Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Quantitative evaluation of automatic ocular removal from simulated EEG signals: regression *vs.* second order statistics methods

Sergio Romero, Miguel Ángel Mañanas, and Manuel José Barbanoj

Abstract—**Analysis of the EEG by means of spectral parameters permit to evaluate the influence of a drug and to diagnose dysfunctional states in neurology, psychiatry and psychopharmacology. Eye movement artifacts contaminate EEG signals and can produce errors in this analysis. Regression based technique is considered the 'gold standard' artifact removal procedure and other techniques have been developed the last years, but few works have shown an objectively evaluation of the efficiency of these methods because it is impossible to record pure EEG and EOG signals. In this study, an artificially reproduction of bidirectional contaminated EEG and EOG data is proposed in order to simulate a real case. A comparative study between automatic second-order statistics techniques (PCA, AMUSE and SOBI) and multiple regression analysis is performed. Effectiveness of removal techniques is evaluated by calculating the errors in spectral parameters between sources and corrected EEG signals. Average values and topographic brain distribution of these errors are considered. Errors are located in the anterior leads especially in the frontopolar ones. Results show that AMUSE and SOBI methods preserve more cerebral activity than other techniques. We conclude that AMUSE and SOBI algorithms overcome the limitations of the regression based approach in the bidirectional contamination between ocular and neural activity.**

I. Introduction

QUANTITATIVE analysis of the electroencephalogram (EEG) allows to understand brain function in order to help in the clinical diagnosis of dysfunctional states in neurology, psychiatry and psychopharmacology. Changes in some parameters extracted from the EEG power spectral density function (PSD) have been found to be a valuable method for evaluating the influence of a drug and in the diagnosis. Relation between EEG rhythms and coarse physiological changes in brain states (alertness vs. sleep, sleep phases, coma and epileptic seizures), as well as between EEG rhythms and pharmacodynamic effects, are well established and analyzed [1]. There are four standard spectral bands of clinical interest: delta (0.5-3.5 Hz), theta (3.5-7.5 Hz), alpha (7.5-13 Hz) and beta (13-35 Hz).

This work was partially supported by a grant from CICYT (TEC2004-02274).

Sergio Romero is with the Automatic Control Department (ESAII), Technical University of Catalonia (UPC), Barcelona, Spain. (e-mail: sergio.romero-lafuente@upc.edu).

Miguel A. Mañanas is with the Automatic Control Department (ESAII), Technical University of Catalonia (UPC), Barcelona, Spain. (e-mail: miguel.angel.mananas@upc.edu).

Manuel J. Barbanoj is with the Drug Research Center (CIM), Research Institute of Sant Pau Hospital, Department of Pharmacology and Therapeutics, Autonomous University of Barcelona (UAB), Barcelona, Spain. (email: mbarbanoj@santpau.es).

EEG leads record not only cerebral activity but also interferences like beating of the heart, movement of the eyes and muscular activity from the face. Procedures to detect these artifacts and techniques to remove them are very important and necessary. The contamination of these artifacts in EEG signals could lead up to wrong results, conclusions and clinical decisions. Ocular artifacts are the most relevant interference because they occur very frequently [2]. The eye forms an electric dipole with the cornea (positive) and the retina (negative). As the eyeball moves, the electric field around the eye changes producing an electrical signal known as electrooculogram (EOG). Ocular activity propagates along the scalp. Electrodes for the acquisition of EOG signals also record EEG interference. Thus, EEG and EOG leads record a mixture of electrical ocular field and the brain activities located at the scalp sites. This is known as bidirectional contamination.

Multiple linear regression, in time or frequency domain, is the most commonly technique used to minimize eye movement artifacts [3]. This procedure estimates and removes the EOG component that is present in the EEG. As regression methods do not take into account the bidirectional contamination between EEG and EOG signals, they do not only reduce ocular activity but also potentially interesting cerebral information is cancelled out [4].

In order to solve the limitations of regression methods, other approaches based on a linear decomposition of the EEG and EOG data into components have shown to play an important role in EEG data analysis. These procedures identify artifactual components, and then they reconstruct the EEG signals without these components. There are several algorithms to estimate the components such as second-order statistics techniques and Independent Component Analysis based on higher-order statistics.

The best situation for the objective evaluation of the methods mentioned above is by means of a simulation study using EEG and EOG data. This work reproduces artificially corrupted EEG and EOG signals as a real situation in order to quantify their performance in removing interferences.

The aim of this paper is to evaluate objectively the efficiency of several automatic eye-movement filtering methods based on regression analysis and SOS techniques on simulated EEG and EOG data. Drawbacks are quantified by means of the errors in the standard spectral parameters estimated from the EEG signals.

II. METHODOLOGY

A. Subjects and instrumentation

Forty adult subjects participated in the study. Ten 3-minute epochs of EEG and EOG signals, sampled at 100 Hz, were recorded during vigilance-controlled EEG with eyes closed. During the experiment, the technician tries to keep the volunteers alert by acoustic estimulation. In addition to vertical and horizontal EOG (VEOG and HEOG, respectively), 19 EEG channels (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1 and O2) were recorded referenced to the averaged mastoid electrodes.

B. Simulated data

As the propagation of the signals (EEG and EOG) in the brain is very fast, a simultaneous linear mixture between real EEG and EOG signals is proposed in order to simulate a real case. Twenty 3-minute mixtures were simulated. These mixtures are called simulated subjects.

Regarding the ocular activity, twenty 3-minute epochs with vertical and horizontal ocular movements were selected from twenty subjects of the database. Propagation factors were calculated by regression analysis in these epochs [3]. EOG sources corresponded to VEOG and HEOG signals that were low-pass filtered with a cut-off frequency of 7.5 Hz in order to guarantee only most ocular activity remaining [5].

EEG sources were obtained following the criteria of the lowest ocular activity. Twenty 3-minute epochs were selected from the remainder twenty subjects of the database, where none of both EOG signals had amplitudes higher than 40μV. EEG sources were obtained by high-pass filtering with a cut-off frequency of 0.5 Hz the 19 EEG leads. This filtering removed the very slow and slow ocular components that could have the recorded EEG signals. Besides, spectral parameters that will be obtained for the evaluation of the EOG correction methods are calculated at frequencies higher than 0.5 Hz.

Ocular interferences were considered by means of the addition of VEOG and HEOG signals (weightened by their propagation factors) to the EEG sources. In the opposite side, cerebral interferences were added to EOG sources by means of average frontopolar (Fp1, Fp2) and average lateral-frontal (F7, F8) channels. Both average signals were added to VEOG and HEOG sources respectively, wheightening with their corresponding propagation factors (the same values in both directions: from EOG to EEG and from EEG to EOG).

C. Regression-based approach

Multiple regression analysis assumes that the recorded EEG signals (EEG_{raw}) are a linear, time-invariant superposition of different sources. Corrected EEG signals (EEG_{corr}) are calculated by subtracting a fraction of the VEOG and the HEOG signals from each EEG channel:

$$EEG_{corr}(t) = EEG_{raw}(t) - \alpha \cdot VEOG(t) - \beta \cdot HEOG(t) \quad (1)$$

where α and β represents the propagation factors of the VEOG and HEOG signals, respectively, to the EEG channel. Equation (1) is applied to each EEG lead with its corresponding factors α and β. These factors are estimated using only samples of each EEG and EOG leads with high EOG activity [3].

D. Component-based approaches

Component-based techniques decompose the multi-channel EEG and EOG data into a mixture of source signals. There are several techniques to solve the so-called blind source separation (BSS) problem and the following were used in this study: Principal Component Analysis (PCA), second-order statistics procedures (SOS). The BSS problem assumes than a set of *m* recorded channels are composed by a mixture of *n* source components with $n \le m$.

PCA technique finds components that are spatially orthogonal and uncorrelated. In general, no reason exists for EEG and EOG signals to be orthogonal. Second-order statistics are sufficient to solve the linear instantaneous BSS problem, if temporal information is taken into account. In this paper, two SOS algorithms are applied exploiting the spatio-temporal decorrelation: AMUSE [6] and SOBI [7]. Both are based on singular or eigen-value decomposition of a linear combination of several time-delayed covariance matrices [8]. Different lengths of epochs are used for the application of these algorithms.

E. Automatic filtering criteria

Automatic artifact identification is based on frequency and scalp topography aspects of the component and was previously described in [9]. Then, corrected EEG signals are obtained by means of reconstruction of the components removing those which are considered as ocular artifacts [4].

F. Validation of correction

Spectral analysis is performed for all EEG leads. PSD functions are calculated by means of Welch-periodogram using a Hanning window of 5-second duration as usual [10]. Several parameters are calculated from the PSD functions: total power (0.5-35 Hz), absolute and relative power of the following bands: delta (0.5-3.5 Hz), theta (3.5-7.5 Hz), alpha (7.5-13 Hz) and beta (13-35 Hz). Relative parameters are calculated with respect to the total power. These band power values are calculated for further comparison before and after EOG correction by all the mentioned techniques.

The mean square error (MSE) between the sources and the corrected EEG signals were evaluated for all the methods. Percentage error in each spectral parameter between the sources (real values) and the corrected EEG was calculated for all the EEG leads. Decomposition procedures are performed using the functions included in the ICALAB toolbox v2.2 for Matlab [11].

III. Results

As an example of the simulation procedure, Fig. 1 shows an epoch corresponding to the EOG and EEG sources coming from two different subjects in order to guarantee the signal independence. The bidirectional contamination between EOG and EEG data can be observed in the result of the mixing simulation procedure.

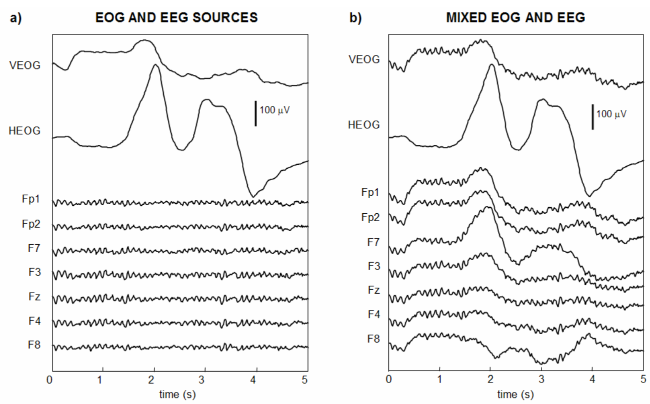

Fig. 1. a) A five second epoch of EOG and EEG sources. b) Mixed EOG and EEG signals after applying the simulation procedure.

In order to demonstrate the visual effectiveness of the proposed automatic artifact correction procedures on spontaneous EEG signals, two EEG leads, as an example, corresponding to the same 5-second epoch of Fig. 1 before and after applying the EOG removal procedures to the mixed EEG data can be seen in Fig. 2. Performance of the proposed automatic artifact correction procedures on spontaneous EG signals can be observed visually in this figure. Because the propagation factors are very high at the anterior electrodes (Fp1: VEOG 0.972 ± 0.072, HEOG 0.011 ± 0.070; F7: VEOG 0.439 ± 0.155, HEOG 0.302 ± 0.061), regression-based removal also reduces neural activity recorded in EOG data, as it can be seen in the corrected Fp1 and F7 signals. It can be observed that PCA also removes interesting cerebral activity at both channels.

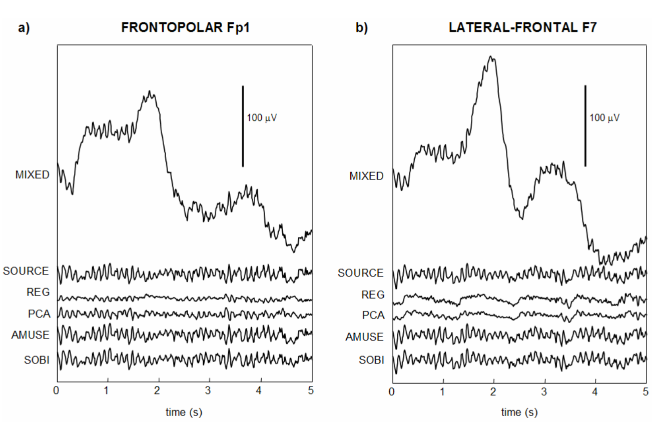

Fig. 2. Five second epoch corresponding to mixed and corrected EEG signals by different correction techniques: a) Fp1; b) F7.

Component-based procedures were applied to different epoch durations: 30, 60, 90, 120 and 180 seconds. Additional sample points can improve the decomposition, only if relative stationarity of the structure of the sources can be assumed. For regression analysis removal, propagation factors only depend on the subject and the electrode location at the scalp. Thus, these factors are obtained considering all 180 seconds available in each simulated subject. Better estimations of the propagation factors are obtained using all the 3-minutes data available. If few seconds are used for the computation of the propagation coefficients, including few ocular artifacts, it is difficult to afford reliable factors because there are not enough ocular movements for the estimation.

Fig. 3 shows the MSE and several percentage errors in mean values of all the EEG leads and all the simulated subjects using different decomposition duration segments. Similar errors are obtained in each component-based algorithm for the different epoch durations. Errors for AMUSE and SOBI algorithms are smaller than regression and PCA. Only slight higher errors are obtained in 30 second duration with respect to the other segment durations.

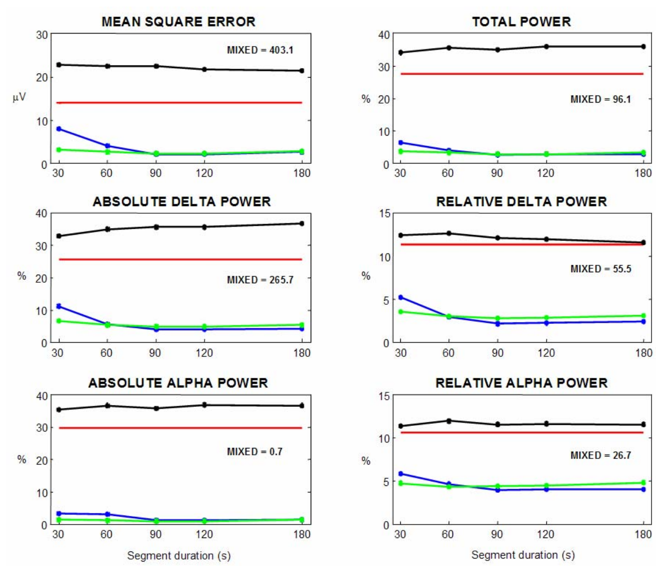

Fig. 3. Mean (all the EEG channels) MSE and percentage errors between simulated and estimated source signals calculated by means of: regression (red), PCA (black), AMUSE (blue) and SOBI (green). Errors between source and mixed EEG are also indicated in the plots.

The specific influence of the ocular removal in each EEG lead is evaluated by representing the distribution of the MSE and the spectral percentage errors along the scalp. Fig. 4 shows the topographic distribution of the same errors shown in the Fig. 3 for each EOG correction method using 90 second duration epochs for the decomposition procedures. Absolute and relative powers for theta and beta bands present similar error distributions but smaller values than absolute and relative powers for delta and alpha bands, respectively. The bidirectional mixing simulation procedure introduces several distortions in the broadband spectral power of the EOG and the EEG signals. Simulated ocular artifacts only affect delta and theta EEG bands. Topographic results are coherent with the distribution across the scalp of the electrical potentials generated by the eye activity. Errors are located in the anterior leads and especially in the frontopolar ones. SOBI and AMUSE algorithms present lower errors between the sources and the corrected EEG signals. These SOS methods also preserve more cerebral activity in the anteriorly placed electrodes than regression and PCA procedures.

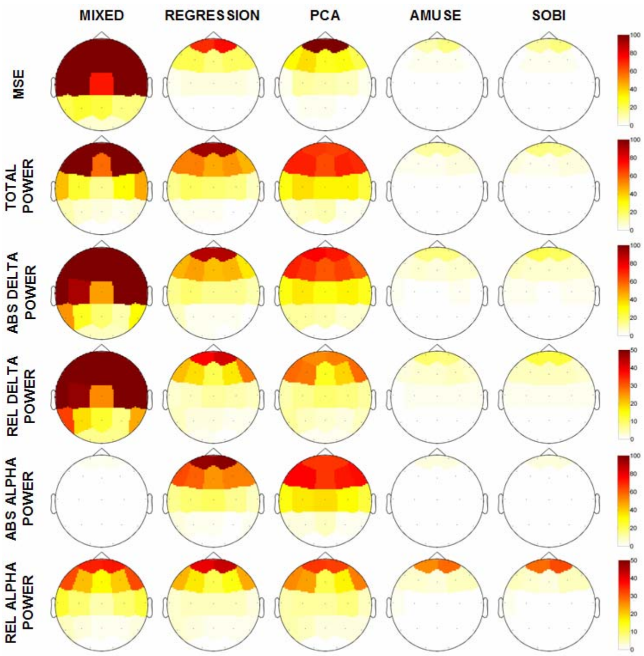

Fig. 4. Topographic maps of the errors (MSE and percentage) between the sources and the corrected EEG signals by regression, PCA, AMUSE and SOBI. Mixed column denotes that no correction is made. The color key depicts μV^2 for MSE, and % for percentage errors.

IV. Discussion

EOG contamination of EEG data is a very important and common problem in the diagnosis of neurobiological events because PSD functions of EEG and EOG signals overlap in frequency. Results extracted from EEG signals not treated in a satisfactory way could be spurious and wrong. In general, artifact minimization is preferable to artifact rejection, since no loss of data is entailed when limited data are available, or eye movements occur too frequently.

In this work, an empirically automatic component-based method for the reduction of ocular artifacts in the EEG signals is applied in simulated mixtures for evaluating the bidirectional contamination problem between the EEG and the EOG signals.

A comparison between regression analysis and other component-based approaches is performed in order to show which ocular reduction technique is the best to apply in the particular spontaneous EEG condition.

Time and frequency results indicate that spatio-temporal decorrelation procedures, like AMUSE and SOBI algorithms, preserve and recover more brain activity than regression and PCA in decomposing real EEG data. These SOS-based techniques find source components more realistic in physical sense, because they do not force the mixing matrix to be orthogonal like PCA algorithm. Regression analysis, considered as the ‘gold standard’ artifact removal procedure, provides a relatively simple mathematical solution for EOG correction by removing ocular component from the EEG channels. However, a part of the neural component is also subtracted by the regression process due to the mutual contamination between EEG and EOG channels. This statement is shown by percentage errors in the absolute alpha power obtained with regression analysis especially in anterior located EEG leads, although there is no influence of ocular artifacts in this parameter in the mixed signals (see Fig. 4).

Spatio-temporal decorrelation procedures are simple, fast to compute and robust algorithms. Unlike, other more complicated HOS procedures, SOS techniques also permit the separation of Gaussian signals and need less data length in order to estimate reliable sources. SOS techniques appear to be a generally applicable and effective method for removing ocular artifacts from spontaneous EEG recordings since they simultaneously separate both the EEG and EOG data into components, without relying on the availability of clear reference signals.

References

[1] E. Basar, *Dynamics of sensory and cognitive processing by the brain*, Springer Verlag, 1988.

[2] J.S. Barlow, “Artifact processing (rejection and minimization) in EEG data processing” in *Handbook of electroencephalography and clinical neurophysiology*, Ed. by Lopes da Silva, Storm van Leeuwen, A. Rémond. Elsevier, vol. 2, 1986, pp. 15-62.

[3] P. Anderer, H.V. Semlitsch, B. Saletu and M.J. Barbanoj, “Artifact processing in topographic of electroencephalographic activity in neuropsychopharmacology”, *Psychiatry Research*, vol. 45, 1992, pp. 79-93.

[4] T-P. Jung, S. Makeig, C. Humphries, et al., “Removing electroencephalographic artifacts by blind source separation”, *Psychophysiology*, vol. 37, 2000, pp. 163-178.

[5] T. Gasser, P. Ziegler, W.F. Gattaz, “The deleterious effect of ocular artifacts on the quantitative EEG, and a remedy”, *Eur Arch Psychiatry Clin Neurosci*, vol. 241, 1992, pp. 352-356.

[6] L. Tong, R-W. Liu, V.C. Soon, Y-H. Huang, “Indeterminacy and identifiability of blind identification”, *IEEE Transactions on Circuits and Systems*, vol. 38, no. 5, 1991, pp. 499-509.

[7] A. Belouchrani, K. Abed-Meraim, J.F. Cardoso, E. Moulines, “A blind source separation technique using second-order statistics”, *IEEE Trans. Signal Process.*, vol. 45, no. 2, 1997, pp. 434-444.

[8] A. Hyvärinen, J. Karhunen, E. Oja, *Independent Component Analysis*, John Wiley&Sons, 2001.

[9] S. Romero, M.A. Mañanas, J. Riba, et al., “Evaluation of an automatic ocular filtering method for awake spontaneous EEG signals based on Independent component analysis", *Proc. Conf. IEEE-EMBC*, vol. 1, 2004, pp. 925-928.

[10] P. Anderer, B. Saletu, K. Kinsperger, H. Semlitsch, “Topographic brain mapping of EEG in neurophsychopharmacology – Part I. Methodological aspects”, *Meth and Find Exptl Clin Pharmacol*, vol. 9(6), 1987, pp.371-384.

[11] A. Cichocki, S. Amari, K. Siwek, et al., *ICALAB toolboxes*. Available at: http://www.bsp.brain.riken.go.jp/ICALAB/