$, $\mathbf{E}_{sa}^{(k)}$ and $\mathbf{E}_{po}^{(k)}$ of size 20×63. Each extracted signal is centered. Then, we stack the trials such that $\mathbf{E}_{pr} = [\mathbf{E}_{pr}^{(1)\,T}, \dots, \mathbf{E}_{pr}^{(K)\,T}]^T$ where $(\cdot)^T$ is the transpose operator. We do the same for $\mathbf{E}_{sa}$ and $\mathbf{E}_{po}$. Finally, we compute two vectors of generalized eigenvalues (GEV):

$$\mathbf{d_1} = \mathsf{GEV}(\mathsf{Cov}(\mathbf{E}_{pr}), \mathsf{Cov}(\mathbf{E}_{po})) \text{ and } \mathbf{d_2} = \mathsf{GEV}(\mathsf{Cov}(\mathbf{E}_{sa}), \mathsf{Cov}(\mathbf{E}_{po})), \quad (10)$$

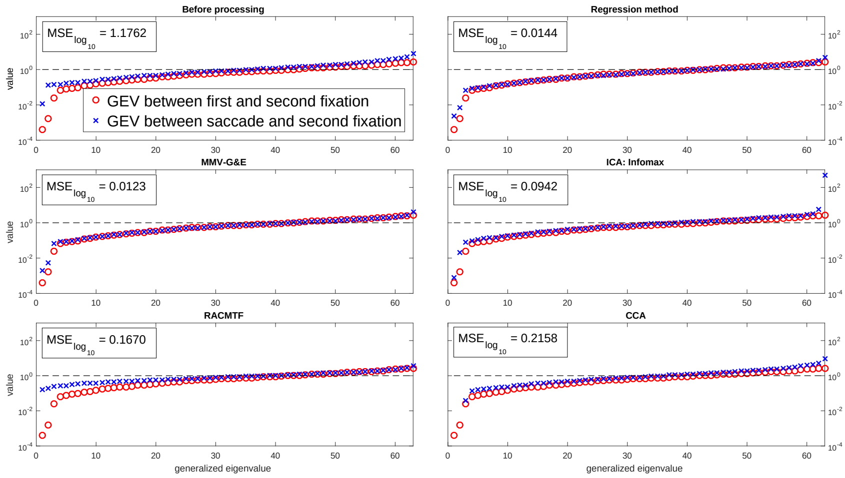

Fig. 5: Validation comparing the GEVD between the covariance matrices of pre and post-saccadic fixations and GEVD between the covariance matrices of pre-saccadic fixation and denoised saccade. MMV-G&E is confronted with four popular method.

where $\mathsf{Cov}(\cdot)$ is the covariance operator. The vectors $\mathbf{d_1}$ and $\mathbf{d_2}$ are displayed in Fig. 5. The brain activity can be assumed to be stationary for long time segments. Here, as the trials are stacked in the matrices $\mathbf{E}_{pr}$, $\mathbf{E}_{sa}$ and $\mathbf{E}_{po}$, we can expect that each $\mathbf{d_1}$ entry should tend to 1. Due to the saccade power, we have $\mathbf{d_2}_{,i} \geq \mathbf{d_1}_{,i}$ $(i \in \{1, \ldots, 63\})$. This is confirmed before denoising (see Fig. 5). After this processing, we expect to reduce the distance bewteen each pair of generalized eigenvalues (ideally $\mathbf{d_2}_{,i} = \mathbf{d_1}_{,i}$). For indicative information, a measurement between $\mathbf{d}_2$ and $\mathbf{d}_1$ is provided, in Fig. 5, using the mean square error in logarithmic scale ($\text{MSE}_{\log_{10}}$). In this figure, we can observe that the proposed method obtains slightly better results than regression one and outperforms the three other algorithms on this example.

4 Conclusions and Perspectives

In this paper, we propose a multimodal approach to tackle the eye-movement artifact removal in EEG recordings. The gaze signals, used as a reference, share a similar shape with the ocular artifacts. The considered MMV method allows to exploit this property decomposing the data in a row sparse way in a same well-chosen dictionary. Only the structure shared by gaze and EEG recordings is estimated and is used to extract the ocular artifacts from the EEG data. One may notice that the use of MMV-G&E for single trial processing is straightforward, yet, it is more complicated for multitrial processing. Indeed the signals between different trials have to share the sought similar temporal structure. In order to enforce this constraint, we propose to use the GTW method that warps time in order to align the signals. The experiments on gaze and EEG real-data have shown the proposed method efficiency for the ocular artifact extraction. Moreover MMV-G&E compares favorably to classical methods from the literature. Future work will consider other extensions as a clever choice for the MMV thresholding parameter or some additional constraints for the sparse representation linked to the temporal structure of gaze and EEG data. It will also consist in testing MMV-G&E performance on various criteria.

References

1. Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., Mckeown, M.-J., Iragui, V., Sejnowski, T.: Removing electroencephalographic artifacts by blind source separation. In: Psychophysiology, vol. 37, no. 2, pp. 163–178 (2000)
2. Iwasaki, M., Kellinghaus, C., Alexopoulos, A.-V., Burgess, R.-C., Kumar, A.-N., Han, Y.-H., Lüders, H.-O., Leigh, R.-J.: Effects of eyelid closure, blinks, and eye movements on the electroencephalogram. In: Clinical Neurophysiology, vol. 116, no. 4, pp. 878–885 (2005)
3. Semlitsch, H.-V., Anderer, P., Schuster, P., Presslich, O.: A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. In: Psychophysiology, vol. 23, no. 6, pp. 695–703 (1986)
4. Schlögl, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., Pfurtscheller, G.: A fully automated correction method of EOG artifacts in EEG recordings. In: Clinical neurophysiology, vol. 118, no. 1, pp. 98–104 (2007)
5. Comon, P., Jutten, C.: Handbook of Blind Source Separation: Independent component analysis and applications. Academic press (2000)
6. Rivet, B., Duda, M., Guérin-Dugué, A., Jutten, C., Comon, P.: Multimodal approach to estimate the ocular movements during eeg recordings: a coupled tensor factorization method. In Proc.: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 6983–6986 (2015)
7. Cotter, S.-F., Rao, B.-D., Engan, K., Kreutz-Delgado, K.: Sparse solutions to linear inverse problems with multiple measurement vectors. In: IEEE Transactions on Signal Processing, vol. 53, no. 7, pp. 2477–2488 (2006)
8. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. In: Foundations and Trends® in Machine Learning, vol. 3, no. 1, pp. 1–122 (2011)
9. Eldar, Y.-C., Rauhut, H.: Average case analysis of multichannel sparse recovery using convex relaxation. In: IEEE Transactions on Information Theory, vol. 51, no. 1, pp. 505–519 (2010)
10. Kowalski, M.: Sparse regression using mixed norms. In: Applied and Computational Harmonic Analysis, vol. 27, no. 3, pp. 303–324 (2009)
11. Berndt, D.-J., Clifford, J.: Using Dynamic Time Warping to Find Patterns in Time Series. In Proc.: KDD workshop, 2012 IEEE Conference on, vol. 10, no. 16, pp. 359–370 (1994)
12. Zhou, F., De la Torre, F.: Generalized time warping for multi-modal alignment of human motion. In Proc.: Computer Vision and Pattern Recognition (CVPR), IEEE Conference on, pp. 1282–1289 (2012)
13. Cassisi, C., Montalto, P., Aliotta, M., Cannata, A., Pulvirenti, A.: Similarity measures and dimensionality reduction techniques for time series data mining. In: Advances in data mining knowledge discovery and applications (2012)
14. Kristensen, E., Guerin-Dugué, A., Rivet, B.: Comparison between Adjar and Xdawn algorithms to estimate eye-fixation related potentials distorted by overlapping. In Proc.: 7th International IEEE/EMBS Conference on Neural Engineering (NER), pp. 976–979 (2015)
15. Hardoon, D., Szedmak, S., Shawe-Taylor, J.: Canonical correlation analysis: An overview with application to learning methods. In: Neural computation, vol. 16, no. 12, pp. 2639–2664 (2004)