A new muscle artifact removal technique to improve the interpretation of the ictal scalp electroencephalogram

Wim De Clercq[1,*], Anneleen Vergult[1], Bart Vanrumste[1], Johan Van Hees[2], André Palmini[2], Wim Van Paesschen[2] and Sabine Van Huffel[1]

[1]Department of Electrical Engineering (ESAT/SCD), Katholieke Universiteit Leuven, Leuven, Belgium
[2]Department of Neurology, University Hospital Gasthuisberg, Katholieke Univeristeit Leuven, Leuven, Belgium
[*] E-mail: wim.declercq@esat.kuleuven.ac.be

Abstract–**In this paper a new method for muscle artifact removal in EEG is presented, based on Canonical Correlation Analysis (CCA) as a Blind Source Separation technique (BSS). This method is demonstrated on a synthetic data set. The method outperformed a low pass filter with different cutoff frequencies and an Independent Component Analysis (ICA) based technique for muscle artifact removal. The first preliminary results of a clinical study on 26 ictal EEGs of patients with refractory epilepsy illustrated that the removal of muscle artifact results in a better interpretation of the ictal EEG, leading to an earlier detection of the seizure onset and a better localization of the seizures onset zone. These findings make the current method indispensable for every Epilepsy Monitoring Unit.**

Keywords– **Muscle artifact removal, canonical correlation analysis, blind source separation, ictal EEG.**

I INTRODUCTION

The EEG is frequently contaminated by electrophysiological potentials associated with muscle contraction due to biting, chewing and frowning. These muscle artifacts obscure the EEG and complicate the interpretation of the EEG or even make the interpretation unfeasible [1].

Low-pass filters are commonly used to remove muscle artifact. However, as the frequency spectrum of the muscle artifacts overlaps with that of interesting brain signals [2], frequency filters not only suppress muscle artifacts but also valuable information, such as ictal beta activity [3]. A more recently explored approach is Independent Component Analysis (ICA) which separates the EEG into statistical independent components. However, cross-talk can be observed when the separation of brain and muscle activity is considered [4,5]. Moreover, when applying ICA, the identification of the components containing artifacts in general, and muscle activity in particular, is not obvious, and requires additional user attention.

In [6] a new method for muscle artifact correction in EEG is presented, based on the statistical Canonical Correlation Analysis (CCA) applied as a Blind Source Separation (BSS) technique, further referred to as BSS-CCA. In this paper the performance of the method is tested on synthetic data and compared with a low pass filter with different cutoff frequencies and an ICA based technique for muscle artifact removal. Furthermore, we present the first results of an evaluation of the method on a group of 26 patients with refractory partial epilepsy. The aim of this clinical study was to evaluate the impact of our new muscle artifact removal algorithm on the readability of ictal EEG.

II METHODS

A. *Blind source separation by CCA for artifact removal*

In the blind source separation problem, the observed time course $\mathbf{x}(t) = [x_1(t), x_2(t), ..., x_K(t)]^T$, with $t = 1, ..., N$, with N the number of samples and K the number of sensors, is the result of an unknown linear mixture of a set of unknown source signals $\mathbf{s}(t) = [s_1(t), s_2(t), ..., s_K(t)]^T$:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \tag{1}$$

where $\mathbf{A}$ is the unknown mixing matrix. The goal is to estimate the mixing matrix and recover the original source signals $\mathbf{s}(t)$. This is carried out by introducing the de-mixing matrix $\mathbf{W}$ such that

$$\mathbf{z}(t) = \mathbf{W}\mathbf{x}(t), \tag{2}$$

approximates the unknown source signals in $\mathbf{s}(\mathbf{t})$, by a scaling factor. Unless there are extra constraints imposed, it is in general impossible to solve this problem. Canonical correlation analysis solves the problem by forcing the sources to be mutually uncorrelated and maximally correlated with a predefined function [7].
Let $\mathbf{x}(t)$ be the observed data matrix with K mixtures and N samples, then we define the predefined function $\mathbf{y}(t)$ as a temporally delayed version of the original data matrix, to enforce the sources to be maximally autocorrelated [7]:

$$\mathbf{y}(t) = \mathbf{x}(t-1). \tag{3}$$

For the remainder of the chapter, BSS-CCA refers to this autocorrelation version. Consider the linear

combinations of the components in $\mathbf{x}$ and $\mathbf{y}$:

$$\begin{aligned} x &= w_{x_1}x_1 + ... + w_{x_K}x_K = \mathbf{w_x}^\mathbf{T}\mathbf{x}, \\ y &= w_{y_1}y_1 + ... + w_{y_K}y_K = \mathbf{w_y}^\mathbf{T}\mathbf{y}, \end{aligned} \tag{4}$$

CCA finds the vectors $\mathbf{w_x} = [\mathbf{w_{x_1}}, ..., \mathbf{w_{x_K}}]^\mathbf{T}$ and $\mathbf{w_y} = [\mathbf{w_{y_1}}, ..., \mathbf{w_{y_K}}]^\mathbf{T}$ that maximize the correlation ρ between x and y by solving the following maximization problem:

$$\max_{\mathbf{w_x},\mathbf{w_y}} \rho(x,y) = \frac{E[xy]}{\sqrt{E[x^2]E[y^2]}} \tag{5}$$

The canonical correlations ρ, which are in the BSS-CCA case equal to the autocorrelations of the sources, are given by the cosines of the principal angles between the row spaces of $\mathbf{x}^\mathbf{T}$ and $\mathbf{y}^\mathbf{T}$ [8]. The canonical variates x^T and y^T correspond to the principal directions in the corresponding row space of $\mathbf{x}^\mathbf{T}$ and $\mathbf{y}^\mathbf{T}$. Let $\mathbf{x}^\mathbf{T}$= $\mathbf{Q_x R_x}$ and $\mathbf{y}^\mathbf{T}$= $\mathbf{Q_y R_y}$ be the QR decompositions of $\mathbf{x}^\mathbf{T}$ and $\mathbf{y}^\mathbf{T}$, respectively. The canonical correlations and canonical variates can then be obtained from the SVD of $\mathbf{Q_x}^\mathbf{T}\mathbf{Q_y}$ as follows [8]:

$$\mathbf{Q_x}^\mathbf{T}\mathbf{Q_y} = \mathbf{ECF}^\mathbf{T}. \tag{6}$$

The autocorrelations ρ are given by the diagonal elements of $\mathbf{C}$. The columns of $\mathbf{z}^\mathbf{T} = \mathbf{Q_x E}$ give the canonical variates of $\mathbf{x}^\mathbf{T}$, corresponding to the estimates of the sources $\mathbf{s}_i(t)$. The columns of $\mathbf{z}^\mathbf{T}$ are ordered by decreasing autocorrelation.

When BSS-CCA is applied to the EEG and the sources, or components, contributing to the EEG are derived, the muscle artifact can be removed by setting the columns representing the activations of the artifactual sources equal to zero in the reconstruction $\mathbf{x}_{clean}(t)$:

$$\mathbf{x}_{clean}(t) = \mathbf{A}_{clean}\mathbf{z}(t), \tag{7}$$

with $\mathbf{z}(t)$ the sources obtained by BSS-CCA, and $\mathbf{A}_{clean}$ the mixing matrix with the columns representing activations of the muscle artifactual sources, set to zero. We observed that the muscle artifact components are well distinguished from the components related to the brain activity [6]. Moreover, we observed that the muscle activity is present in the lowest components in the CCA decomposition. These observations were exploited to establish a semi-automatical removal of muscle artifacts [6]. CCA components are gradually removed from bottom upwards based on the ordering of the components.

III SIMULATION STUDY

The aim of the simulation study was to evaluate the performance of the proposed method in removing muscle artifacts. Several simulations with different signal-to-noise ratios were performed.

A. *Brain activity*

An EEG epoch of 10 s containing mainly delta ($0.5-4$Hz) activity was selected as underlying brain signal. In the selected EEG no muscle artifact was present according to the visual inspection of an experienced neurophysiologist. The data was collected from 21 scalp electrodes placed according to the International 10-20 System [9] with additional electrodes $T1$ and $T2$ on the temporal region. The sampling frequency was 250 Hz and an average reference was used. The EEG epoch was stored in a 21-by-2500 dimensional matrix $\mathbf{B}$.

B. *Muscle activity*

To obtain only muscle activity, necessary for the simulation study, it is not sufficient to select muscle artifacts in the EEG as these events will not only contain muscle activity but also activity from the brain. To remove the EEG activity we used ICA (SOBI) [10] to decompose these events. Note that for a large number of events, which we visually inspected, no clear separation was established. For those events where a clear separation between muscle and brain activity was obtained, we selected a component that accounted for muscle artifact together with the corresponding field distribution. This procedure was repeated for three different 10 s EEG epochs recordings from three different subjects. The resulting average referenced signal containing all three independent muscle activities was stored in a 21-by-2500 dimensional matrix $\mathbf{M}$.

C. *Data and measures to test the performance*

In the simulation study the average referenced muscle artifact signal $\mathbf{M}$ was superimposed on the average referenced signal $\mathbf{B}$ containing only brain activity:

$$\mathbf{X}(\lambda) = \mathbf{B} + \lambda \cdot \mathbf{M}, \tag{8}$$

with λ representing the contribution of muscle activity. The Root Mean Squared (RMS) value of the brain signal is then equal to

$$RMS(\mathbf{B}) = \sqrt{\frac{1}{K \cdot N}\sum_{k=1}^{K}\sum_{n=1}^{N}\mathbf{B^2}(k,n)}, \tag{9}$$

with N equal to the number of time samples and K equal to the number of EEG channels. An important measure is the Signal-to-Noise Ratio (SNR) which is defined as follows,

$$SNR = \frac{RMS(\mathbf{B})}{RMS(\lambda \cdot \mathbf{M})}. \tag{10}$$

Changing the λ parameter alters the signal-to-noise ratio of the simulated signal. Several simulations with different signal-to-noise ratios were performed.

The performance is expressed in terms of the Relative Root Mean Squared Error (RRMSE):

$$RRMSE = \frac{RMS(\mathbf{B} - \hat{\mathbf{B}})}{RMS(\mathbf{B})}, \qquad (11)$$

where $\hat{\mathbf{B}}$ is the signal after muscle artifact removal.

D. Methods for comparison

For comparison, the proposed technique for muscle artifact removal is compared with other commonly used techniques for this purpose:

- Low pass filter: a low pass butterworth filter of order 8 was used with four different cutoff frequencies equal to 10Hz, 15Hz, 20Hz and 30Hz.
- ICA: as it is not in the scope of this paper to compare all different ICA algorithms, we limited ourself to the Joint Approximate Diagonalization of Eigen-matrices (JADE) algorithm [11]. This ICA algorithm was previously used in a study for eye and muscle artifact removal in ictal scalp EEG recordings [5]. The muscle artifact was removed as follows. The ICA components were calculated by using the JADE algorithm. The selection of the components accounting for the muscle artifact was based on visual inspection and the signal was reconstructed, excluding the components related to the artifact.

E. Results of simulation study

The RRMSE as a function of signal-to-noise ratio on the for different artifact removal techniques (BSS-CCA, ICA(JADE) and the low pass filter with cutoff frequency at 10Hz, 15Hz, 20Hz and 30Hz) is shown in Fig. 1. Compared to the low pass filters, the performance of ICA was better for almost all signal-to-noise ratios. The BSS-CCA method outperformed all methods for all SNRs.

IV METHOD APPLIED TO ICTAL EEG RECORDINGS

A. Materials

One ictal EEG of 26 patients with refractory partial epilepsy and an ictal onset zone that was well-defined during a presurgical evaluation, were processed with the muscle artifact removal technique. The artifact-free recordings were compared with the band pass (0.3-35 Hz) filtered original EEGs by an unblinded neurologist.

B. Results

In 24 of 26 cases (92%) muscle artifact contaminated the ictal EEGs significantly. In all 24 cases the ictal EEG was easier to interpret after the muscle artifact was removed. In 11 out of 24 cases (42%) the interpretation of the ictal EEG improved in a clinically significant way; a more precise overview of the improvements is given below. The time of ictal onset on EEG was detected earlier in 9 out of 26 cases (35%), the onset of the seizure was better localized in 7 out of 26 (27%) and the ictal-pattern of the onset was located in a higher frequency range for 8 out of 26 cases (31%). Furthermore, we observed localized ictal-onset beta activity only after removal of muscle artifact in 5 patients (19%). Worell et al. (2002) have shown that focal ictal beta discharges localize the ictal onset zone accurately and are highly predictive of excellent postsurgical outcome [3]. The muscle artifact removal technique did not degrade the EEG signal in any of the patients. The results are summarized in Table 1.

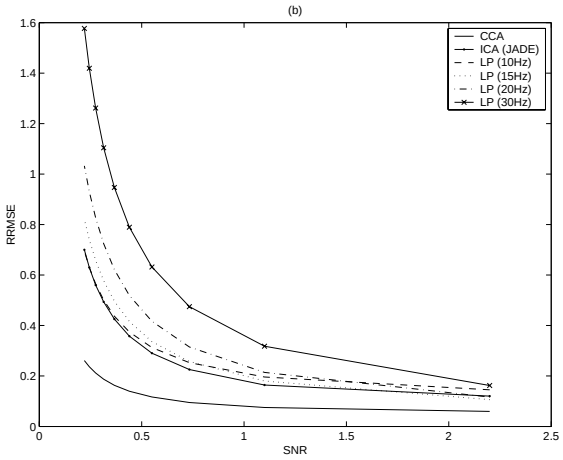


Figure 1. The RRMSE as a function of SNR.

Fig. 2 shows an ictal EEG epoch of one of the 11 cases for which the muscle artifact removal method improved the interpretation. In Fig. 2 (a) the band pass (0.3-35 Hz) and notch (50 Hz) filtered EEG epoch is shown. Fig. 2 (b) shows the resulting ictal EEG epoch after muscle artifact removal by BSS-CCA. Notice that the muscle artifact was removed completely, while preserving the sharp quality ictal activity on the $T6$ and $O2$ electrodes. The ictal onset, which was masked by muscle artifact activity, becomes visible in the BSS-CCA processed ictal EEG recording.

Table 1. Preliminary results of the clinical study.

Muscle artifact contamination	24/26 (92%)
Easier interpretation	24/24 (100%)
Improved interpretation	11/24 (42%)
Earlier detection of ictal-onset in EEG	9/26 (35%)
Better localization of onset	7/26 (27%)
Onset pattern of higher frequency	8/26 (31%)
Localized onset beta activity only after removal of muscle artifact	5/26 (19%)
Degradation of EEG by the method	0/26 (0%)

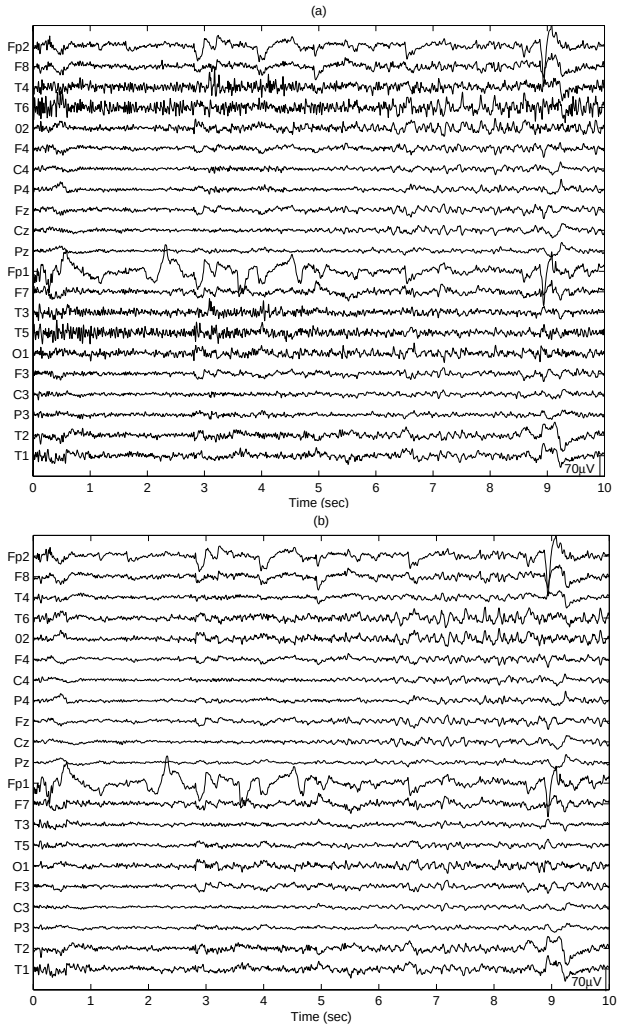


Figure 2. (a) Band pass (0.3-35 Hz) and notch (50 Hz) filtered ictal EEG epoch of 10 s; (b) ictal EEG epoch after muscle artifact removal by BSS-CCA

V DISCUSSION

In this paper the performance of a new method for muscle artifact removal in EEG was tested on synthetic data. The method outperformed a low pass filter with different cutoff frequencies and an Independent Component Analysis (ICA) based technique for muscle artifact removal. The first preliminary results of a clinical study on 26 ictal EEGs of patients with refractory epilepsy illustrated that the removal of muscle artifact improves the interpretation of the ictal EEG in a clinically significant way in around 40% of the patients. These findings make the current method indispensable for every Epilepsy Monitoring Unit. The method is now being evaluated on a larger patient group by three neurologists.

ACKNOWLEDGMENTS

This research is sponsored by the Research Council of the K.U.Leuven (GOA-AMBioRICS), the Flemish Government (FWO: projects G.0360.05, research communities ICCoS & ANMMM), the Belgian Federal Government (DWTC: IUAP V-22 (2002-2006)), and the EU (PDT-COIL: contract NNE5/2001/887; BIOPATTERN: contract FP6-2002-IST 508803). Bart Vanrumste is funded by the 'Programmatorische Federale Overheidsdienst Wetenschapsbeleid' of the Belgian Government.

REFERENCES

[1] S. Spencer, P. Williamson, S. Bridgers, R. Mattson, D. Cicchetti, and D. Spencer, "Reliability and accuracy of localization by scalp ictal EEG," *Neurology*, vol. 30, pp. 1567–1575, 1985.

[2] I. Goncharova, D. McFarland, T. Vaughan, and J. Wolpaw, "EMG contamination of EEG: Spectral and topographical characteristics," *Clinical Neurophysiology*, vol. 114, pp. 1580–1593, 2003.

[3] G. Worell, E. So, J. Kazemi, T. O'Brien, R. Mosewich, G. Cascino, F. Meyer, and W. Marsh, "Focal ictal beta discharge on scalp EEG predicts excellent outcome of frontal lobe epilepsy surgery," *Epilepsia.*, vol. 43, no. 3.

[4] H. Nam, T.-G. Yim, S. Han, J.-B. Oh, and S. Lee, "Independent component analysis of ictal EEG in medial temporal lobe epilepsy," *Epilepsia*, vol. 43, no. 2, pp. 160–164, 2002.

[5] E. Urrestarazu, J. Iriarte, M. Alegre, M. Valencia, C. Viteri, and J. Artieda, "Independent component analysis removing artifacts in ictal recordings," *Epilepsia*, vol. 45, no. 9, pp. 1071–1078, 2004.

[6] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram," *IEEE Trans. Biomed. Eng., submitted for publication.*

[7] O. Friman, *Adaptive analysis of functional MRI Data.* PhD thesis, Linköpings University, Linköping, 2003.

[8] G. Golub and C. F. Van Loan, *Matrix computations.* Baltimore: John Hopkins University Press, third ed., 1996.

[9] M. Nuwer, G.Comi, R. Emmerson, *et al.*, "IFCN standards for digital recording of clinical EEG," *Electroencephalogr Clin Neurophysiol*, vol. 106, pp. 259–261, 1998.

[10] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique usingsecond order statictics," *IEEE Trans. Signal Processing*, vol. 45, pp. 434–444, 1997.

[11] J.-F. Cardoso, "Higher-order contrast for independent component analysis," *Neural Comput*, vol. 11, pp. 157–192, 1999.