Subspace Techniques to Remove Artifacts From EEG: a quantitative analysis

Article *in* Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference · February 2008

DOI: 10.1109/IEMBS.2008.4650185 · Source: PubMed

4 authors:

Ana Rita Teixeira
University of Aveiro
51 PUBLICATIONS **328** CITATIONS
SEE PROFILE

Ana Maria Tomé
University of Aveiro
211 PUBLICATIONS **1,413** CITATIONS
SEE PROFILE

Elmar Lang
Universität Regensburg
86 PUBLICATIONS **1,488** CITATIONS
SEE PROFILE

Armando Silva
University of Porto
130 PUBLICATIONS **826** CITATIONS
SEE PROFILE

Subspace Techniques to Remove Artifacts From EEG: a quantitative analysis

A. R. Teixeira, A. M. Tomé, E. W. Lang and A. Martins da Silva

***Abstract*— In this work we discuss and apply projective subspace techniques to both multichannel as well as single channel recordings. The single-channel approach is based on singular spectrum analysis(SSA) and the multichannel approach uses the extended infomax algorithm which is implemented in the open-source toolbox EEGLAB. Both approaches will be evaluated using artificial mixtures of a set of selected EEG signals. The latter were selected visually to contain as the dominant activity one of the characteristic bands of an electroencephalogram (EEG). The evaluation is performed both in the time and frequency domain by using correlation coefficients and coherence function, respectively.**

I. INTRODUCTION

Electrical scalp recordings do not only contain cerebral activity (EEG) but they are often contaminated by artifacts of physiological or non-physiological origin. Eye movements and blinking (EOG) are of larger amplitude than cortical signals (EEG). As they propagate over the scalp, they are recorded in most EEG derivations. These artifacts complicate the EEG interpretation as, for instance, a seizure onset will be difficult to detect in epileptic data analysis. Consequently, ocular artifacts pose a significant problem to the clinicians and neurologists either by the loss of data or by masking significant events in the data.

The availability of digitalized EEGs makes possible the application of more sophisticated techniques than simple linear filtering. The primary goal of these correction methods is to remove artifacts without distorting the underlying brain signals of interest. A variety of automatic procedures have been proposed to correct or remove ocular artifacts from EEG recordings. The ocular artifacts usually spread over the scalp but are particularly prominent on frontal leads. Therefore most of the techniques correct simultaneously the multichannel recordings [5], [9]. However there exist techniques which, when applied to single-channel recordings, need the EOG signal as reference [2], [4]. More recently in [8] a method was described to eliminate from single-channel EEG recordings ocular artifacts as well as other high-amplitude and low frequency artifacts like head movements and baseline drifts. In spite of the variety of methods already existing, studies are still going on [10], [6], [7], [4] where the main goal is to evaluate quantitatively their performance.

A.M. Tomé, A.R. Teixeira are with DETI/IEETA-Universidade Aveiro, 3810-193 Aveiro Portugal ana@ieeta.pt

E.W. Lang is with Institute of Biophysics, University of Regensburg, D-93040 Regensburg, Germany elmar.lang@biologie.uni-regensburg.de

A. Martins da Silva is with HGSA and ICBAS/IBMC, University of Porto 4099-001 Porto, Portugal ams@icbas.up.pt

The quantification of possible distortions resulting from the application of these methods has been addressed differently. But most use artificial mixtures and propose measures both in frequency and in time to compare the outcome of the algorithms with the original signals. Anyway, no quantitative comparison on a representative and common dataset is available. In this work we discuss projective subspace techniques both for multichannel and single channel recordings. We study the performance of these approaches based on the artificial mixtures of a set of selected EEG signals with a frequency content that covers all the significant bands of the EEG.

II. PROJECTIVE TECHNIQUES

Projective techniques are used to generate an alternative representation of the data that can be more easily interpreted. The projective models are described by a matrix (or a couple of matrices) and generally comprise three steps: the projection of the data, the selection of the relevant components and the reconstruction. The projecting matrix $\mathbf{U}$, the columns of which represent basis vectors in an M - dimensional space, will transform the input data vector $\mathbf{x}$ by

$$\mathbf{y} = \mathbf{U}^T\mathbf{x} \tag{1}$$

where $\mathbf{y} = (y_1, \ldots, y_L), L \leq M$ constitute a new representation of the data. The most widely used techniques to compute $\mathbf{U}$ are: principal component analysis (PCA), blind source separation (BSS) and independent component analysis (ICA). In the first case the projection matrix has orthogonal columns, i.e. $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, and it is possible to identify the components which keep the largest variance of the raw data. The goal of ICA is a decomposition into statistically independent components and the projecting matrix is usually a non-orthogonal matrix. BSS separation methods also achieve an non-orthogonal matrix but often only use second-order statistics to estimate the components of the data. However, after identifying the relevant (or irrelevant) components in $\mathbf{y}$, all methods are used in a similar way to obtain reconstructed data without the influence of undesired components. Considering a new vector $\hat{\mathbf{y}}$ with zero entries in place of the irrelevant components of $\mathbf{y}$, the reconstruction is defined as

$$\hat{\mathbf{x}} = \mathbf{B}\hat{\mathbf{y}} = (\mathbf{U}^T)^\dagger\hat{\mathbf{y}} \tag{2}$$

where † denotes the pseudo-inverse and $\mathbf{B} = \mathbf{U}$ in case of PCA. Note that as the vector $\hat{\mathbf{y}}$ has null entries there are

columns in $\mathbf{B}$ which do not contribute to the reconstruction of $\hat{\mathbf{x}}$.

A. Multidimensional Signals

The projecting techniques rely on a multidimensional representation of the data. In multi-sensor signal processing the data vector $\mathbf{x}$ is naturally formed with samples of the different sensors. However projective subspace techniques can also be applied to single sensor signals by forming vectors with windows of the signal. In what concerns EEG analysis such techniques have been proposed in the literature. In the following section we formalize both approaches by forming a representation of the data segment to compute the projective model parameters ($\mathbf{U}$).

1) Spatial approach: An electroencephalogram (EEG) is a multichannel recording and the multidimensional vector can be formed with samples from the M EEG channels at a given sampling time $t = nT$, $\mathbf{x}[n] = (x_1, x_2 \dots x_M)^T$. The projective techniques can be applied directly by forming a data matrix with N samples where each column represents a time sample of the multichannel recording

$$\mathbf{X} = \begin{bmatrix} x_1[0] & x_1[1] & \dots & x_1[N-1] \\ x_2[0] & x_2[1] & \dots & x_2[N-1] \\ x_3[0] & x_3[1] & \dots & x_3[N-1] \\ \vdots & \vdots & \ddots & \vdots \\ x_M[0] & x_M[1] & \dots & x_M[N-1] \end{bmatrix} \quad (3)$$

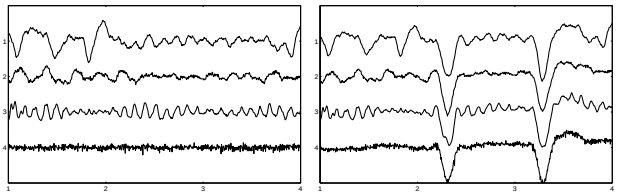
2) Temporal Approach: Time series analysis techniques often rely on embedding one dimensional sensor signals in the space of their time-delayed coordinates. Embedding can be regarded as a mapping that transforms a one-dimensional time series to a multidimensional sequence of lagged vectors. Considering a segment of the ith EEG channel $(x_i[0], x_i[1], \dots, x_i[N-1])$ the multidimensional signal is obtained by $\mathbf{x}_k = (x_i[k-1+M-1], \dots, x_i[k-1])^T, k = 1, \dots, K$. The lagged vectors lie in a space of dimension M, that constitute the columns of the *trajectory matrix*

$$\mathbf{X} = \begin{bmatrix} x_i[M-1] & x_i[M] & \dots & x_i[N-1] \\ x_i[M-2] & x_i[M-1] & \dots & x_i[N-2] \\ x_i[M-3] & x_i[M-2] & \dots & x_i[N-3] \\ \vdots & \vdots & \ddots & \vdots \\ x_i[0] & x_i[1] & \dots & x_i[N-M] \end{bmatrix} \quad (4)$$

Note that the matrix has identical entries along its diagonals. This trajectory matrix forms the data set to apply the projective techniques. The reconstruction then leads to $\hat{\mathbf{X}}$ which in general does not show identical elements along each descending diagonal of $\hat{\mathbf{X}}$ like in case of the original trajectory matrix $\mathbf{X}$. In singular spectrum analysis (SSA) these differing entries in each diagonal by are replaced by their average recovering again a Toeplitz matrix $\mathbf{X}_r$. An univariate $\hat{x}[n]$ is then obtained by reverting the embedding, i.e. by forming the signal with the mean of the values along each descendent diagonal of $\hat{\mathbf{X}}$.

III. Data collection

Twelve multi-channnel (16 leads with reference to Cz) EEG segments of $10s$ were selected by three specialists with a predominance in one of the characteristic bands and clean of artifacts (EOG, EMG or movements of the patient) to form the EEG data set. The data set will then be classified into four types according to the visibility/dominance of one of the activities: Type A - delta activity ; Type B - theta activity; Type C - alpha band and Type D - beta activity. For quantitative comparison purposes only one channel of each multichannel segments will be used, 3 for each of the characteristic band but not necessarily corresponding to the same lead. Figure 1shows an example for each segment. To facilitate the exposition the set of single-channel signals will be called reference signals.



(a) Original (b) Artificial mixtures

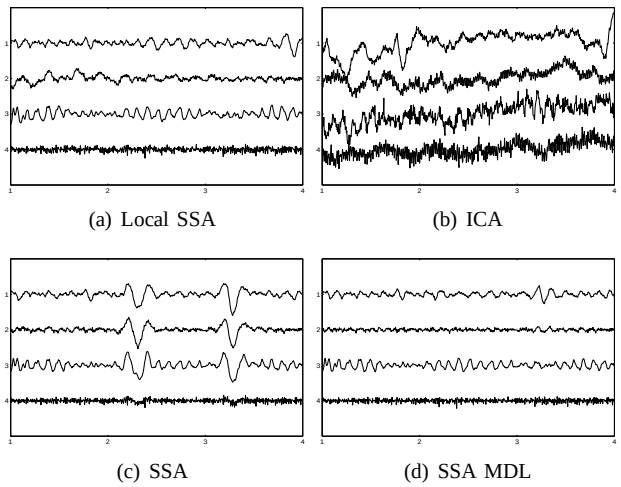
Fig. 1. Sub-segments EEG signals with different activities *Top to Down*: 1st - Delta (type A), 2nd - Theta (type B), 3rd- Alpha (type C) and 4th-Beta (type D).

The artifacts are obtained processing segments of EEG with clear interferences of EOG (usually corresponding to a frontal lead). For that purpose the EEG segments, belonging to different subjects, with 10s segment were processed by SSA and the artifact component is extracted as related with the largest eigenvalue. In total 10 EOG artifact were select with different amplitudes and a number of blinks ranging from 1-7 in the segment of $10s$. Note that SSA with one direction extracts always a component that is related with an artefact, however in most of the cases the corrected version of the EEG still shows some remnants of the artifact. By visual inspection the signals were recognized as clearly defined artifacts not showing any EEG relevant information.

3) Artificial mixtures: In this study it was followed an strategy proposed in [1] where the multichannel data set is obtained via $\mathbf{X} = \mathbf{AS}$, where S has the first M rows with M EEG signals and the last row is EOG artifact and the mixing matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \dots & a_1 \\ 0 & 1 & \dots & a_2 \\ 0 & 0 & \dots & a_3 \\ \vdots & \vdots & \ddots & \vdots \\ 0.5 & 0 & \dots & 1 \end{bmatrix} \quad (5)$$

Notice that each row of $\mathbf{X}$ is obtained by linearly adding two signals: the EEG and EOG. The coefficients, $\mathbf{a}_i, i \dots, M$ are chosen so that the amplitude of the artefact decreases from the frontal channels to the channels placed on the back of the head. The last row of $\mathbf{A}$ simulates the contamination of the EOG signal with a frontal EEG. For quantitative comparison purposes, between spacial and temporal approaches, only one channel of each of the multichannel segments will be used. In order to compare with the local SSA the coefficients



(a) Local SSA (b) ICA

(c) SSA (d) SSA MDL

Fig. 2. Corrected EEG subsegments using different algorithms.

(a_i) that correspond to the set of reference signals are chosen to have always the same value. Figure 1 illustrate an artificial mixture where the artefact is added to each original signal.

IV. Computing Projective Model

The data set $\mathbf{X}$ formed either using a spacial approach or a temporal approach can be used to remove artifacts in EEG recording systems. It has to be notice that with a temporal approach the artefact is removed from a single channel while with a spacial approach the artefact is removed from a set of channels.

A. Multichannel Approach

The spacial approach have been often used with ICA component algorithms. The matrix $\mathbf{X}$ might include or not rows with the EOG recording channels by considering that the mixing matrix has or not the last row. The EEGLAB [3] platform has built-in facility to remove the artifacts based on informax algorithm. In this case it only removed the component that is visually assigned as the EOG artifact and it has to be noticed that in all runs the algorithm finds the number of independent components equal to the number of mixtures provided (16 or 17). After the identification of the component related with EOG the multichannel recording is reconstructed using the pseudo-inverse of the the separation matrix. Figure 2-b) illustrate the output of the algorithm for sub-segments. Note that the signals on this figure correspond to different runs of the algorithm.

B. Uni-channel Approach

In a previous work [8]it was proposed an algorithm, called local SSA, that introduces a clustering step into SSA technique. The Singular Spectrum analysis uses PCA approach to compute the projecting matrix which is formed by keeping the eigenvectors that are related with the relevant components. In local SSA the PCA is performed in each cluster and as the goal is to extract the highest amplitude component the number of relevant components is identified

TABLE I

Correlation coefficients between original EEG and corrected EEG. The values correspond to 10s segments of the signals used on illustrations

Type	Local SSA	SSA MDL	SSA	ICA
A	0.54	0.33	0.38	0.38
B	0.63	0.39	0.37	0.11
C	0.87	0.92	0.67	0.045
D	0.95	0.95	0.91	0.34

by using an order selection criterium (minimum description length, MDL). The implementation of the algorithm requires the assignment of the following parameters: the embedding dimension (M) and the number of clusters. The heuristic rules proposed in the literature to choose the embedding dimension M point towards a minimum value that is related with the period of the signal to be extracted. An heuristic was also proposed to find the number of cluster [8]. Experimentally we conclude that the heuristics work quite consistently if the Local SSA is applied to all this signals but not for SSA. In particular, the embedding dimension can be chosen according to that heuristic ($M = 75, 250Hz$ is the sampling rate) while with SSA the M has also to be chosen according to the frequency contents of the EEG. In the latter case we show results for varying embedding dimensions M corresponding to the outcome with the highest correlation coefficient between the original and corrected EEG. The results for the one-channel approach are concerned with the following variants: 1)- Local SSA using MDL criterium to find the number of projections in each cluster;2)-SSA using only one direction; 3) SSA using MDL to select the number of projections. Note that the algorithm extracts the artefact, then the corrected EEG is obtained by subtracting the artefact to the corrupted signal. Figure 2 illustrates the output of the algorithm for sub-segment signals. Note that with SSA , with one projection the artefact is still visible except for the segments where beta dominates.

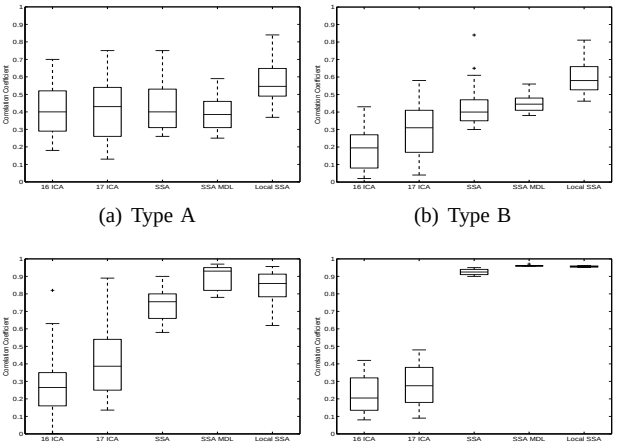
V. Discussion

The performance of the algorithms will be evaluated using measures in time domain and in frequency domain. The time domain measures the correlation coefficient computed between the corrected signals and the original (reference) signals. The frequency domain parameter pretends to evaluate the distortion in frequency by measuring the coherence function between the spectra of the two segments. Coherence Function values range between 0 and 1 and it is computed in the periodograms of subsegments i of the segments to be compared. For each k bin in frequency it is defined as

$$c_{xy}(k) = \frac{|\sum X_i^*(k)Y_i(k)|^2}{|X_i(k)|^2|Y_i(k)|^2} \ldots k = 0, 1, 2...K \qquad (6)$$

where X_i and Y_i represent the discrete fourier transform of the ith subsegment of the segments.

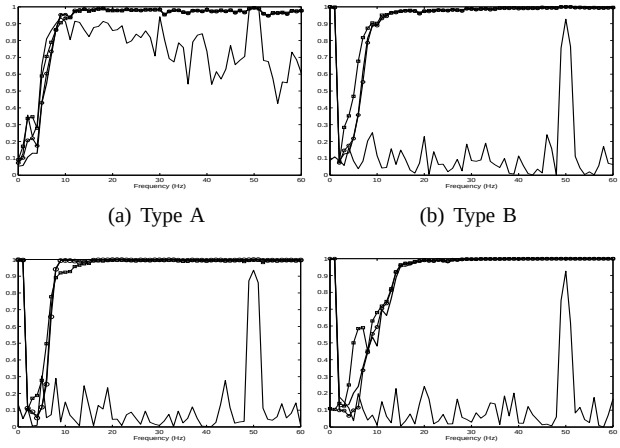
Figure 3 represents time domain results for the data set



(a) Type A (b) Type B

(c) Type C (d) Type D

Fig. 3. Correlation Coefficients between the corrected EEG signal and original



(a) Type A (b) Type B

(c) Type C (d) Type D

Fig. 4. Coherence in spectra between original and correct EEGs: RunICA (-); Local SSA (-□-); SSA (-o-) and SSA MDL (-)

considered. Globally we can see that ICA performs better when 17 mixtures are considered but worse than the SSA variants. Local SSA has an high correlation coefficient (> 0.8) for the segments where beta and alpha are dominant (Type C and Type D). When the dominant activity is theta (Type B) local SSA is better than the other variants. And when delta (Type A) dominates all the algorithms have the correlation coefficient in the same range. Table I presents the correlation coefficients for the segments represented in the fig. 2 and we can confirm that when correlation coefficient is low the distortions are visible.

In the frequency domain the results confirm the time domain ones. It has to be noticed that the coherence in frequency of ICA is very low, with the exception of $50Hz$. The table II shows the mean of the coherence in the frequency range of each characteristic band for signals of fig. 2. The coherence values for local SSA are high in the beta and alpha range whatever is the type of segment used in the mixture. For this method the distortion is on the delta band and is even higher when that reference signal is of delta type. In [10] it was also reported a distortion in the frequency range $5-25Hz$ when ICA algorithms are used but [6] reports that the distortion is higher for ICA algorithms based on higher order statistics.

TABLE II

COHERENCE IN FREQUENCY

	TYPE	Delta	Theta	Alpha	Beta
Local SSA	A	0.17	0.55	0.93	0.98
	B	0.66	0.73	0.89	0.95
	C	0.39	0.51	0.88	0.98
	D	0.55	0.71	0.90	0.99
ICA	A	0.21	0.52	0.32	0.11
	B	0.38	0.22	0.08	0.13
	C	0.40	0.05	0.105	0.13
	D	0.37	0.11	0.08	0.13

VI. ACKNOWLEDGMENTS

A. R. Teixeira received a PhD Scholarship (SFRH/BD/28404/2006) supported by the Portuguese Foundation for Science and Technology (FCT).

REFERENCES

[1] Charles W. Anderson, James N. Knight, Tim O'Connor, Michael J. Kirby, and Artem Sokolov. Geometric subspace methods and time-delay embedding for EEG artifact removal and classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):142–146, 2006.

[2] R. J. Croft and R. J. Barry. Removal of ocular artifact from the EEG: a review. *Neurophysiol. Clin.*, 30:5–19, 2000.

[3] A. Delorme and S. Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, 134:9–21, 2004.

[4] P. He, G. Wilson, C. Russel, and M. Gerschutz. Removal of ocular artifacts from the EEG: a comparison between time-domain regression method and adaptive filtering method using simulated data. *Medical & Biological Engineering & Computing*, 45(5):495–503, 2007.

[5] Tzyy-Ping Jung, Scott Makeig, Colin Humphries, Te-Won Lee, Martin J. Mckeown, Vicente Iragui, and Terrence j. Sejnowski. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37:163–178, 2000.

[6] Sergio Romero, Miguel A. Mañanas, and Manel J. Barbanoj. A comparative study of automatic techniques for ocular reduction in spontaneous eeg signals based on clinical target variables: a simulation case. *Computers in Biology and Medicine*, 38:348–360, 2008.

[7] A. Schögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller. A fully automated correction method of eog artifacts in eeg recordings. *Clinical Neurophysiology*, 118:98–104, 2007.

[8] A. R. Teixeira, A. M. Tomé, E.W. Lang, P. Gruber, and A. Martins da Silva. Automatic removal of high-amplitude artifacts from single-channnel electroencephalograms. *Computer Methods and Programs in Biomedicine*, 83(2):125–138, 2006.

[9] Elena Urrestarazu, Jorge Iriarte, Manuel Alegre, Miguel Valencia, César Viteri, and Julio Artieda. Independent component analysis removing artifacts in ictal recordings. *Epilepsia*, 45(9):1071–1078, 2004.

[10] 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 and component-based methods. *International Journal of Psychophysiology*, 53:105–119, 2004.