A TREATMENT OF EEG DATA BY UNDERDETERMINED BLIND SOURCE SEPARATION FOR MOTOR IMAGERY CLASSIFICATION

Zbynek Koldovsky[1,2], *Anh Huy Phan*[3], *Petr Tichavsky*[2], *and Andrzej Cichocki*[3]

[1]Faculty of Mechatronic and Interdisciplinary Studies,
Technical University of Liberec, Studentská 2, 461 17 Liberec, Czech Republic.
[2]Institute of Information Theory and Automation,
Pod vodárenskou věží 4, P.O.Box 18,182 08 Prague 8, Czech Republic.
[3]Brain Science Institute, RIKEN, Wakoshi, Japan.

ABSTRACT

Brain-Computer Interfaces (BCI) controlled through imagined movements cannot work properly without a correct classification of EEG signals. The difficulty of this problem consists in low signal-to-noise ratio, because EEG may contain strong signal components that are not related to motor imagery. In this paper, these artifact components are to be suppressed using a recently proposed underdetermined blind source separation method and a novel MMSE beamformer. We use these tools to remove unwanted components of EEG to increase the classification accuracy of the BCI system. In our experiments with several datasets, the classification is improved by up to 10%.

***Index Terms*—** Underdetermined Blind Source Separation, Beamforming, Electroencephalogram, Brain-Computer Interface

1. INTRODUCTION

Some Brain-Computer Interfaces (BCI) are based on recognition of a human's intention from noninvasive measurements of brain activity such as electroencephalography (EEG) [1]. They enable to control a computer without using the peripheral nervous system of the human. We focus on BCI utilizing EEG with a small number of channels (six), because it is the most accessible measurement device.

A detection of imagined movements is possible thanks to the fact that they generate components in EEG related to the activity. The need is to isolate these components, because there are many artifacts in EEG caused by other background activities of the brain. Moreover, the signal-to-noise ratio is usually low. One way is the standard temporal filtering that passes selected frequency bands only, in which the target components show most of their energy [2, 3]. Since EEG is a multichannel signal, spatial filterings (linear combinations of channels) are possible as well. A popular method of this kind is the common spatial patterns algorithm (CSP) [4]. Combining both filtering approaches provides a powerful strategy.

However, CSP and similar supervised methods suffer from artifactual learning data. The reason is that the energy of artifacts can be greater than the energy of the target components related to motor imagery (MI). Since CSP consists in maximizing the ratio of variances of extracted EEG from two classes (e.g., left and right-hand imagined movements), it may focus on the artifacts instead.

A great potential here provide blind source separation methods (BSS), because they decompose multichannel EEG into components based on their "information content". In other words, the energy (variance) of components is not the primary criterion in BSS. For example, a popular method for BSS is the Independent Component Analysis (ICA) that computes spatial filters such that their outputs are as independent as possible [5]. Since independent components originate from independent brain activities, it is possible to isolate artifacts not related to MI [6] and apply CSP then.

In BSS, the measured EEG data are modeled as

$$\mathbf{X} = \mathbf{A}\mathbf{S} \tag{1}$$

where $\mathbf{X}$ is an $m \times N$ matrix whose rows consist of N samples of signals from m electrodes, $\mathbf{S}$ is a $d \times N$ matrix of d independent components, and $\mathbf{A}$ is an $m \times d$ mixing matrix. While $\mathbf{X}$ is known, $\mathbf{A}$ and $\mathbf{S}$ are to-be retrieved. Most ICA methods consider the determined case where the number of unknown components is the same as the number of electrodes, that is $m = d$. The problem is that this need not be satisfied in case of EEG, especially, when a small number of electrodes is available.

In this respect, the underdetermined BSS (UBSS) case $d > m$ can be taken into account [7]. Contrary to the determined model, the tasks to find $\mathbf{A}$ and $\mathbf{S}$ are not equivalent here. Most UBSS methods therefore consist of two steps where $\mathbf{A}$ is identified first, and $\mathbf{S}$ are separated based on the previously estimated $\mathbf{A}$.

This work was supported by the Grant Agency of the Czech Republic through the projects P103/11/1947.

In this paper, we focus on the classification of EEG recordings that were recorded during imagined right or left hand movement. Our goal is to utilize a recent UBSS algorithm UDSEP [8], for the identification of $\mathbf{A}$, and a novel MMSE beamformer [10], for the retrieval of $\mathbf{S}$, in order to remove unwanted artifacts from EEG recordings before they are sent to a BCI system (either for learning or testing). The BCI system considered here consists of a preprocessing, feature extraction based on CSP, and a Linear Discriminative Analysis (LDA) classifier. We evaluate the performance of the BCI system on several datasets and compare the cases when data are or are not treated via UBSS.

The paper is organized as follows. The following section describes the problem to recognize MI and details of the BCI system. Section 3 introduces the UDSEP algorithm and the MMSE beamformer and describes their utilization for the preprocessing of EEG. Section 4 is devoted to experimental evaluations and comparisons on several datasets.

2. SYSTEM FOR EEG CLASSIFICATION

The classification of motor imagery (MI) can rely on event-related phenomena that consist in either decreases or increases of power in given frequency bands caused by decreases or increases in synchrony of underlying neuronal populations [9]. A decrease and an increase in synchronization is, respectively, called the event-related desynchronization (ERD) and event-related synchronization (ERS).

For example, it was observed that preparation and imagination of movements are accompanied by ERD of the mu (8-13 Hz) and beta (13-30 Hz) rhythms over the contralateral primary sensorimotor area. Some subjects show an ipsilateral or contralateral beta ERS following the beta ERD [9].

This gives us a guideline to recognize and classify signal components related to MI: For the right hand imagery movement, an ERD distributes over the left hemisphere while an ERS distributes over the right hemisphere, and vice versa. We use a setup of few electrodes C3, C4, Cz, and electrodes surrounding them (e.g. CP3, CP4 and CPz) [3, 14].

2.1. Preprocessing

A common preprocessing is to filter the original EEG signals by a band-pass filter. The selected range should correspond with the MI activity. Therefore, the band-pass filter is designed to pass the range of mu and beta rhythms, that is 8-30 Hz [2, 3].

2.2. CSP and Feature Extraction

CSP consists in finding a transform of EEG data by a matrix $\mathbf{W}_{\text{CSP}}$ (a spatial filter) such that it maximizes the difference (in terms of variance) between two training populations of EEG trials: one for left and one for right motor imagery. The problem is solved by generalized eigenvalue decomposition of covariance matrices of the two populations; see [4] for details.

Then, the features for a given trial data $\mathbf{X}$ are computed as the variances of rows of

$$\mathbf{Y} = \mathbf{W}_{\text{CSP}}\,\mathbf{X} \tag{2}$$

where $\mathbf{W}_{\text{CSP}}$ is a $2r \times m$ matrix. The rows of $\mathbf{W}_{\text{CSP}}$ are the generalized eigenvectors corresponding to the r largest and smallest generalized eigenvalues, which are the most suitable for the discrimination of the two populations of EEG trials [4].

Finally, the classification of EEG proceeds by an LDA classifier[1] that was learned on a dataset of training trials with the default setting of its parameters.

3. UBSS PREPROCESSING

As stated in the introduction, the learning process and the classification itself can be badly affected by strong components in EEG data that are not related to MI. Here we propose to remove the artifacts by means of UBSS. The use of blind separation to remove unwanted signals from EEG has been already studied e.g. in [12]. In this paper, we focus on the use of two complementary methods recently published in [8, 10]. The first is the UDSEP algorithm for blind identification of the mixing matrix $\mathbf{A}$ in (1). The second method is the MMSE beamformer that extracts $\mathbf{S}$ from $\mathbf{X}$ based on the estimated $\mathbf{A}$.

3.1. UDSEP

UDSEP belongs to a class of UBSS algorithms that estimate $\mathbf{A}$ through a tensor decomposition [7]. In UDSEP, the components $\mathbf{S}$ are modeled as white block-stationary Gaussian processes, and the tensor to-be decomposed is a three-way tensor built-up of estimated covariance matrices of blocks of $\mathbf{X}$. Let $\mathbf{X}_k$, $k = 1, \ldots, M$, be the kth block of $\mathbf{X}$ and M be the number of blocks. Owing to (1) and the block-stationary model of $\mathbf{S}$, the covariance matrix of $\mathbf{X}_k$ has the structure

$$\begin{aligned} \mathbf{R}_k &= \mathrm{E}[\mathbf{X}_k\mathbf{X}_k^T]/N_1 \\ &= \mathbf{A}\,\mathrm{E}[\mathbf{S}_k\mathbf{S}_k^T]/N_1\mathbf{A}^T \\ &= \mathbf{A}\,\mathrm{diag}[D_{1k}, \ldots, D_{dk}]\mathbf{A}^T \end{aligned} \tag{3}$$

where diag$[\cdot]$ denotes a diagonal matrix, E$[\cdot]$ stands for the expectation operator, and N_1 is the length of blocks. $D_{1k}, \ldots, D_{dk}$ are elements of a matrix $\mathbf{D}$ and denote variances of $\mathbf{S}$ within the kth block.

In fact, the tensor decomposition consists in fitting the estimated covariances

$$\widehat{\mathbf{R}}_k = \mathbf{X}_k\mathbf{X}_k^T/N_1 \tag{4}$$

[1]We use Discriminant Analysis Toolbox v. 0.3 by M. Kiefte available at `www.mathworks.com/matlabcentral/fileexchange/189`.

to their model values, which entails the finding of $\mathbf{A}$ and $\mathbf{D}$. Most tensor decomposition algorithms minimize the mean square error criterion

$$Q_1(\mathbf{A},\mathbf{D}) = \sum_{k=1}^{M} \mathrm{trace}[(\widehat{\mathbf{R}}_k - \widetilde{\mathbf{R}}_k)(\widehat{\mathbf{R}}_k - \widetilde{\mathbf{R}}_k)], \quad (5)$$

where $\widetilde{\mathbf{R}}_k = \mathbf{A}\,\mathrm{diag}[D_{1k},\dots,D_{dk}]\mathbf{A}^T$. UDSEP minimizes the following weighted criterion

$$Q_2(\mathbf{A},\mathbf{D}) = \sum_{k=1}^{M} \mathrm{trace}[\mathbf{C}_k(\widehat{\mathbf{R}}_k - \widetilde{\mathbf{R}}_k)\mathbf{C}_k(\widehat{\mathbf{R}}_k - \widetilde{\mathbf{R}}_k)] \quad (6)$$

where $\mathbf{C}_k$ are weighting matrices. It was shown that the choice $\mathbf{C}_k = (\widehat{\mathbf{R}}_k + \epsilon\mathbf{I})^{-1}$ with a small positive ϵ to restrain regularity is asymptotically optimum within the assumed model of signals $\mathbf{S}$ [8].

UDSEP also exploits the fact that the elements of $\mathbf{D}$ are non-negative. Together with the asymptotic optimality, UDSEP was shown to be more accurate than its competitors, and it succeeded in difficult scenarios such as the separation of 16 speech signals mixed into 9 channels [8]. Its applicability in EEG data processing is promising but was not tested yet.

3.2. MMSE Beamformer

UDSEP is endowed by a beamformer that assumes the same model of $\mathbf{S}$ and separates them block-by-block using the estimated $\mathbf{A}$ and signal variances in $\mathbf{D}$. It estimates $\mathbf{S}$ through time-varying spatial filtering of $\mathbf{X}$ based on maximizing the theoretical signal-to-interference ratio. A disadvantage is that this beamformer does not apply temporal filtering of signals, which may improve the separation. Moreover, it must follow the same block structure of signals that is assumed by UDSEP. This is impractical in situations where the separated signals should be continuous, because the estimated variances of signals do not change continuously on endings of blocks.

Therefore, we have proposed a novel MMSE beamformer in [10] that is independent of UDSEP. It minimizes theoretical mean square distance of estimated signals from the original ones. It assumes that each signal is locally stationary and its autocovariance (spectrum) is changing in time. In other words, non-whiteness of signals is taken into account, which leads to a better separation thanks to a spatio-temporal filtering of signals.

The MMSE beamformer processes signals block-by-block, but the length of blocks N_s is a free integer parameter and can be even equal to one. A further free parameter is the length of temporal filters of signals denoted by P. Here, N_1 has the meaning of the length of block within which signals are assumed to be stationary, which enables to estimate their autocovariances.

The beamformer, represented by a vector $\mathbf{w}_{j,t}$ of length $mP \times 1$, estimates the tth sample of the jth signal as

$$\widehat{S}_{jt} = \mathbf{w}_{j,t}^T \widetilde{\mathbf{X}}_{:,t} \quad (7)$$

where

$$\widetilde{\mathbf{X}}_{:,t} = \begin{bmatrix} \mathbf{X}_{:,t} \\ \mathbf{X}_{:,t-1} \\ \vdots \\ \mathbf{X}_{:,t-P+1} \end{bmatrix}, \quad (8)$$

and $\mathbf{X}_{:,t}$ denotes the tth column of $\mathbf{X}$. The MMSE beamformer is defined to minimize

$$\mathrm{E}\left[(S_{jt} - \widehat{S}_{jt})^2\right] = \mathrm{E}\left[(S_{jt} - \mathbf{w}_{j,t}^T\widetilde{\mathbf{X}}_{:,t})^2\right]. \quad (9)$$

All details are provided in [10], and the m-code implementation is available at [10]. Here, we summarize the steps of the beamformer only. The beamformer begins with $k = 1$.

1. Put $t_k = (k-1)N_s + 1$.
2. Compute the sample covariance matrices
$$\widehat{\mathbf{R}}_{t_k}[\tau] = \frac{1}{N_1} \sum_{\ell=t_k-N_1/2+1}^{t_k+N_1/2} \mathbf{X}_{:,\ell}\mathbf{X}_{:,\ell-\tau}^T \quad (10)$$
for $\tau = 0,\dots,P-1$.
3. Estimate signal auto-covariances for $\tau = 0,\dots,P-1$ as
$$\begin{aligned} \mathbf{d}_{t_k}^{\tau} &= [D_{1t_k}^{\tau},\dots,D_{dt_k}^{\tau}]^T \\ &= \arg\min_{\mathbf{d}} \|\widehat{\mathbf{R}}_{t_k}[\tau] - \mathbf{A}\,\mathrm{diag}(\mathbf{d})\,\mathbf{A}^T\|_F^2 \\ &= (\widetilde{\mathbf{A}}^T\widetilde{\mathbf{A}})^{-1}\widetilde{\mathbf{A}}^T \mathrm{vec}(\widehat{\mathbf{R}}_{t_k}[\tau]) \end{aligned} \quad (11)$$
where $\|\cdot\|_F$ denotes the Frobenius norm, vec$(\cdot)$ is the vectorization operator, and $\widetilde{\mathbf{A}} = \mathbf{A} \odot \mathbf{A}$ where $\odot$ is the Khatri-Rao product.
4. Adjust the resulting signal autocovariances, for the jth signal denoted by $\mathbf{d}_{j,t_k} = [D_{jt_k}^0,\dots,D_{jt_k}^{P-1}]^T$, so that they are positive definite; see [10] for details.
5. For each $j = 1,\dots,d$, put
$$\mathbf{w}_{j,t_k} = \mathbf{G}_{t_k}^{-1}(\mathbf{d}_{j,t_k} \otimes \mathbf{A}_{:,j}) \quad (12)$$
where $\otimes$ is the Kronecker product and
$$\mathbf{G}_t = \begin{bmatrix} \mathbf{R}_t[0] & \mathbf{R}_t[1] & \dots & \mathbf{R}_t[P-1] \\ \mathbf{R}_t[1] & \mathbf{R}_t[0] & \dots & \mathbf{R}_t[P-2] \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{R}_t[P-1] & \mathbf{R}_t[P-2] & \dots & \mathbf{R}_t[0] \end{bmatrix}. \quad (13)$$
The covariance matrices in (13) are estimated by their model values following from (1), that is $\mathbf{R}_t[\tau] = \mathbf{A}\,\mathrm{diag}[\mathbf{d}_t^{\tau}]\mathbf{A}^T$. The jth original signal for $t = (k-1)N_s+1,\dots,kN_s$ is estimated by
$$\widehat{S}_{jt} = \mathbf{w}_{j,t_k}^T \widetilde{\mathbf{X}}_{:,t}. \quad (14)$$
6. Put $k \leftarrow k+1$ and go to 1 unless all available data are processed.

3.3. Artifacts Removal

An important property of the MMSE beamformer is that the separated signals, denoted by $\widehat{\mathbf{S}}$ satisfy

$$\mathbf{A}\widehat{\mathbf{S}} = \mathbf{X}, \tag{15}$$

which enables a consistent reconstruction of $\mathbf{X}$ without selected components (rows of) $\widehat{\mathbf{S}}$. Here, the goal is to withdraw components that are not related to MI.

As justified in Section 2, components can be distinguished based on their spatial activity, which is determined by elements of $\mathbf{A}$. Components related to MI should be dominant over the left-hand sensorimotor area (electrodes C3 and CP3) or over the right-hand sensorimotor area (C4 and CP4), while the other components should be active over the whole surface on a topographic map. Consequently, our steps are as follows. Given a bundle of EEG data in $\mathbf{X}$,

1. estimate $\mathbf{A}$ from $\mathbf{X}$ by UDSEP,
2. determine d_{MI} columns of $\mathbf{A}$ (the components) showing the most different activity on (C3,CP3) and (C4, CP4) in terms of the Fisher score [11],
3. estimate the corresponding rows of $\widehat{\mathbf{S}}$ (MI components) by the MMSE beamformer; the other rows put equal to zero (let $\widehat{\mathbf{S}}_{MI}$ denote the resulting matrix), and
4. put the reconstructed EEG equal to $\mathbf{X}_{MI} = \mathbf{A}\widehat{\mathbf{S}}_{MI}$.

4. EXPERIMENTAL EVALUATION

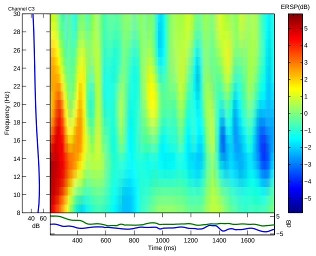
We compare three BCI systems of the same type. The first one uses CSP [4] for feature extraction, the second system utilizes SWCSP [17], and the third system combines the proposed artifacts removal procedure and CSP (UBSS+CSP). In CSP as well as in SWCSP[2], we select $r = 2$, so there are 4 features of each trial data. We use default values for the other parameters in SWCSP. Our tests are done using the following datasets

- Dataset CBCMI for single trial BCI EEG classification provided by [13]. We used EEG signals of subjects S1 and S2 in section 1.
- ABSP EEG motor imagery dataset [14]: subjects A (session 1 and 2), B, and C (sessions 2-4). Datasets where the subject was asked to imagine movement of left hand and both feet (instead of the left or right hand) are denoted by LF.
- BCI III MI dataset (4a) [15]: subjects al and ay.

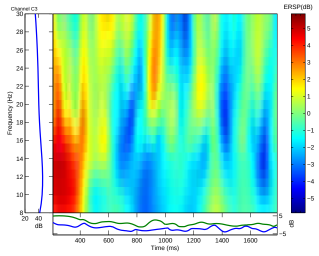
In total, we use 12 datasets on which we evaluate the classification accuracy by the 10×10-fold cross-validation. Results in terms of percents of correctly classified trials are shown in Table 1. The last column of the table shows the parameters of UDSEP and of the MMSE beamformer.

The results confirm that the proposed artifact removal procedure can significantly improve the performance. For instance, the classification accuracy is improved up to by 10% compared to the case when CSP is used directly to untreated EEG. The approach also competes with state-of-the-are SWCSP.

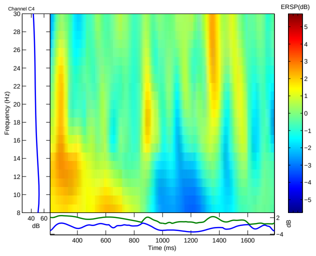
[2] We use the implementation of SWCSP from the BCILAB toolbox by C. Kothe available at `http://mloss.org/software/view/368/`.



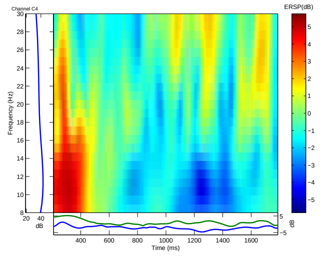
(a) ERSP for the right-hand MI, channel C3.



(b) ERSP for the right-hand MI with removed artifacts, channel C3.



(c) ERSP for the left-hand MI, channel C4.



(d) ERSP for the left-hand MI with removed artifacts, channel C4.

Fig. 1. ERSPs for the dataset S2s1 [13] calculated from EEG before and after the artifacts removal. The ERD patterns (blue areas) are more apparent in (b) and (d) compared to untreated EEG in (a) and (c). The left plots show the mean frequency power subtracted from the ERSP maps. The plots below the maps show the maximum and minimum power over time.

For illustration, Fig. 1 visualizes the event-related spectral perturbation (ERSP) [16], which is an average of power spectrograms of a given channel over trials. A blue area in the ERSP map corresponds to a decrease of power (ERD), while a red area is caused by ERS. Fig. 1(a) illustrates the ERSP map for the right-hand MI for the dataset “S2s1”, which was calculated from 30 trials over channel C3. The same data are evaluated in Fig. 1(b) but after the proposed UBSS procedure for artifacts removal was applied. We conclude that ERD areas in Fig. 1(b) are much more significant than those in Fig. 1(a). Similar conclusions can be drawn when comparing Figures 1(c) and 1(d) for the left-hand MI over C4.

5. CONCLUSIONS

Our experiments confirm that the UDSEP algorithm and the MMSE beamformer can be useful for removal of EEG signal components that are not related to a target activity, here, the motor imagery. The parameters of the methods can be tuned during the learning process to optimize the classification accuracy of a BCI system for a given subject. In this paper, we have demonstrated an improvement of a classical BCI system

Table 1. Classification score of motor imagery EEG

subj.,session [dataset]	f_s(Hz), trial length(s), #trials	CSP	SWCSP	UBSS+CSP	(M, d, P, d_{MI})
S1,1 [13]	500, 2, 60	82.10 $\pm$ 1.80	79.50 $\pm$ 2.61	91.67 $\pm$ 0.42	(240, 12, 9, 2)
S2,1 [13]	500, 2, 60	81.08 $\pm$ 2.33	82.50 $\pm$ 3.54	95.32 $\pm$ 0.81	(120, 13, 11-17, 4)
A,1 [14]	256, 3, 130	87.53 $\pm$ 1.04	86.92 $\pm$ 0.81	89.15 $\pm$ 0.77	(520, 7, 1-20, 6)
A,2 [14]	256, 3, 134	83.81 $\pm$ 1.00	78.21 $\pm$ 2.35	86.34 $\pm$ 0.71	(134, 9, 1-20, 5)
B [14]	250, 4, 162	87.96 $\pm$ 1.28	87.22 $\pm$ 2.00	94.32 $\pm$ 0.57	(810, 11, 1-5, 3)
C,2 [14]	256, 3, 158	85 .43$\pm$ 0.76	79.87 $\pm$ 2.76	87.01 $\pm$ 0.53	(158, 7, 1-20, 3)
C,3 [14]	256, 5, 48	91.46 $\pm$ 0.66	87.50 $\pm$ 3.67	96.25 $\pm$ 1.31	(3072, 10, 1-20, 3)
C,4 [14]	256, 3, 120	88.89 $\pm$ 0.78	86.08 $\pm$ 1.89	91.33 $\pm$ 0.90	(1440, 14,1-20, 4)
C,2(LF) [14]	256, 3, 180	88.78 $\pm$ 0.35	90.72 $\pm$ 0.53	91.67 $\pm$ 0.64	(2912, 15, 1-20, 3)
C,3(LF) [14]	256, 3, 102	85.59 $\pm$ 1.14	92.06 $\pm$ 1.49	90.00 $\pm$ 0.62	(612, 6-12, 1-20, 2)
ay [15]	100, 3.5, 280	89.82 $\pm$ 0.51	88.21 $\pm$ 0.56	93.25 $\pm$ 0.20	(560, 7, 2-20, 5-6)
al [15]	100, 3.5, 280	90.75 $\pm$ 0.36	97.46 $\pm$ 0.26	94.14 $\pm$ 0.42	(280, 8, 18, 5)

based on CSP. A combination of the proposed method with SWCSP will be subject of a further research.

6. REFERENCES

[1] M. Grosse-Wentrup, C. Liefhold, K. Gramann, and M. Buss, "Beamforming in noninvasive brain computer interfaces," *Biomedical Engineering, IEEE Transactions on*, vol. 56, no. 4, pp. 1209–1219, April 2009.

[2] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Muller, "Optimizing spatial filters for robust EEG single-trial analysis," *Signal Processing Magazine, IEEE*, vol. 25, no. 1, pp. 41–56, 2008.

[3] M. Naeem, C. Brunner, and G. Pfurtscheller, "Dimensionality reduction and channel selection of motor imagery electroencephalographic data," *Intell. Neuroscience*, vol. 2009, pp. 5:1–5:8, January 2009.

[4] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *Rehabilitation Engineering, IEEE Transactions on*, vol. 8, no. 4, pp. 441–446, Dec 2000.

[5] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*, J. Wiley, 2001.

[6] T. P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. J. McKeown, V. Iragui, T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, Vol. 37, No. 2, pp. 163-178, March 2000.

[7] L. De Lathauwer and J. Castaing, "Blind identification of underdetermined mixtures by simultaneous matrix diagonalization," *Signal Processing, IEEE Transactions on*, vol. 56, no. 3, pp. 1096–1105, March 2008.

[8] P. Tichavský and Z. Koldovský, "Weight adjusted tensor method for blind separation of underdetermined mixtures of nonstationary sources," *IEEE Transactions on Sig. Proc.*, vol. 59, no. 3, pp. 1037–1047, March 2011.

[9] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related eeg/meg synchronization and desynchronization: basic principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842 – 1857, 1999.

[10] Z. Koldovský, P. Tichavský, A. H. Phan, and A. Cichocki, "Minimum Mean Square Error Beamformer for Underdetermined Signal Separation," *submitted*, March 2012, the Matlab code and preprint available at `http://itakura.ite.tul.cz/zbynek`.

[11] A. H. Phan and A. Cichocki, "Tensor Decompositions for Feature Extraction and Classification of High Dimensional Datasets," *Nonlinear Theory and Its Applications*, IEICE", vol. 1, no. 1, pp. 37–68, 2010.

[12] H. Liu, P. H. Schimpf, G. Dong, X. Gao, F. Yang, and S. Gao, "Standardized shrinking LORETA-FOCUSS (SSLOFO): a new algorithm for spatio-temporal EEG source reconstruction," *IEEE Trans Biomed Eng*, vol. 52, no. 10, pp. 1681–1691, 2005.

[13] L. Zhang et al., "Data set for single trial EEG classification in BCI.", Center for Brain-Like Computing and Machine Intelligence, available at `http://bcmi.sjtu.edu.cn`.

[14] A. Cichocki and Q. Zhao, "EEG motor imagery dataset," Tech. Rep., Laboratory for Advanced Brain Signal Processing, BSI, RIKEN, Saitama, Japan, 2011, available at `http://www.bsp.brain.riken.jp/~qibin/homepage/Datasets.html`.

[15] G. Dornhege, B. Blankertz, G. Curio, and K.-R. Müller, "Boosting bit rates in non-invasive EEG single-trial classifications by feature combination and multi-class paradigms," *IEEE Trans. Biomed. Eng.*, vol. 51, 2004.

[16] S. Makeig, "Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones.," *Electroencephalography and clinical neurophysiology*, vol. 86, no. 4, pp. 283–293, Apr. 1993.

[17] R. Tomioka, G. Dornhege, G. Nolte, B. Blankertz, K. Aihara, and K.-R. Müller, "Spectrally Weighted Common Spatial Pattern Algorithm for Single Trial EEG Classification, " *Mathematical Engineering Technical Reports (METR-2006-40)*, The University of Tokyo, July 2006.