35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Multiple Time-Lag Canonical Correlation Analysis for Removing Muscular Artifacts in EEG

Kaiquan Shen, Ke Yu, Aishwarya Bandla, Yu Sun, Nitish Thakor, Xiaoping Li

Abstract— **In this work, a new approach for joint blind source separation (BSS) of datasets at multiple time lags using canonical correlation analysis (CCA) is developed for removing muscular artifacts from electroencephalogram (EEG) recordings. The proposed approach jointly extracts sources from each dataset in a decreasing order of between-set source correlations. Muscular artifact sources that typically have lowest between-set correlations can then be removed. It is shown theoretically that the proposed use of CCA on multiple datasets at multiple time lags achieves better BSS under a more relaxed condition and hence offers better performance in removing muscular artifacts than the conventional CCA. This is further demonstrated by experiments on real EEG data.**

I. INTRODUCTION

Scalp EEG does not only record cerebral activity but is also contaminated by non-cerebral electrical sources like line noise, cardiac signals, eye blinks and movements, and muscle activities. Line noise and other extraphysiological artifacts can be greatly attenuated by proper calibration, right instrumentation or notch filters. Likewise, cardiac artifacts can be removed, almost completely, by adaptive filters using a reference ECG channel or by a surface Laplacian spatial filter. Also, eye blinks and movements can be effectively handled by adaptive filters with extra sensors HEOG and VEOG placed near the eyes, or by a combination of a blind source separation (BSS) method and a machine learning technique (see e.g. [1]). Thus, in most practical settings, only muscular artifacts, i.e. electromyographic (EMG) artifacts, pose a real challenge for the inferential interpretation of the electroencephalogram (EEG).

Low-pass filters are commonly used to remove muscular artifacts. However, since the frequency spectrum of muscular artifact overlaps significantly with that of cerebral signals, low-pass filters not only suppress muscular artifacts but also cerebral signals. Linear regression methods and adaptive filters are relatively effective in suppressing muscular artifacts, but they usually require dedicated reference EMG channels (which are usually not desirable or not possible due to application constraints). More recently explored approaches are the so-called transform methods or spatial filtering methods, such as independent component analysis (ICA), second order blind identification (SOBI) and canonical correlation analysis (CCA). Although they were not originally designed for EEG artifact removal, they appear very promising for the purpose of EEG artifact removal.

K. Shen, K. Yu, A. Bandla, Y. Sun, N. Thakor are with the Singapore Institute of Neurotechnology (SINAPSE), the National University of Singapore. (phone: 65-66012658; fax: 65-68733905; e-mails: lsiskq@nus.edu.sg; c_y@nus.edu.sg; lsisu@nus.edu.sg; sinapsedirector@gmail.com).

X. Li is with the department of mechanical engineering, the National University of Singapore. (e-mail: mpelixp@nus.edu.sg).

ICA [2-4] has been extensively studied for artifact removal in EEG [5] [1, 6-12]. This is mainly motivated by the fact that ICA is effective in decomposing raw EEG recordings into spatially-fixed but temporally-dynamic components. Artifact-free EEG data can then be reconstructed by subtracting components due to artifacts from the raw EEG data. Two points are worth mentioning. First, ICA exploits only the spatial diversity of the source signals and it uses only the marginal distribution of the observations in searching for spatially-fixed and temporally-dynamic independent sources. Therefore, they are applicable even when source signals are temporally independent. However, EEG artifact typically has unique temporal correlation which can be exploited for better source separation. Second, ICA has shown consistent success in removing ocular artifacts and some success in removing continuous EMG artifact in the context of epileptic seizures. However, the separation between EEG and EMG sources was in many cases not optimal, see e.g. [13]. Recent attempts to study the sensitivity and specificity of ICA as an EMG removal tool also concluded that ICA was not able to provide perfect EMG artifact removal, although it did outperform conventional regression-based correction techniques [14, 15].

SOBI [16] or equivalently TDSEP by [17] seeks a transformation that simultaneously diagonalizes several correlation matrices at different lags. Since, in general, no transformation may exist to accommodate such a stringent condition, a function that objectively measures the degree of joint (approximate) diagonalization (JD) at different lags is employed instead. It has been reported that SOBI showed significant improvement over ICA. However, traditional SOBI methods consider stationary sources and it may suffer when non-stationary sources are present, such as transient EMG episode [18].

Besides ICA and SOBI, it has already been shown that it is also feasible to apply CCA as a BSS tool for separating EMG artifacts from EEG [13]. In such an approach, CCA solves the BSS problem by forcing the sources to be maximally autocorrelated and mutually uncorrelated. Due to the broad frequency spectrum of EMG artifacts, they resemble temporal white noise, thus having lower autocorrelations compared to cerebral EEG signals [19]. In this paper, a new CCA-based method for removing muscular artifact in EEG is proposed. Similar to SOBI, multiple time lags are introduced into CCA and BSS is achieved by maximizing the sum of canonical correlations over multiple datasets at different time lags. It will be shown theoretically that the proposed use of CCA on multiple datasets at multiple time lags achieves better BSS under a more relaxed condition and hence offers better performance in removing muscular artifacts than the conventional CCA.

II. CONVENTIONAL CCA

CCA is a transfer method involving linear transformation of two sets of random vectors (denoted by $\mathbf{x}_1 \in \mathbb{R}^{d_1}$ and $\mathbf{x}_2 \in \mathbb{R}^{d_2}$) onto a joint subspace [20]. This concept was further extended to solve the BSS problem by considering the source signal and a temporally delayed version of the source signal as the two vectors [21]. This multivariate statistical method exploits the fact that muscle artifacts have low autocorrelation compared to the EEG signals. The aim would be to determine two sets of basis vectors $\omega_1 \in \mathbb{R}^{d_1}$ and $\omega_2 \in \mathbb{R}^{d_2}$, for $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively so as to maximize the canonical correlation coefficient ρ between the linear combination $\mathbf{z}_1 = \omega_1^T \mathbf{x}_1$ and $\mathbf{z}_2 = \omega_2^T \mathbf{x}_2$, as shown in (1). The thus determined weight vectors, ω_1 and ω_2 are used to calculate the first pair of canonical variates $\mathbf{z}_1 = \omega_1^T \mathbf{x}_1$ and $\mathbf{z}_2 = \omega_2^T \mathbf{x}_2$. Similarly, consequent pairs of transformation vectors are found such that they are maximally correlated and also, uncorrelated to the previously determined set of variates. This process is repeated to find all such vectors, with the total number restricted to minimum dimension of $\mathbf{x}_1$ and $\mathbf{x}_2$.

$$\max\ \rho(\mathbf{x}_1, \mathbf{x}_2)_{\omega_1, \omega_2} = \frac{\omega_1^T C_{x_1x_2} \omega_2}{\sqrt{(\omega_1^T C_{x_1x_1} \omega_1)(\omega_2^T C_{x_2x_2} \omega_2)}} \tag{1}$$

where $C_{x_1x_1}$ and $C_{x_2x_2}$ are the autocovariance matrices and $C_{x_1x_2}$ is the cross covariance matrix of $\mathbf{x}_1$ and $\mathbf{x}_2$. Normalizing the vectors ω_1 and ω_2 by letting $\omega_1^T C_{x_1x_1} \omega_1 = 1$ and $\omega_2^T C_{x_2x_2} \omega_2 = 1$, we observe that it has been reduced to a generalized eigenvalue decomposition problem:

$$\begin{cases} C_{x_1x_1}^{-1} C_{x_1x_2} C_{x_2x_2}^{-1} C_{x_2x_1} \omega_1 = \rho^2 \omega_1 \\ C_{x_2x_2}^{-1} C_{x_2x_1} C_{x_1x_1}^{-1} C_{x_1x_2} \omega_2 = \rho^2 \omega_2 \end{cases} \tag{2}$$

Thus BSS-CCA finds sources which are uncorrelated with each other, maximally autocorrelated and ordered by decreasing autocorrelation index. The above described method of CCA is constrained by a few assumptions. It extends only to data sets which have a linear relationship. Furthermore it is worthy to note it has been demonstrated that under more relaxed separability conditions, improved source separation can be achieved with increasing number of data sets.

III. MULTIPLE DATASET CCA

A. Algorithm

Multiple dataset CCA (MCCA) can be viewed as a generalized extension of the conventional canonical correlation analysis. MCCA involves mapping a linear relationship between several sets of variables to achieve maximum overall correlation. Given M datasets $\mathbf{x}_m \in \mathbb{R}^{d_m}$, m=1…M, the aim of this method would be to determine a set of M vectors ω_m and compute pairwise correlation coefficients for each pair of the variable $\mathbf{z}_m = \omega_m^T \mathbf{x}_m$, such that the sum of all the pairwise correlation is highest.

The MCCA algorithm cannot be solved as a simple eigenvalue decomposition problem as in the case of CCA, due to high computational costs. It works in multiple stages, with a linear combination being found for each random vector in each stage such that correlation among the group of resulting variates, i.e. the canonical variates, is maximized.

The optimization of overall correlation among the canonical variates can be realized by the following objective functions [22]: SUMCORR, SSQCOR, MAXVAR, MINVAR, and GENVAR. Of these, MAXVAR and MINVAR result in direct solutions while the other three follow an iterative process implying that the initial condition has an impact on the final solution.

B. Source Separability Conditions for conventional CCA

The source separability condition can be investigated by assuming the following generative model [23]:

1) Consider a group of M datasets $\{\mathbf{x}_m \in \mathbb{R}^{d_m}\} | m=1 \ldots M$, containing linear mixture of K sources $\mathbf{s}_m \in \mathbb{R}^K$ where $K \le M$, mixed by a nonsingular mixing matrix A_m, i.e., $\mathbf{x_m} = \mathbf{A_m s_m}$.

2) Assume that sources $\mathbf{s}_m$ are uncorrelated within each dataset and have zero mean and unit variance, i.e., $E\{\mathbf{s_m}\} = 0, m = 1, \cdots, M$ and $E\{\mathbf{s}_m \mathbf{s}_m^{\mathbf{T}}\} = I, m = 1, 2, \cdots M$ where I is the identity matrix.

3) Assume that sources from any pair of datasets $m, n \in \{1, 2, \cdots, M\}$; $m \ne n$ have nonzero correlation only on their corresponding indices. Without loss of generality, it is further assumed that corresponding sources have correlation in non-decreasing order in magnitude, i.e., $\left|r_{m,n}^{(1)}\right| \ge \left|r_{m,n}^{(2)}\right|, \cdots, \ge \left|r_{m,n}^{(K)}\right|$, where $\left|r_{m,n}^{(k)}\right| = E\{\mathbf{s}_m^{(k)} \mathbf{s}_n^{(k)}\}$, and $^{(k)}$ represents the kth element of random vector $\mathbf{s}_m$.

Let's consider the case of two datasets, i.e. M=2. As explained earlier, the demixing vectors ω_1 and ω_2 are obtained by maximizing the correlation coefficient ρ between the two sources $\mathbf{z}_1 = \omega_1^T \mathbf{x}_1$ and $\mathbf{z}_2 = \omega_2^T \mathbf{x}_2$. The first pair of canonical variates $\mathbf{z}_1 = \omega_1^{\mathrm{T}} \mathbf{x}_1$ and $\mathbf{z}_2 = \omega_2^{\mathrm{T}} \mathbf{x}_2$ are then jointly extracted by ω_1 and ω_2. The source separability condition is given by:

$$| r_{1,2}^{(1)} | \ne | r_{1,2}^{(2)} | \tag{3}$$

To see this, let's express $\mathbf{x}_1^{(1)}$ and $\mathbf{x}_2^{(1)}$ in terms of linear mixtures of sources $\mathbf{s}_1$ and $\mathbf{s}_2$, respectively, i.e., $\mathbf{x}_1^{(1)} = \sum_{k=1}^K \mathbf{a}^{(k)} \mathbf{s}_1^{(k)}$ and $\mathbf{x}_2^{(1)} = \sum_{k=1}^K \mathbf{b}^{(k)} \mathbf{s}_2^{(k)}$, where $\mathbf{a}$ and $\mathbf{b}$ are vectors of mixing coefficients (they are corresponding to the first row of $\mathbf{A}_1$ and $\mathbf{A}_2$, respectively). Assuming that $\mathbf{x}_1^{(1)}$ and $\mathbf{x}_2^{(1)}$ have unit variance, the correlation between them can be rewritten as $r_{1,2} = \sum_{k=1}^K \mathbf{a}^{(k)} \mathbf{b}^{(k)} r_{1,2}^{(k)}$.

Due to the triangle inequality rule, we have

$$| \tilde{r}_{1,2} | \le \sum_{k=1}^{K} | \mathbf{a}^{(k)} || \mathbf{b}^{(k)} || r_{1,2}^{(k)} |. \tag{4}$$

The RHS of the above can be rewritten as an inner product of two vectors ζ_1 and ζ_2, where

$$\zeta_1 = \left| \| \mathbf{a}^{(1)} | \sqrt{| r_{1,2}^{(1)} |}, | \mathbf{a}^{(2)} | \sqrt{| r_{1,2}^{(2)} |}, \cdots, | \mathbf{a}^{(K)} | \sqrt{| r_{1,2}^{(K)} |} \right|^T, \quad (5)$$

and

$$\zeta_2 = \left| \| \mathbf{b}^{(1)} | \sqrt{| r_{1,2}^{(1)} |}, | \mathbf{b}^{(2)} | \sqrt{| r_{1,2}^{(2)} |}, \cdots, | \mathbf{b}^{(K)} | \sqrt{| r_{1,2}^{(K)} |} \right|^T. \quad (6)$$

Note that vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ have unit variance in the generative model. As a result, we then have

$$\sum_{k=1}^{K} | \mathbf{a}^{(k)} |^2 = 1, \quad (7)$$

and

$$\sum_{k=1}^{K} | \mathbf{b}^{(k)} |^2 = 1. \quad (8)$$

Therefore, vectors ζ_1 and ζ_2 are confined on a K-dimensional hyper-ellipsoid with semi-axes

$$\left\{ \sqrt{| r_{1,2}^{(1)} |}, \sqrt{| r_{1,2}^{(2)} |}, \cdots, \sqrt{| r_{1,2}^{(K)} |} \right\}. \quad (9)$$

According to the generative model and the condition specified in (3), we have

$$| r_{1,2}{}^{(1)} | > | r_{1,2}{}^{(k)} |, \forall k > 1. \quad (10)$$

Hence,

$$< \zeta_1, \zeta_2 > \leq \| \zeta_1 \| \| \zeta_2 \| \leq | r_{1,2}{}^{(1)} |, \quad (11)$$

where the equality holds and the inner product is maximized if the angle between ζ_1 and ζ_2 is zero , i.e., ζ_1 and ζ_2 are parallel and pointing to the same direction, and $\| \zeta_1 \| = \| \zeta_2 \| = \sqrt{| r_{1,2}{}^{(1)} |}$. In other words, $| \tilde{r}_{1,2} |$ achieves its maximum when

$$| \mathbf{a}^{(k)} | = | \mathbf{b}^{(k)} | = \begin{cases} 1, & \text{if } k = 1 \\ 0, & \text{if } k \neq 1 \end{cases}. \quad (12)$$

Correspondingly, we have

$$| \zeta_1^{(1)} | = | \mathbf{s}_1^{(1)} |, \quad (13)$$

$$| \zeta_2^{(1)} | = | \mathbf{s}_2^{(1)} |. \quad (14)$$

Therefore, if two demixing vectors ω_1 and ω_2 can be found to maximize the correlation between the extracted sources from each dataset, the extracted sources are certainly the first pair of corresponding "true" sources, up to a phase ambiguity.

Now, suppose that the first pair of sources are extracted and removed from $\mathbf{x}_1$ and $\mathbf{x}_2$. Similarly proceeding, the kth pair of sources can be jointly recovered from $\mathbf{x}_1$ and $\mathbf{x}_2$ if

$$| r_{1,2}{}^{(k)} | \neq | r_{1,2}{}^{(k+1)} |. \quad (15)$$

In summary, the separability condition for CCA to recover all the K pairs of sources is as follows:

$$\left| r_{1,2}^{(k)} \right| \neq \left| r_{1,2}^{(l)} \right|, \quad 1 \leq k < l \leq K. \quad (16)$$

C. MCCA Achieves Better BSS Under More Relaxed Conditions

We know that MCCA operates in multiple stages, extracting at each stage, sources from dataset pairs while maximizing their correlation. This multistage deflationary process aims at unifying the pairwise correlations into an overall correlation measure, which achieves a maximum only when each of the pairwise correlations is maximized. Bearing this in mind, the source separability condition of joint BSS for two datasets can be extended to M datasets:

$$\forall m, n \in \{1, 2, \ldots, M\}, \quad \left| r_{m,n}^{(k)} \right| \neq \left| r_{m,n}^{(l)} \right|, \quad 1 \leq k < l \leq K \quad (17)$$

However, the above condition can be further relaxed without compromising its BSS capability. To corroborate this, consider that the first source in all M datasets has to be extracted. Following the separability condition to the mth and nth datasets, we have that the first source in these two datasets are jointly extracted if

$$| r_{m,n}{}^{(1)} | \neq | r_{m,n}{}^{(2)} |. \quad (18)$$

Therefore, to guarantee that the first source in all M datasets are extracted, we just need to find, for each $m = 1, 2, \cdots, M$, an index $n \in \{1, 2, \cdots, M\}$ such that (18) is satisfied. In other words,

$$\forall m \in \{1, \cdots, M\}, \exists n \neq m \text{ such that } | r_{m,n}^{(1)} | \neq | r_{m,n}^{(2)} |. \quad (19)$$

Extending this condition to K sources, we have the separability condition for multiple dataset CCA: $\forall m \in \{1, \cdots, M\}$, $\exists n \neq m$ such that $| r_{m,n}{}^{(k)} | \neq | r_{m,n}{}^{(l)} |$, subject to $1 \leq k < l \leq K$. Details of the proof can be found in the work by Li et al. [23].

This is a much more relaxed condition compared to (17). In other words, joint BSS on a larger group of datasets is easier to be achieved than on a smaller group of datasets. This provides the theoretical basis for the proposed idea of using multiple datasets CCA to remove muscle artifacts.

D. Multiple Time-Lag Canonical Correlation Analysis for Removing Muscular Artifacts in EEG

Motivated by the theoretical advantage of multiple dataset CCA over conventional two dataset CCA, we propose to use multiple time-lag CCA for removing muscular artifacts in EEG. For a given EEG epoch, multiple datasets are generated by introducing different time lags to the original EEG epoch. The procedure is shown in Fig.1.

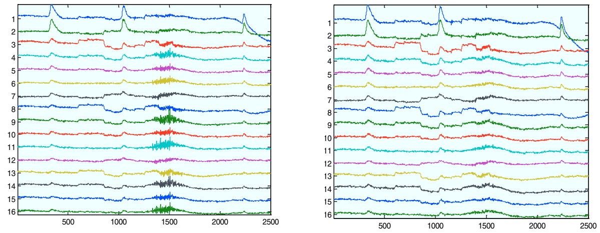
Fig.1. Implementation steps for Multiple Time-Lag CCA.

E. Experiments

The proposed method was tested on a large EEG database. The data were self-collected with control of several of muscle artifacts (caused by various facial expressions, jaw clenching, eyebrow raising, etc.).

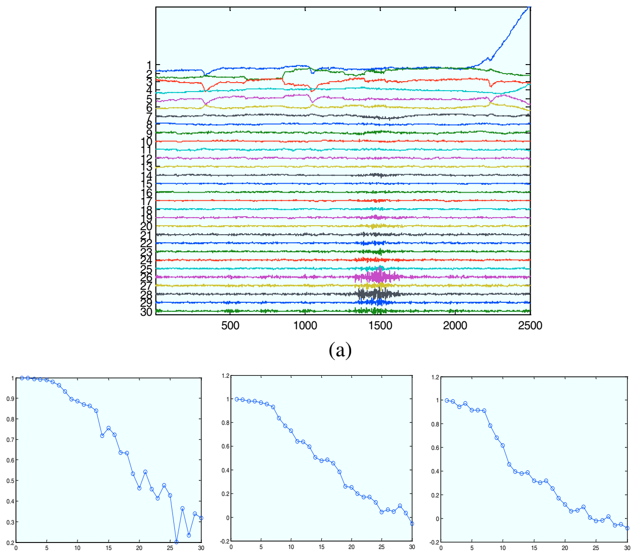
To demonstrate the effectiveness of the proposed method, Fig. 2a shows a 10s epoch of a scalp EEG recording with sampling frequency of 250 Hz. The muscle artifact was caused by transient facial muscle movement. The corresponding reconstructed artifact-free EEG epoch is shown in Fig. 2b.

Fig. 3a shows the sources that the proposed multiple time lag canonical correlation analysis decomposed. It is clear that muscle artifacts are relatively well contained in the last few components, while cerebral EEG components and eye blinking components appeared in the first few components.



(a) (b)

Fig.2. (a) Sample EEG epoch with muscular artifacts, (b) reconstructed artifact-free EEG epoch after muscular artifact removal using the multiple time lag CCA.



(a)

(b) (c) (d)

Fig.3. (a) Blind Source Separation results by the proposed method (Number of time lags = 10). Cerebral EEG components appear in the first several components, while EMG artifacts appear in the last several components. The autocorrelation of the separated sources for different number of time lags used: (b) time lag =1, (c) time lag = 5, (d) time lag =10. It appears that it provides better separation of muscular artifacts from cerebral EEG components with increasing number of time lags.

The advantage of introducing multiple time lags in canonical correlation analysis can be seen in Figs. 3b, 3c, and 3d, where different numbers of time lags were used in the proposed method. It can be seen that joint BSS on a larger group of datasets (for a larger number of time lags used) is easier to be achieved than on a smaller group of datasets (for a smaller number of time lags used).

IV. Discussion and Conclusion

A new method to remove muscular artifacts from EEG is proposed. The method is based on joint blind source separation scheme by using Multiple Dataset Canonical Correlation Analysis. Theoretical study shows that the proposed joint blind source separation method can achieve source recovery under a more relaxed separability condition. In other words, the proposed method is easier to achieve blind source separation than conventional method on a smaller group of datasets. The experiment shows promising results.

References

[1] S.-Y. Shao, K.-Q. Shen, C.-J. Ong, E. P. V. Wilder-Smith, and X.-P. Li, "Automatic EEG Artifact Removal: a Weighted Support-Vector-Machine Approach With Error Correction," IEEE Trans Biomed Eng, vol. 56, pp. 336 - 344, 2008.

[2] P. Common, "Independent component analysis - A new concept?," Signal processing, vol. 36, pp. 287-314, 1994.

[3] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," IEEE transactions on neural networks, vol. 10, pp. 626-634, 1999.

[4] A. Hyvarinen, "Independent component analysis: algorithms and applications," neural networks, vol. 13, pp. 411-430, 2000.

[5] E. Urrestarazu, J. Iriarte, M. Alegre, M. Valencia, C. e. Viteri, and J. Artieda, "Independent component analysis removing artifacts in ictal recordings," Epilepsia, vol. 45, pp. 1071-1078, 2004.

[6] T.-P. Jung, C. Humphries, T.-W. Lee, S. Makeig, M. J. McKeown, V. Iragui, et al., "Removing electroencephalographic artifacts: comparison between ICA and PCA," in Proc. IEEE Signal Processing Society Workshop Neural Networks for Signal Processing VIII, ed, 1998, pp. 63-72.

[7] T. P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. J. McKeown, V. Iragui, et al., "Removing electroencephalographic artifacts by blind source separation.," Psychophysiology, vol. 37, pp. 163-178, 2000.

[8] G. L. Wallstrom, R. E. Kass, A. Miller, J. F. Cohn, and N. A. Fox, "Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods.," Int J Psychophysiol, vol. 53, pp. 105-119, 2004.

[9] N. P. Castellanos and V. A. Makarov, "Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis.," J Neurosci Methods, vol. 158, pp. 300-312, 2006.

[10] S. Makeig and T. Jung, "Tonic phasic, and transient EEG correlates of auditory awareness in drowsiness," cognitive brain research, vol. 4, pp. 15-25, 1996.

[11] R. N. Vig\'ario, "Extraction of ocular artefacts from EEG using independent component analysis," Electroencephalography And Clinical Neurophysiology, vol. 103, pp. 395-404, 1997.

[12] R. Vig\'ario, J. S\"arel\"a, V. Jousm\"aki, M. H\"am\"al\"ainen, and E. Oja, "Independent component approach to the analysis of EEG and MEG recordings.," IEEE Trans Biomed Eng, vol. 47, pp. 589-593, 2000.

[13] W. D. Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical Correlation Analysis Applied to Remove Muscle Artifacts From the Electroencephalogram," IEEE Transactions on Biomedical Engineering, vol. 53, pp. 2583-2587, 2006.

[14] B. W. McMenamin, A. J. Shackman, J. S. Maxwell, D. R. W. Bachhuber, A. M. Koppenhaver, L. L. Greischar, et al., "Validation of ICA-based myogenic artifact correction for scalp and source-localized EEG.," Neuroimage, vol. 49, pp. 2416-2432, Feb 2010.

[15] B. W. McMenamin, A. J. Shackman, J. S. Maxwell, L. L. Greischar, and R. J. Davidson, "Validation of regression-based myogenic correction techniques for scalp and source-localized EEG.," Psychophysiology, vol. 46, pp. 578-592, May 2009.

[16] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," Signal Processing, IEEE Transactions on, vol. 45, pp. 434-444, 1997.

[17] A. Ziehe, K.-R. Muller, G. Nolte, B.-M. Mackert, and G. Curio, "Artifact reduction in magnetoneurography based on time-delayed second-order correlations," IEEE Trans Biomed Eng, vol. 47, pp. 75-87, 2000.

[18] S. Choi, A. Cichocki, and A. Beloucharni, "Second Order Nonstationary Source Separation," The Journal of VLSI Signal Processing, vol. 32, pp. 93-104, 2002.

[19] I. I. Goncharova, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "EMG contamination of EEG: spectral and topographical characteristics.," Clin Neurophysiol, vol. 114, pp. 1580-1593, Sep 2003.

[20] H. Hotelling, "Relations Between Two Sets of Variates," Biometrika, vol. 28, pp. pp. 321-377, 1936.

[21] O. Friman, M. Borga, P. Lundberg, and H. Knutsson, "Exploratory fMRI analysis by autocorrelation maximization.," Neuroimage, vol. 16, pp. 454-464, Jun 2002.

[22] J. R. Kettenring, "Canonical Analysis of Several Sets of Variables," Biometrika, vol. 58, pp. pp. 433-451, 1971.

[23] Y. O. Li, T. Adali, W. Wang, and V. D. Calhoun, "Joint Blind Source Separation by Multi-set Canonical Correlation Analysis," IEEE Trans Signal Process, vol. 57, pp. 3918-3929, Oct 1 2009.