Loughborough University

This item was submitted to Loughborough's Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/

Proceedings of the 29th Annual International
Conference of the IEEE EMBS
Cité Internationale, Lyon, France
August 23-26, 2007.

SuA02.3

A Robust Minimum Variance Beamforming Approach for the Removal of the Eye-Blink Artifacts from EEGs

Kianoush Nazarpour, Yodchanan Wongsawat, Saeid Sanei, Soontorn Oraintara, and Jonathon A. Chambers

Abstract— In this paper a novel scheme for the removal of eye-blink (EB) artifacts from electroencephalogram (EEG) signals based on the robust minimum variance beamformer (RMVB) is proposed. In this method, in order to remove the artifact, the RMVB is provided with *a priori* information, i.e., an estimation of the steering vector corresponding to the *point source EB artifact*. The artifact-removed EEGs are subsequently reconstructed by deflation. The *a priori* knowledge, namely the vector corresponding to the spatial distribution of the EB factor, is identified using a novel space-time-frequency-time/segment (STF-TS) model of EEGs, provided by a four-way parallel factor analysis (PARAFAC) approach. The results demonstrate that the proposed algorithm effectively identifies and removes the EB artifact from raw EEG measurements.

I. Introduction

Electroencephalogram (EEG) is the manifestation of brain activity recorded as changes in electrical potentials at multiple locations over the scalp. The electrooculogram (EOG) generated by eye movements or blinks is found to be the most significant and common artifact in EEG [1]. The EOG is of the order of ten times larger in amplitude than average cortical signals and lasts approximately 300 msec. Due to the reasonably high magnitude of the blinking artifacts and the high resistance of the skull and scalp tissues, EOG may contaminate the majority of the electrode signals, even those in the occipital area.

In recent years, various methods for EB artifact removal from EEGs have been proposed which are mainly based on linear regression [2] and independent component analysis (ICA) [1]. Approaches such as trial rejection, eye fixation, EOG subtraction, principal component analysis (PCA) [3], blind source separation (BSS) using ICA [4], [5], and H^∞ [6] adaptive filters have also been documented as having varying success. Despite no quantitative comparison for any reference dataset being available, it has been shown that the regression- and BSS-based methods are the most reliable ones [1], [2], [4], and [5]. Although beamforming-based methods have been recently utilized in the EEG source localization problem [7], to the authors' best knowledge, they have not been considered in removing the EB artifacts from the EEGs. This is understandable since these schemes suffer a significant performance degradation when the array response vector for the source of interest, i.e. EB, is not exactly known [8].

Statistically nonstationary EEGs yield information about active parts of the brain. This spatial knowledge has been efficiently exploited for localizing the sources of background EEG using PCA. However, although by using ICA, the inherent nonuniqueness problem of PCA is resolved, the statistical independence constraint is imposed. Moreover, in conventional PCA/ICA, no other prior knowledge, such as frequency band, is exploited during EEG analysis. A topographic time-frequency decomposition method is proposed in [9] and followed by [10] where the space-time-frequency (STF) model of multi-channel EEGs is introduced. More recently, we have utilized the STF model for the identification and removal of EB artifacts and brain computer interfacing [1], [5], [11], and [12]. Although, STF modeling is effective, it suffers from high computational complexity when applied to long term data sequences recorded from a high number of electrodes [12].

In this paper, a novel technique for removing the EOG artifacts from multi-channel EEGs is presented. Our method is based on the robust minimum variance beamformer [13], where the spatial knowledge of the mixing process, obtained by PARAFAC analysis, is exploited as an estimation of the steering vector corresponding to the EB source. Aiming at reducing the computational complexity in the estimation of the STF model using PARAFAC, the time domain is subdivided into a number of segments and a four-way array is then set to estimate the space-time-frequency-time/segment (STF-TS) model of the data using the four-way PARAFAC. Subsequently, the STF-TS model results in the classic STF model, with significantly lower computational cost.

This paper is organized as follows. In Section II, we briefly review the RMVB and introduce the spatial signature of the STF-TS model as an estimation of the array response vector following by the proposed STF-TS based STF model estimation methodology. The results are subsequently reported in Section III, followed by concluding remarks in Section IV.

II. Algorithm Development

Assume N zero-mean real mutually uncorrelated point *geometrically stationary* sources $\mathbf{s}(t) = [s_1(t), s_2(t), \cdots, s_N(t)]'$, where $[\cdot]'$ denotes the vector transpose, are mixed by an $N \times N$ full column rank matrix $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_N]$ where $\mathbf{a}_i$ is the ith column of $\mathbf{A}$. The vector of time mixture samples $\mathbf{x}(t) = [x_1(t), x_2(t), \cdots, x_N(t)]'$ is given as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{v}(t) \qquad (1)$$

where $\mathbf{v}(t) = [v_1(t), v_2(t), \cdots, v_N(t)]'$ is the additive white Gaussian zero-mean noise which is assumed to be spatially uncorrelated with the sensor data and temporally uncorrelated. The sources are presumed to be uncorrelated, therefore, the time lagged symmetrized autocorrelation matrix $\mathbf{R}_{\mathbf{xx}}^k$ can be calculated as $\mathbf{R}_{\mathbf{xx}}^k = E[\mathbf{x}(t)\mathbf{x}'(t - \tau_k)]$ for $k = 1, 2, \cdots, K$, where K is the maximum number of time lags τ_K and $E[\cdot]$ denotes the statistical expectation operator. The vector $\mathbf{x}(t)$ in (1) is a linear combination of the columns of the mixing matrix, i.e., the $\mathbf{a}_i$s, weighted by the associated source and contaminated by the noise $\mathbf{v}(t)$.

A. Robust Minimum Variance Beamformer

The most straightforward way to extract the jth source is to project $\mathbf{x}(t)$ onto the space orthogonal to, denoted by $\perp$, all of the columns of $\mathbf{A}$ except $\mathbf{a}_j$, i.e., $\{\mathbf{a}_1, \cdots, \mathbf{a}_{j-1}, \mathbf{a}_{j+1}, \cdots, \mathbf{a}_N\}$. Since $\mathbf{a}_j$ performs as the steering vector of the jth source, by defining a vector, *a spatial filter*, $\mathbf{w}_j$, we may write $y(t) = \mathbf{w}_j'\mathbf{x}(t)$ where $y(t)$ is an estimation of the source $s_j(t)$ corresponding to $\mathbf{a}_j$. The spatial filter can be determined by applying the unit-gain constraint, $\mathbf{w}_j'\mathbf{a}_j = 1$ and by minimizing the variance of the filter output, i.e., $y(t)$ [7]. However, in practice, the steering vector $\mathbf{a}_j$ is not always known. Hence, theoretically rigorous worst-case performance optimization-based approaches have been recently developed in [8] in order to compensate the deviation vector, δ of the $\hat{\mathbf{a}}_j$ from the actual steering vector $\mathbf{a}_j$, i.e., $\delta = \mathbf{a}_j - \hat{\mathbf{a}}_j$. Note that δ is l_2-, denoted by $\|.\|$, norm-bounded by some known constant ε. As outlined in [13], the beamformer is obtained by minimizing $J_c = \sum_{k=1}^{K} \mathbf{w}_j'\mathbf{R}_{\mathbf{xx}}^k\mathbf{w}_j$ subject to $\min_{\|\delta\| \le \varepsilon} |\mathbf{w}_j'\hat{\mathbf{a}}_j + \mathbf{w}_j'\delta| = 1$ where $|\varsigma|$ denotes the absolute value of ς. Equivalently [13], we may rewrite J_c as $J_c = \sum_{k=1}^{K} \mathbf{w}_j'\mathbf{R}_{\mathbf{xx}}^k\mathbf{w}_j$ s.t. $|\mathbf{w}_j'\hat{\mathbf{a}}_j - \varepsilon\|\mathbf{w}_j\|| = 1$. Using the Lagrange multiplier method, differentiating J_c with respect to $\mathbf{w}_j$, and setting to zero, we have $\mathbf{R}\mathbf{w}_j + \lambda\varepsilon\frac{\mathbf{w}_j}{\|\mathbf{w}_j\|} = \lambda\hat{\mathbf{a}}_j$ where $\mathbf{R} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{R}_{\mathbf{xx}}^k$. After dropping the unimportant constant λ, and considering the reasoning provided in [13], the spatial filter can be computed using

$$\mathbf{w}_j = \left[\mathbf{R} + \frac{\varepsilon}{\rho}\mathbf{I}\right]^{-1}\hat{\mathbf{a}}_j \quad (2)$$

where $\rho \triangleq \|\mathbf{w}_j\|$ and $\mathbf{I}$ denotes the identity matrix. In (2), the main concern in estimating $\mathbf{w}_j$ is to have an estimation of ρ which may be determined by using the following procedure.

Eigenvalue decomposition of $\mathbf{R}$, i.e., $\mathbf{R} = \mathbf{U}\Xi\mathbf{U}'$ results in the $N \times N$ unitary matrix $\mathbf{U}$ whose columns are the eigenvectors of $\mathbf{R}$, and Ξ, the diagonal matrix of the real positive eigenvalues of $\mathbf{R}$, i.e., $\xi_1 \ge \xi_2 \ge \cdots \ge \xi_N > 0$. By defining $\Psi(\rho) \triangleq \Xi + \frac{\varepsilon}{\rho}\mathbf{I}$ and following the procedure suggested in [13], we may write

$$\|\mathbf{U}\Psi^{-1}(\rho)\mathbf{U}'\hat{\mathbf{a}}_j\|^2 - \rho^2 = \|\Psi^{-1}(\rho)\mathbf{g}'\| - \rho^2 = 0 \quad (3)$$

where $\mathbf{g} = [g_1, g_2, \cdots, g_M]' = \mathbf{U}'\hat{\mathbf{a}}_j$. Introducing

$$f(\rho) \triangleq \|\Psi^{-1}(\rho)\mathbf{g}'\| - \rho^2 = \sum_{i=1}^{M}\left[\frac{|g_i|}{\varepsilon + \rho\xi_i}\right]^2 - 1 = 0, \quad (4)$$

in [13], it is shown than that the necessary and sufficient condition for (4) to have a unique real positive solution for ρ is that the norm of the mismatch vector is upper bounded by the norm of estimated signal steering vector, i.e., $\|\delta\| = \varepsilon < \|\hat{\mathbf{a}}_j\|$. Considering $\|\mathbf{g}\| = \|\hat{\mathbf{a}}_j\|$ and (4), the upper bound of $f(\rho)$ is achieved as

$$f(\rho) < \frac{\sum_{i=1}^{M}|g_i|^2}{(\varepsilon + \rho\xi_M)^2} - 1 = \frac{\|\hat{\mathbf{a}}_j\|^2}{(\varepsilon + \rho\xi_M)^2} - 1 \triangleq f_{max}(\rho). \quad (5)$$

Note that $f(\rho)$ and $f_{max}(\rho)$ are both decreasing functions of ρ and the root of $f(\rho)$, say ρ_0, is positive. Hence, we have $0 < \rho_0 < \rho_{max} = \frac{\|\hat{\mathbf{a}}_j\| - \varepsilon}{\xi_M}$ [13]. Therefore, the problem of estimating ρ and consequently the spatial filter, $\mathbf{w}_j$, can be solved within an iterative scheme as in [13].

B. STF and SFT-TS Modeling

In this work, by exploiting PARAFAC, we extract the factor relevant to the EB artifact to be used within the beamforming procedure. The resulting spatial signature of the EB-related factor is exploited to formulate (2). Importantly, we have considered that the spatial signature of this factor is directly related to the level of EB contamination for each electrode. This assumption is rational since the EB can be considered as a strong point source which is just attenuated while propagating from the frontal area to the central and occipital parts of the brain. Hence, the column of the mixing matrix $\mathbf{A}$, i.e., $\hat{\mathbf{a}}_j$ corresponding to the EB source, is estimated by PARAFAC and used in (2).

Here, we briefly review the adopted approach from [5] and [12]. Note that the MATLAB matrix notation has been utilized. In order to decompose the EEGs into spatial, temporal, and spectral signatures, the three-way PARAFAC is applied to the three-way EEG data $\underline{\check{\mathbf{Y}}}^{N \times T \times F} \triangleq \underline{\check{\mathbf{Y}}}(1 : N, 1 : T, 1 : F)$ where N, T, and F are respectively the number of EEG channels, time instants, and frequency bins. Therefore, as in sequel, $\check{\mathbf{A}}^{N \times M}$, $\check{\mathbf{C}}^{F \times M}$, and $\check{\mathbf{D}}^{T \times M}$ are respectively the spatial, spectral, and temporal signatures of $\underline{\check{\mathbf{Y}}}^{N \times F \times T}$ where their elements are denoted as $\check{a}(t, m)$, $\check{c}(n, m)$, and $\check{d}(f, m)$. While retaining the consistency of formulation, we occasionally drop the superscripts to simplify the presentation.

The STF model is presented as $\underline{\check{\mathbf{Y}}}^{N \times F \times T} = \underline{\hat{\check{\mathbf{Y}}}} + \underline{\check{\mathbf{E}}}^{N \times F \times T}$ where $\underline{\hat{\check{\mathbf{Y}}}} = \sum_{m=1}^{M}\check{a}(n, m)\check{c}(f, m)\check{d}(t, m)$ is an estimation, denoted by $\hat{}$, of $\underline{\check{\mathbf{Y}}}$, M is the maximum possible number of factors, and $\underline{\check{\mathbf{E}}}$ is the three-way array of the residue of the model which is mostly omitted for brevity.

In order to find M, we utilize the known core consistency diagnostic (CORCONDIA) measure [14]. The signatures $\check{\mathbf{A}}$, $\check{\mathbf{C}}$, and $\check{\mathbf{D}}$ can be estimated by using the alternating least squares (ALS) algorithm where the cost function is $[\hat{\check{\mathbf{A}}}, \hat{\check{\mathbf{C}}}, \hat{\check{\mathbf{D}}}] = \arg\min_{\check{a}, \check{c}, \check{d}} \|\underline{\check{\mathbf{Y}}} - \underline{\hat{\check{\mathbf{Y}}}}\|^2$.

Intuitively, the spatial signature $\check{\mathbf{A}}$ obtained from the STF model represents the weighting parameters of the inter-channel correlation among time-frequency representations of each channel. However, in order to surpass the high computational cost occurring in using STF with three-way PARAFAC [14], in the sequel, we introduce a novel method

6212

for estimating the STF model. The strategy is based on the *divide and conquer* philosophy where as will be detailed in sequel, instead of calculating the model signatures from the original data, we estimate these signatures by joining the weighted versions of their local temporal signatures.

For long-term EEG measurements, the calculations of both the time-frequency transform and STF modeling are computationally intensive. Therefore, aiming to reduce this computational complexity, we divide the time domain into a number of segments. Thereafter, the time-frequency transform is applied [5] individually to each segment forming a four-way array. We set up the four-way array $\underline{\mathbf{Y}}^{N \times S \times F_s \times T_s} \triangleq \underline{\mathbf{Y}}(1:N, 1:S, 1:F_s, 1:T_s)$ where N is the channel index and S is the maximum time/segment index, and compute the energy of the time-frequency transform for T_s time instants and F_s frequency bins. The four-way PARAFAC model of the four-way array $\underline{\mathbf{Y}}$ may be formulated the same way as in [14] where $\mathscr{A}^{N \times M}$ is the spatial signature, $\mathscr{B}^{S \times M}$ is the temporal/segment signature, $\mathscr{C}^{F_s \times M}$ is the spectral signature, and $\mathscr{D}^{T_s \times M}$ is the temporal signature with matrix elements denoted respectively as $a(n,m)$, $b(s,m)$, $c(f_s,m)$, and $d(t_s,m)$. Hence,

$$\underline{\mathbf{Y}}^{N \times S \times F_s \times T_s} = \underline{\hat{\mathbf{Y}}} + \underline{\mathscr{E}}^{N \times S \times F_s \times T_s} \qquad (6)$$

where $\underline{\hat{\mathbf{Y}}} = \sum_{m=1}^{M} a(n,m)b(s,m)c(f_s,m)d(t_s,m)$ and $\underline{\mathscr{E}}^{N \times S \times F_s \times T_s}$ is the negligible four-way residual of the model array. In order to find the model we have used the following cost function $[\hat{\mathscr{A}}, \hat{\mathscr{B}}, \hat{\mathscr{C}}, \hat{\mathscr{D}}] = \arg\min_{a,b,c,d} \|\underline{\mathbf{Y}} - \underline{\hat{\mathbf{Y}}}\|^2$.

By decomposing the multi-channel EEGs using the STF-TS model, the number of free parameters P_4, i.e., the number of elements that has to be estimated by PARAFAC, is $M(N + S + F_s + T_s)$, while the number of free parameters of the STF model P_3 is as high as $M(N + F + T)$. Evidently, when T is large, $P_4 << P_3$. This means that less parameters need to be estimated and therefore the computational complexity of the PARAFAC algorithm is reduced. According to (6), the temporal signatures of the long-term EEGs are estimated by joining all S segments of the temporal signatures $\mathscr{D}$ which are weighted by their corresponding time/segment signatures $\mathscr{B}$. In order to effectively estimate the STF model from the STF-TS model, the suggested number of segments S and the number of components M should maximize the CORCONDIA value as

$$[S, M] = \arg\max_{S} \left\{ \max_{M} \{\text{CORCONDIA}(\underline{\mathbf{Y}}, \mathscr{A}, \mathscr{B}, \mathscr{C}, \mathscr{D})\} \right\}. \qquad (7)$$

The concept behind (7) is that by decomposing $\underline{\mathbf{Y}}$ to as many as M possible factors for the STF model, we firstly guarantee that the correct number of factors for STF-TS is achieved and then, we progress to the process of temporal segmentation. In other words, since the ultimate goal of the STF-TS model is to approximate the STF model, M should be identified for the STF model using the conventional approach of [14] before adjusting S to maximize the CORCONDIA criterion for STF-TS. When the residual is considered negligible, the STF model can be written in a matrix form as $\underline{\check{\mathbf{Y}}}_{N \times F \times T} = \check{\mathbf{D}} \Sigma_{\check{\mathbf{A}}_n} \check{\mathbf{C}}'$, where $\Sigma_{\check{\mathbf{A}}_n}$ is the diagonal matrix with the n-th row

of $\check{\mathbf{A}}$ as its diagonal elements, $n = 1, 2, \cdots, N$. Similarly, the STF-TS model (6) is written in matrix form as

$$\underline{\mathbf{Y}}_{N \times S \times F_s \times T_s} = \mathscr{D} \Sigma_{\mathscr{B}_s} \Sigma_{\mathscr{A}_n} \mathscr{C}', \qquad (8)$$

where $\Sigma_{\mathscr{A}_n}$ is a diagonal matrix with the n-th row of $\mathscr{A}$ as its diagonal elements, $n = 1, 2, \cdots, N$. Similarly, $\Sigma_{\mathscr{B}_s}$ is a diagonal matrix with the s-th row of $\mathscr{B}$ as its diagonal elements for $s = 1, 2, \cdots, S$. Hence, $\check{\mathbf{D}}$ for the STF can be estimated by the scaled version of $\mathscr{D}$ from the STF-TS as

$$\check{\mathbf{D}} \approx [\mathscr{D}\Sigma_{\mathscr{B}_1}, \cdots, \mathscr{D}\Sigma_{\mathscr{B}_S}]' \qquad (9)$$

The spectral signature $\check{\mathbf{C}}$ is well approximated by $\mathscr{C}$, while the spatial signature $\check{\mathbf{A}}$ is approximately equal to $\mathscr{A}$.

III. SIMULATION RESULTS

We applied our algorithm to real EEG measurements. The database was provided by the School of Psychology, Cardiff University, UK, and represent a wide range of EBs, i.e. more than 500 EB contaminated EEG recordings. The scalp EEG was obtained using 25 Silver/Silver-Chloride electrodes placed at locations defined by the conventional 10-20 system [1]. The data were sampled at 200 Hz, and bandpass filtered with cut-off frequencies of 1 Hz and 30 Hz. The performance of the algorithm can be observed by comparing the EEGs obtained at the electrodes in the left subplot of Fig. 1 and the same segment of data after being processed by the proposed algorithm in the right subplot of Fig. 1.

The CORCONDIA value has been computed for the methods of STF and STF-TS modeling. In Fig. 2, the number of components M is selected as $M = 2$ according to the CORCONDIA value, i.e, 84.425% whereas the CORCONDIA value for the proposed STF-TS model was 32.339% when the number of segments was $S = 10$. Figs. 3 (a) to (d) illustrate respectively the estimated spectral, temporal, and spatial signatures of the under-studied EEGs. The results of the STF-TS model in comparison to that of the STF model, i.e., Fig. 2, demonstrates the reliability of the STF-TS modeling, since both methods result in approximately the same signatures, except, as expected, the STS-TS method is a faster algorithm.

By using the STF model, we have to calculate the parallel factors of the three-way array of size $N \times F \times T$. This process takes a longer period of time due to the calculations of more free parameters P_3 compared to that of the STF-TS model P_4. The number of free parameters is greatly reduced by using the STF-TS model, where the size of the three-way $\underline{\check{\mathbf{Y}}}_{N \times F \times T}$ for the STF model is $25 \times 1800 \times 180$, i.e., 4,010 parameters to be estimated, and the size of the four-way $\underline{\mathbf{Y}}_{N \times S \times F_s \times T_s}$ for the STF-TS model is $25 \times 10 \times 180 \times 180$, i.e., 790 parameters to be estimated. Consequently, the relative calculation time of the STF-TS model, presuming that the calculation time of the STF model is 1, is 0.197.

We are only interested in the spatial signature of the EB artifact relevant factor to be used in the RMVB algorithm as an approximation to $\mathbf{a}_j$. Note that the first component (Factor 1) of both the STF and STF-TS models demonstrates
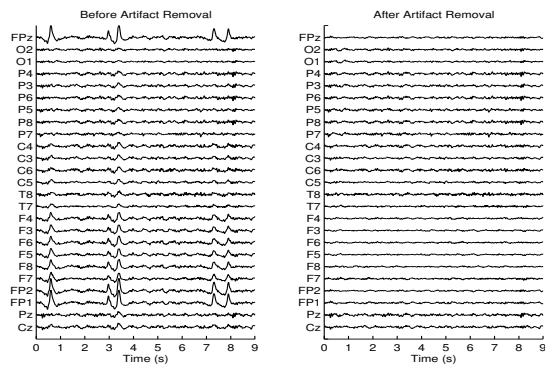
6213

Fig. 1. The results of the proposed EB artifact removal method for a set of real EEGs. The left subplot depicts highly EB contaminated EEGs before artifact removal while in the right subplot the segment of EEGs after being corrected for EB artifact is illustrated.
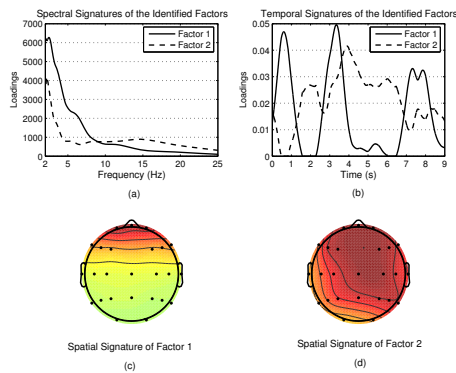

Fig. 2. The extracted factor by using STF modeling; (a) and (b) illustrate respectively the spectral and temporal signatures of the extracted factors. (c) and (d) represent the spatial distribution of the factors, respectively. Evidently, Factor 1 demonstrates the EB phenomenon since it occurs in frequency band of around 5 Hz (a), it is indeed transient in the time domain (b) and it is confined to the frontal area.

the eye-blink-relevant factor, since they mainly occur in the frequency band of around 5 Hz, while the other factors exist in the entire band and represent the ongoing activity of the brain or perhaps a broadband white noise-like component. The temporal signature of the first factor definitely shows a transient phenomenon such as eye-blink while that of Factor 2 consistently exists during the course of the EEG segment. Unlike Fig. 2- and Fig. 3-(d), in Fig. 2- and Fig. 3-(c), the spatial distribution of the extracted factor, to be used as $\hat{\mathbf{a}}_j$, is confined to the frontal area, which clearly demonstrates the effect of eye-blink. The other factor shows the background activity of the brain as it almost spreads over the scalp. Using $\hat{\mathbf{a}}_j$ in (2), we find the beamformer $\mathbf{w}_j$ and extract the EB source. The artifact removed EEGs are then reconstructed by using the batch deflation method [1].

IV. CONCLUDING REMARKS

We have presented a method for removing EOG from EEG recordings by employing the robust minimum variance beamforming method to estimate the steering vector corresponding to the EB source. The vector of spatial distribution of the EB factor has been identified using the proposed STF-TS model which enjoys much less computational complexity
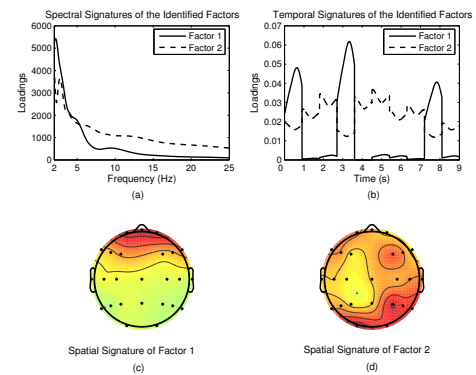

Fig. 3. The extracted factor by using STF-TS modeling; (a) and (b) illustrate respectively the spectral and temporal signatures of the extracted factors and (c) and (d) represents the spatial distributions of those extracted factors. Interestingly, as expected, the spectral and spatial signatures of the extracted components are similar to those of Fig. 2 and the temporal signatures effectively track the transient EBs of the ongoing EEGs.

in comparison to the conventional STF model [5]. For the first time in this work, we have utilized the vector of spatial signature of the EB factor as the estimation of the steering vector that introduces the EB source to the EEGs. Probable deviations of the $\hat{\mathbf{a}}_j$ from the actual steering vector $\mathbf{a}_j$ are compensated by exploiting the robustness of the RMVB.

REFERENCES

[1] S. Sanei and J. A. Chambers, *EEG Signal Processing*. Wiley, 2007.
[2] A. Schlögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller, "A fully automated correction method of EOG artifacts in EEG recordings," *Clin. Neurophysiol.*, vol. 118, pp. 98–104, 2007.
[3] M. Fatourechi, A. Bashashati, R. K. Ward, and G. E. Birch, "EMG and EOG artifacts in brain computer interface systems: A survey," *Clin. Neurophysiol.* 2007, in Press.
[4] L. Shoker, S. Sanei, and J. A. Chambers, "Artifact removal from electroencephalograms using a hybrid BSS-SVM algorithm," *IEEE Sig. Process. Lett.*, vol. 12, no. 10, pp. 721–724, 2005.
[5] K. Nazarpour, S. Sanei, and J. A. Chambers, "A novel semi-blind signal extraction approach incorporating PARAFAC for the removal of eye-blink artifact from EEGs," *IEEE Trans. Biomed. Eng.* submitted.
[6] S. Puthusserypady and T. Ratnarajah, "Robust adaptive techniques for minimization of EOG artefacts from EEG signals," *Sig. Process.*, vol. 86, pp. 2351–2363, 2006.
[7] B. D. Van Veen and W. van Drongelen and M. Yuchtman and A. Suzuki, "Localization of brain electrical activity via linearly constrained minimum variance spatial filter," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 867–879, 2000.
[8] S. A. Vorobyov, A. B. Gershman, and Z. Q. Luo, "Robust adaptive beamforming using worst case performance optimization," *IEEE Trans. Sig. Process.*, vol. 51, no. 2, pp. 313–324, 2003.
[9] T. Koenig, F. Marti-Lopez, and P. A. Valdés-Sosa, "Topographic time-frequency decomposition of the EEG," *NeuroImage*, vol. 14, pp. 383–390, 2001.
[10] F. Miwakeichi, E. Martinez-Montes, P. A. Valdés-Sosa, N. Nishiyama, H. Mizuhara, and Y. Yamaguchia, "Decomposing EEG data into space-time-frequency components using parallel factor analysis," *NeuroImage*, vol. 22, pp. 1035–1045, 2004.
[11] K. Nazarpour, S. Sanei, L. Shoker, and J. A. Chambers, "Parallel space-time-frequency decomposition of EEG signals for brain computer interfacing," in *Proc. EUSIPCO 06*, Florence, Italy, Sep. 2006.
[12] Y. Wongsawat, S. Oraintara, and K. R. Rao, "Reduced complexity space-time-frequency model for multi-channel EEG and its applications," in *Proc. IEEE ISCAS*, New Orleans, USA, 2007.
[13] J. Li and P. Stoica, *Robust Adaptive Beamforming*. Wiley, 2005.
[14] R. Bro, "PARAFAC: Tutorial and applications," *Chemometrics Intell. Lab. Syst.*, vol. 38, pp. 149–171, 1997.