

PSVD-XCCA TO SUPPRESS OCULAR ARTEFACTS IN EEG

R. Phlypo¹, Y. D'Asseler¹, I. Lemahieu¹, R. Van de Walle²

¹Ghent University, MEDISIP, Department of electronics and information systems, Belgium

²Ghent University, MMLAB, Department of electronics and information systems, Belgium

Abstract

In this paper we present a technique to suppress ocular artefacts in the EEG (electroencephalogram) based on a combination of local and global correction algorithms. The results are promising when considering distortion, interference and artefact suppression ratios.

1 Introduction

The ocular artefact is a well known source of interference in the EEG. In the past, many techniques have already been presented to cope with the suppression of these artefacts. Some of them rely on the prior knowledge available from the oculogram leads [9]. However the leads are not always available in the measurements set-up and the oculogram is equally contaminated with cerebral activity. Simple regression would thus lead to subtraction of a combination of ocular and cerebral activity. Another viewpoint is given in the techniques based on Autoregressive Modelling, such as the Autoregressive Moving Average (ARMA) and its extensions. Here the method tries to find an optimal filter for the EOG by adapting the filter taps before subtracting the filtered signal from the contaminated EEG. The main drawbacks here are the spectral overlap between cerebral and ocular activity and the choice of filter length [6]. The use of neural networks [13] has also been a subject of research in this domain, next to many other techniques.

It is now approximately fifteen years since the Blind Source Separation (BSS) techniques, and more specifically Independent Component Analysis (ICA), made their entrance in the processing of EEG signals [11]. Since ocular activity is probably the most interfering and most well known source of artefacts in EEG reading, it may not come as a surprise that this was one of the first application dealt with in the context of BSS. However, the recently proposed methods based on higher-order statistics only do well when sufficient data is available. According to [3], $k \cdot m^2$ datapoints are needed, where m is the number of channels and k is a constant. For a 27 channel EEG and a typical value for $k = 30$, this results in 21870 samples.. At a sample rate of 200Hz this is approximately 109s. In general however, we find that $k = 3$ is already sufficient. Due to the large sample size needed, they can be seen as global solutions.

The Singular Value Decomposition (SVD), however, relies only on second order statistics, and needs considerably less data. The drawback of the method is the forced orthogonality of the scalp maps, resulting in unrealistic cerebral activity mapping. Nevertheless, we showed in [12] that by using only the major eigenvector in a sliding window and reprojecting it from the measurements when a matching criteria is fulfilled, the reconstructed EEG is an EEG that is almost free from ocular artefact activities such as blinks and saccades.

In this work we consider the drawbacks and advantages of using piecewise SVD (pSVD) and propose a method to incorporate the complementary advantages of ICA. The solutions to both algorithms are subjected to a cross Canonical Correlation Analysis (xCCA). In the methods section the methods of ICA, pSVD and xCCA are described briefly, next to a simulation set-up which has been used to evaluate the results. For the interested reader we point to the vast literature available on the standard methods in BSS, e.g. [7][10]. The overall algorithm will be sketched and briefly touched upon. In section 3 exemplary figures and results are shown and a statistical comparison is made to show the improvement obtained by combining the methods. The final section concludes with the advantages of this work and tries to reveal the weak points that are to be considered in the future.

2 Methods

2.1 BSS

BSS techniques try to estimate the sources and the mixing matrix, given only the measurements. Since the head can be seen as a volume conductor and the frequencies of interest have a small spectral range, the voltage distributions at the scalp are an instantaneous reflection of the electrical activities of the sources (whether cerebral or extracerebral) and we can reduce the Maxwell equations to their first order (linear) approximation. A simple model would thus be:

$$y(t) = As(t) \quad (1)$$

where $y(t)$ is a measurement of m channels of EEG at a time instance t . The corresponding sources $s(t)$ are linearly mixed using A to obtain the measurements. By imposing some constraints or

characterizations on the sources and/or the mixing matrix, we can estimate the sources by solving

$$s'(t) = Wy(t) \quad (2)$$

to $s'(t)$ and W . It can be shown that when imposing the constraint of statistical independence on the source estimates $y(t)$ and constraining the rows of W to be of unit norm, we can solve equation (2) toward W , which will approximate the inverse of A up to a permutation and a scaling [2]. Many attempts have been made to write fast, robust implementations of the independence criterion and every algorithm so far can be seen as a specific case of the general statistical independence criterion [10]. Henceforth we denote by ICA the use of the fastICA algorithm as implemented by Hyvärinen [7].

2.2 pSVD

The SVD, a general form of the Principal Component Algorithm (PCA), imposes a weaker constraint on the source estimates, namely uncorrelatedness. From a statistical point of view, independence necessarily comprises uncorrelatedness, but uncorrelatedness does not mean the data is independent. To solve for equation (2) only using the uncorrelatedness constraint on the sources, W could only be found up to a rotation. Therefore the rows in W are assumed to be orthogonal as well. The latter restriction has been seen by many researchers as violating the spatial distribution of natural cerebral activity. By using a sliding window and only using the first component we try to overcome this shortcoming of the SVD.

The SVD problem can be stated as the eigen solution of the covariance matrices of Y , where $Y = [y(1) y(2) \dots y(N)]$. The covariance matrices YY^T and Y^TY share common eigenvalues and the corresponding eigenvectors are given in U , respectively V , yielding

$$Y = U\Sigma V^T \quad (3)$$

where the diagonal of Σ contains the singular values or the square roots of the eigenvalues. An advantage of this method is the deliberate ordering of the components, where the component with maximal variance is the eigenvector associated to the largest eigenvalue (which actually is the variance of the source). Since ocular activity almost always exhibits a variance that is five to ten times the variance of the background cerebral activity, we only need to estimate this first component. For consistency in naming we can regroup ΣV^T and denote it as S' , and rename U as A . We then get $Y=AS$, which is the matrix notation of equation (2).

In the setting of pSVD the components are calculated for small windows, to cope with small changes in the the ocular electrical field. These perturbations of the electrical field influence the spatial map of the ocular source over longer time spans and the sliding window approach gives the opportunity to cope with these changes. Since not every window is affected by ocular activity, there is no need to deflate the signal subspace by the first component in each window. This problem is solved by creating an exemplar database, based on the characteristic spatial maps of ocular activity as given in [12][14]. The spatial map associated to the major eigenvalue for each window is then compared to the database, and when a distance criterion to the database is fulfilled, the associated component is deflated from the measurement subspace.

Since the sliding windows make use of overlap, the reconstruction of Y must be done using rescaled versions of the components. The algorithms can thus be seen as an averaging over the major eigenvectors which fulfill the distance criterion.

2.3 xCCA

The xCCA proposed here is nothing more than the intersection of the subspaces retrieved by ICA and the pSVD. If we constrain the ICA components to the set which is fulfilling the distance criterion to the dataset as we proposed in the pSVD, then we can compare the solutions of the pSVD and the ICA by estimating the joint subspace. The algorithm to be used is the same as is given in [5] to compute the angle between subspaces where the component associated with the smallest angle is projected back onto the original subspaces.

Let Y'_{ICA} represent the sources retrieved from an appropriate ICA, selected upon resemblance to the database set, back projected onto the original data space. Let Y'_{pSVD} be analogously defined for the pSVD method. The algorithm then calculates the first component in the joint subspace based upon

$$Q_{ICA}Q_{pSVD}^T = U_2\Sigma_2V_2^T \quad (4)$$

where the Q 's are the normalized matrices of the QR factorization of Y'_{ICA} , respectively Y'_{pSVD} . The solution to the problem is then given by

$$S'_{eye} = U_2(:,1)^T Q_{ICA}^T = QV_2(1,:) \quad (5)$$

where the Matlab® notation is used for indexing. S'_{eye} is then the major direction of intersection between both subspaces and will add the benefits of both methods while trying to overcome their weak points as will be shown in the results.

2.4 Simulated EEG

The underlying sources of patient EEG are not known and by consequence, objective validation of BSS algorithms on realistic EEG is the Achilles' heel to many researchers in the field. We propose to simulate EEG waveforms based on patient EEG characteristics to obtain the background EEG, while simulating the ocular activity by a dipole. The latter is a justified approximation, supported by results in [1][8].

The characteristics derived from the patient EEG are the probabilities of having a certain power for an associated frequency. The spectra are obtained by a sliding window of 2s over a 30s EEG sample and stored without their temporal association. We calculate the cumulative density functions for each of the frequencies so that they can be resampled at a later stage. This approach saves us the effort of calculating or fitting a distribution. Moreover, it has a discretisation given by the data, so no bins are to be defined.

The subsequent step is to define hundred dipoles at random places in the cortex and assign them random temporal activations. The orientation of each of the dipoles is taken radial to the cortex layer since the groups of pyramidal cells in the cortex are known to globally exhibit analogous electrical properties. The obtained activity is then filtered with frequency characteristics per channel derived from uniform sampling of the cumulative density function obtained as described above. The background EEG created in this stage is denoted as Y_{bg} .

The ocular dipole is then added to this Y_{bg} . Since the morphology is not uniform for all the artefact modes or artefact instances, we create two classes of artefacts - blinks and saccades - and allow for variance in their amplitude over time. The aforementioned ocular movements are restricted to rotations in the horizontal or vertical plane for saccades and in the vertical plane only for blinks. It has been shown in [8] that an eye blink consists of an upward turning of the eyeball in combination with an electrical shortcut caused by closing the eyelid(s). This can be modelled perfectly with a rotational dipole and a short time discharge and recharge. The lack of a good model for saccades forces us to test the techniques at first with blinks only.

2.5 SDR, SIR and SAR

To compare the different methods used, we calculate three measures based upon the interference noise, the artefact noise and the target source. Both noise factors are derived from the estimated source and the results give a good

overview of the distortion, given as the ratio of the source to the total of interference and methodological artefact errors (SDR: source-to-distortion ratio), the interference (SIR: source-to-interference ratio) and the methodologically introduced artefact (SAR: source-to-artefacts ratio). The measures are used as defined in the BSS_EVAL toolbox 2.0 [4] for global comparison.

3 Results

The set-up is done using the simulated EEG, where the source is the interfering ocular activity, the remaining sources are covered by the background EEG. We ran 250 simulations to extract some statistically valuable results of the measures SDR, SIR and SAR in the cases of simple pSVD, fastICA and for the combined version by means of xCCA. The results are given in table 1 as the means with their respective standard deviations.

Table 1: mean and standard deviations of SDR, SIR and SAR for the three methods; maxima are typeset in bold

| | pSVD | fastICA | xCCA |
|----------|---------------------|---------------------|---------------------|
| SDR (dB) | 12.99 ± 3.18 | 12.00 ± 3.68 | 14.36 ± 2.44 |
| SIR (dB) | 23.30 ± 4.46 | 14.86 ± 5.07 | 21.19 ± 4.92 |
| SAR (dB) | 13.52 ± 3.02 | 16.87 ± 1.71 | 15.96 ± 2.02 |

In Figure 1 we show the results of the method on a real EEG signal. Figure 2 shows a close-up of the signal to indicate the artefact introduced by the pSVD method and how it is dealt with by the xCCA method in real EEG.

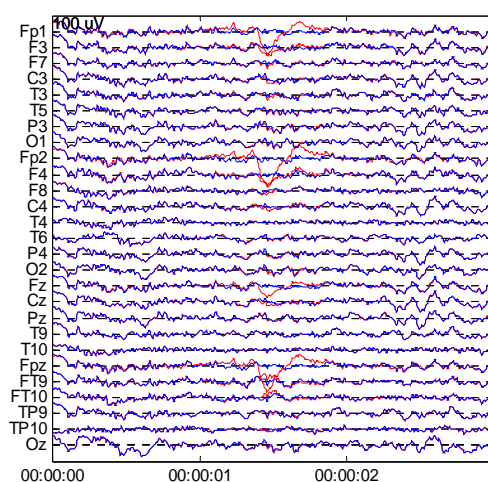


Figure 1: Example of performance on a real 5s EEG fragment (blue: corrected, red: original EEG)

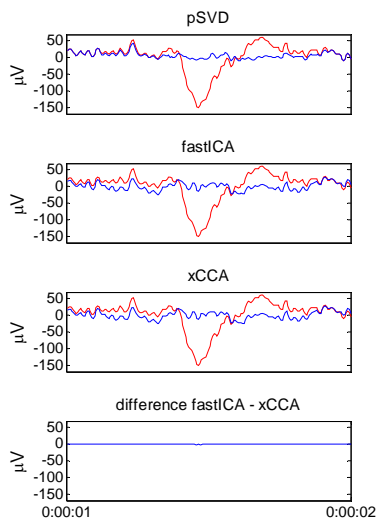


Figure 2: Comparison between the techniques used during the blink (window taken from Figure 1, channel Fp1; blue: corrected, red: original EEG).

4 Discussion

It is clear from the results that ICA and pSVD have both their respective advantages and disadvantages. While ICA is more preferable to reduce the methodological artefacts, pSVD is better in suppression of the interference of other sources. The combination of both methods clearly reduces the nominal values of SIR and SAR for pSVD, respectively ICA, but is still advantageous in its use, looking at the overall dB levels in methodological artefact suppression.

We have made the comparison with other methods, such as CCA with single or multiple delays, with or without a reference (results not shown), full window SVD and SVD with nonlinearities. In all cases the performance of ICA and pSVD for SAR, resp. SIR were considerably higher. The combination of both methods in xCCA still had a higher performance for the ratio parameters in the majority of the cases. In general both the artefact and the interference suppression of the xCCA method were superior to the lowest of both in the other methods.

5 Conclusion

The proposed method, xCCA, sheds new light on the combination of global and local parameters to adjust a raw EEG signal for the ocular artefact. Although the model used to simulate EEGs is not realistic, the results are an indication of the behaviour of the methods on EEG like data. The simulations of the ocular movements apart from the eye blinks (e.g. saccades) have to be refined in order

to show satisfying simulation studies on these data as well.

References

- [1] Berg, P et al. Dipole models of eye movements and blinks. *Electroencephalography and Clinical Neurophysiology* 79(1). p36-44 1991.
- [2] Comon, P. Independent Component Analysis, a new concept? *Signal Processing*, 36(3) p287-314, 1994.
- [3] Delorme A et al. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience methods* 134, p9-21, 2004.
- [4] Févotte, C et al. BSS_EVAL Toolbox user guide, *IRISA Technical Report 1706*, 2005. http://www.irisa.fr/metiss/bss_eval/
- [5] Golub, G et al. Matrix computations, 3rd ed. The John Hopkins University Press, pp694, 1996.
- [6] Haas, S et al. EEG ocular artefact removal through armax model system identification using extended least squares. *Communications in information and systems* 3(1), p19-40, 2003.
- [7] Hyvärinen, A et al. Independent Component Analysis, John Wiley and sons, pp504, 2001.
- [8] Iwasaki, M et al. Effects of eyelid closure, blinks and eye movements on the electroencephalogram. *Clinical Neurophysiology* 116(4). p878-885, 2005.
- [9] Jervis, BW et al. Effect on EEG responses of removing ocular artefacts by proportional EOG subtraction. *Medical and Biological Engineering and Computing* 27(5), p484-490, 2006.
- [10] Lee, T-W et al. A unifying information-theoretic framework for independent component analysis. *Computers and mathematics with applications* 39 p1-21, 2000.
- [11] Makeig, S et al. Independent component analysis of electroencephalographic data, *Advances in Neural Information Processing Systems* 8, p145-151, 1996.
- [12] Phlypo R et al. pSVD: a method for robust, real time eye movement artefact rejection from the EEG. *Proceedings of the MEDSIP2006*, Glasgow, UK, 2006.
- [13] Selvan, S et al. Removal of ocular artefacts from EEG using an efficient neural network based adaptive filtering technique. *IEEE Signal Processing Letters* 6(12), 1999.
- [14] Stern, J et al. Atlas of EEG patterns, Lippincott, Williams & Wilkins, pp320, 2004.