Comparison of Adaptive Spatial Filters with Heuristic and Optimized Region of Interest for EEG Based Brain-Computer-Interfaces

Christian Liefhold[1], Moritz Grosse-Wentrup[1], Klaus Gramann[2,3], and Martin Buss[1]

[1] Institute of Automatic Control Engineering, Technische Universität München, 80333 München, Germany
christian.liefhold@mytum.de, moritz@tum.de, mb@tum.de
[2] Ludwig-Maximilians-Universität München, Department Psychology, 80802 München, Germany
[3] Swartz Center for Computational Neuroscience, Department for Neural Engineering, University of California, San Diego
gramann@psy.uni-muenchen.de

Abstract. Research on EEG based brain-computer-interfaces (BCIs) aims at steering devices by thought. Even for simple applications, BCIs require an extremely effective data processing to work properly because of the low signal-to-noise-ratio (SNR) of EEG signals. Spatial filtering is one successful preprocessing method, which extracts EEG components carrying the most relevant information. Unlike spatial filtering with Common Spatial Patterns (CSP), Adaptive Spatial Filtering (ASF) can be adapted to freely selectable regions of interest (ROI) and with this, artifacts can be actively suppressed. In this context, we compare the performance of ASF with ROIs selected using anatomical a-priori information and ASF with numerically optimized ROIs. Therefore, we introduce a method for data driven spatial filter adaptation and apply the achieved filters for classification of EEG data recorded during imaginary movements of the left and right hand of four subjects. The results show, that in the case of artifact-free datasets, ASFs with numerically optimized ROIs achieve classification rates of up to 97.7 % while ASFs with ROIs defined by anatomical heuristic stay at 93.7 % for the same data. Otherwise, with noisy datasets, the former brake down (66.7 %) while the latter meet 95.7 %.

1 Introduction

Steering wheel chairs, prostheses or technical instruments we use every day only by thoughts is the proximal goal of research in context with brain-computer-interfaces (BCIs). This is mainly beneficial for handicapped people or people suffering from the locked-in-syndrome. In the long term, it could additionally enrich the everyday life for all people.

The technical approach we use for "reading thoughts" consists in recognizing patterns of recorded electroencephalogram data (EEG) from different subjects.

F.A. Hamprecht, C. Schnörr, and B. Jähne (Eds.): DAGM 2007, LNCS 4713, pp. 274–283, 2007.

For example, motor imagery leads to a decrease of the variance of the EEG in specific frequency bands over the motor cortex [1]. This event-related desynchronization (ERD) can be used to infer the intention of the user of the BCI [2]. The main problem in this context is the low signal-to-noise-ratio (SNR) of the recorded EEG. This has led to the development of spatial filters, extracting that component of the EEG that carries the most relevant information about the classification task. One of the most successful algorithms for spatial filtering in the context of BCIs is the Common Spatial Patterns (CSP) algorithm ([3], [4]). For two conditions (i.e., imaginary movement of right and left hand), it minimizes the variance of one dataset, while maximizing the variance of the other dataset. This leads to spatial filters that enable very good classification results in combination with linear classification algorithms. However, the CSP algorithm is not robust against artifacts like eye activity or mental states that are not induced by the imagination of the specific movement.

To overcome this disadvantage we use the Adaptive Spatial Filters (ASF) algorithm [5]. In ASF, it is presumed that the imagination of a specific movement causes a change of the EEG in one specific location of the cortex. Since it is well known that imaginary movements of a limb cause an ERD in that part of the motor cortex representing the specific limb [1], this a-priori knowledge can be used to design a spatial filter that attenuates all EEG activity not originating in this region of interest (ROI). With ASFs, we thus only pick out information originating from the desired sources, and eliminate unwanted EEG activity.

To obtain a good adaptation of the spatial filters to the desired EEG sources, it is necessary to estimate their true origins. This estimation can be based on heuristic a-priori knowledge or on source localization by solving data driven optimization problems. The former does not require time-consuming computation while the latter does. Hence, numerical optimization is not the choice for online capable BCIs. In this context, our intention is to analyze the performance of heuristic filter adaptation compared to numerically optimized filter adaptation. We expect that this will help to improve the performance of ASFs and show the achievable limits. More precisely, in [5] experiments with motor imagery of the left and right hand were performed. Two ASFs were trained with a-priori information about the approximate positions of the left and right motor cortex. These positions were heuristically assumed as spherical regions with 5 mm radius and midpoints located on a sphere around the center of the head with radius 6.6 cm, at the points, where radial projections of C3 and C4 cut this sphere. This assumption was based on anatomical knowledge about the human brain, but for sure, it does not exactly hold for all subjects. This raises the following question: *Is there a better way to define the regions of interest (ROI)?* Or more precise: *How far is the location of a numerically optimized ROI from the heuristic one?* By setting up a method to numerically optimize the position of the sources for preprocessed training data, and by comparing the achieved classification results with the results achieved using heuristically adapted spatial filters, we will give the answer to this question.

The structure of the paper is the following. In Section 2, the ASF algorithm is briefly reviewed and the numerical optimization of the ROI is described. In Section 3, we will apply heuristically adapted ASFs as well as numerically optimized ASFs on EEG data of four subjects, and compare the achieved classification results. A discussion of the different results and some prospects on future research will conclude the paper in Section 4.

2 Methods

2.1 Adaptive Spatial Filters

An adaptive spatial filter tries to minimize the signal variance of sources outside a region of interest, while leaving unchanged the signal variance of sources inside the ROI. It can be described by a filter vector $\boldsymbol{w} \in \mathbb{R}^M$, with M the number of EEG electrodes. Applying it on a signal $\boldsymbol{x}(t) \in \mathbb{R}^M$ returns the signal $y(t) \in \mathbb{R}$, caused by sources inside the ROI:

$$y(t) = \boldsymbol{w}^T \boldsymbol{x}(t). \tag{1}$$

The filter vector $\boldsymbol{w}$ can be found by

$$\boldsymbol{w}^* = \underset{\boldsymbol{w}}{\operatorname{argmax}}\{f(\boldsymbol{w})\} \quad \text{with} \quad f(\boldsymbol{w}) = \frac{\boldsymbol{w}^T R_{\tilde{\boldsymbol{x}}}(t)\, \boldsymbol{w}}{\boldsymbol{w}^T R_{\boldsymbol{x}}(t)\, \boldsymbol{w}}, \tag{2}$$

where $R_{\boldsymbol{x}}(t)$ is the covariance matrix of measured EEG data while $R_{\tilde{\boldsymbol{x}}}(t)$ denotes the covariance matrix of model generated EEG data, that shows only activity of sources originating in the ROI. For the latter, a four-shell spherical head model is used to compute the electrical field at a position $\boldsymbol{r}_i \in \mathbb{R}^3$ (i.e., the EEG signal $\tilde{x}_i(t)$ of the i-th EEG electrode, $i = 1 \ldots M$, with M the number of EEG electrodes) caused by an electric dipole at position $\boldsymbol{r'} \in \mathbb{R}^3$ [6]. The basic equation therefore is described by

$$\Phi(\boldsymbol{r}_i, t) = \tilde{x}_i(t) = \boldsymbol{l}(\boldsymbol{r}_i, \boldsymbol{r'})^T \boldsymbol{p}(t). \tag{3}$$

$\boldsymbol{l}(\boldsymbol{r}_i, \boldsymbol{r'}) : \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$ denotes the leadfield equations, describing the projection strength of a dipole at position $\boldsymbol{r'}$ with moment $\boldsymbol{p}(t) \in \mathbb{R}^3$ on a position $\boldsymbol{r}_i$ outside the head with respect to conductive and geometric properties of the head model. Hence, with (3) we can compute the signal $\tilde{\boldsymbol{x}}(t)$ of all EEG electrodes for a given dipole source and adapt a spatial filter to this dipole source by solving (2). This maximizes signal variance, which may seem contradictory to the goal of finding ERDs of an EEG-Signal, i.e., decreases of variance. But these steps are independent. At first we want to attenuate the variance of all EEG sources not within the ROI with an optimal spatial filter found by (2). Then we can proceed with the variance information originating within the ROI to vote for an ERD or not. The process of predefining a ROI can be carried out by applying a-priori information about cortex positions or by data driven, numerical optimization, i.e., source estimation.

2.2 Optimizing the Region of Interest

While in [5] the ROI was chosen heuristically, we will now address the question how the ROI can be optimized numerically. We will utilize source localization for this purpose, i.e., we will try to localize the origin of the ERD changes observed during motor imagery. The aim in source estimation is to find sources which give the best approximation of EEG data generated by a suitable model of EEG conduction, e.g., a dipole placed inside a four-shell head model, to real EEG data taken from an eligible training set. If a good aproximation is obtained, the estimated source locations should be quite close to the true source locations inside the human brain. Therefore, adapting the ASFs to the estimated sources should increase the achievable classification rate.

As we use the signal variance in each ROI as a classification paradigm, we look for a dipole that gives the best approximation of the variance of the measured EEG data. Hence, we perform source estimation using the variance of EEG data. The main steps for estimating the source locations consist in creating a training dataset and in constructing a suitable cost function.

Creation of the Training Dataset. We create a training dataset from EEG data measured during motor imagery of the left and right hand (four subjects, 150 trials of 10 s per condition, see chapter 3.1 or [5] for details of the experimental setup) using the data processing chain depicted in Fig. 1. At first, to reduce artifacts, we perform Independent Component Analysis (ICA) using the extended Infomax algorithm [7] as implemented in [8], and manually reject those components that are obviously induced by eye activity. All remaining components are reprojected onto the observation space. Then, the most reactive frequency band for each subject is determined heuristically. This is done by computing ERD/ERS changes for electrodes C3 and C4 situated over the motor cortex across all trials, and determining that frequency band that shows the highest ERD during motor imagery relative to a baseline during rest. Afterwards, a sixth-order butterworth bandpass (BP) filter is applied to the data, extracting the most reactive subject specific frequency band. Then, we take the same data intervals of motor imagery that are used for later classification as basis for our optimization (6.5 s per each condition and trial). Subsequently, the variance of the chosen interval is computed for each trial condition, imagination of left-hand-movement (IL) and imagination of right-hand-movement (IR), and averaged over all proper trials. At this point, we have two M-dimensional variance vectors for each subject: variance vector of condition IL and variance vector of condition IR. Now, as we want to extract the distinction of these two conditions, we build the difference between them. Only the negative outcomes of this difference will be used, because we are interested in the decrease of the variance (ERD). Positive outcomes are set to zero. Finally, the absolute values of the remaining variances are computed, since only absolute (positive) variances can be modeled. The result forms the data input for the optimization problem, which is described in the next subsection.

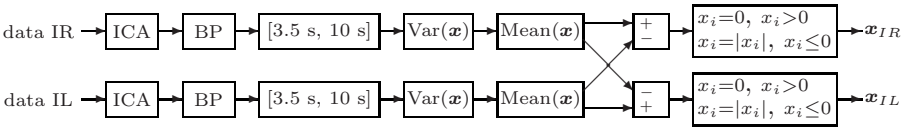

Fig. 1. Data processing chain to compute the training dataset for each subject. See text for more details.

Derivation of the Cost Function. Before we go to the next step of the dipole source estimation, let us assume the following:

- The superposition of many sources originating in each region of interest can be represented by only one single electric dipole.
- For each condition, a dipole can be characterized by a time-invariant dipole position vector $\boldsymbol{r}'$ and a dipole moment vector $\boldsymbol{p}(t)$ with time-invariant orientation, each $\epsilon\, \mathbb{R}^3$.
- The EEG data is of zero mean (due to bandpass filtering) and of stationary variance during each training condition.

Now, we will derive a cost function $J(\boldsymbol{r}', \boldsymbol{p}) : \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathbb{R}$. In general, the cost function should express the difference of the reconstructed signal variance and the signal variance of the training set. Our choice is

$$J[\boldsymbol{r}', \boldsymbol{p}(t)] = \| \boldsymbol{\sigma}^2_{\boldsymbol{x}(t)} - \boldsymbol{\sigma}^2_{\tilde{\boldsymbol{x}}(t)} \|^2_2 \tag{4}$$

with $\boldsymbol{\sigma}^2_{\boldsymbol{x}(t)}$ the variance of the training set data $\boldsymbol{x}(t)$ and $\boldsymbol{\sigma}^2_{\tilde{\boldsymbol{x}}(t)}$ the variance of the model generated data $\tilde{\boldsymbol{x}}(t)$, each $\epsilon\, \mathbb{R}^M$. Now, let us consider the signal of one single EEG channel $x_i(t)\ \epsilon\, \mathbb{R}$. We can expand the difference of the variances by expressing the variance operators as sample variances (zero-mean assumption) such that

$$\sigma^2_{x_i[n]} - \sigma^2_{\tilde{x}_i[n]} = \frac{1}{2N} \sum_{n=-N}^{N} x_i[n]^2 \; - \frac{1}{2N} \sum_{n=-N}^{N} \tilde{x}_i[n]^2, \tag{5}$$

where the variances are computed over $2N$ data samples. According to (3), the model generated signal $\tilde{x}_i(t)$ of the i-th EEG electrode can be written as

$$\tilde{x}_i[n] = \boldsymbol{l}(\boldsymbol{r}_i, \boldsymbol{r}')^T\, \boldsymbol{p}[n] \tag{6}$$

which is the dot product of the leadfield equations and the dipole moment vector $\boldsymbol{p}[n]\ \epsilon\, \mathbb{R}^{3\times 1}$. With (6), we can rewrite the model built signal variance to

$$\sigma^2_{\tilde{x}_i[n]} = \frac{1}{2N} \sum_{n=-N}^{N} [\boldsymbol{l}(\boldsymbol{r}_i, \boldsymbol{r}')^T\, \boldsymbol{p}[n]]^2. \tag{7}$$

Now, we express the dipole moment vector $\boldsymbol{p}[n]$ by an amplitude $s[n] \,\epsilon\, \mathbb{R}$ and a normalized moment $\tilde{\boldsymbol{p}} \,\epsilon\, \mathbb{R}^{3\times 1}$, which is independent of n due to our assumptions (time-invariant orientation):

$$\boldsymbol{p}[n] = \frac{\boldsymbol{p}[n]}{\|\boldsymbol{p}[n]\|_2}\, s[n] = \tilde{\boldsymbol{p}}\, s[n]. \tag{8}$$

We further get

$$\sigma^2_{\tilde{x}_i[n]} = \frac{1}{2N} \sum_{n=-N}^{N} [\boldsymbol{l}(\boldsymbol{r}_i, \boldsymbol{r}')^T\, \tilde{\boldsymbol{p}}]^2\, s[n]^2. \tag{9}$$

The normalized moment vector $\tilde{\boldsymbol{p}}$ is completely determined by two angles, which are denoted as θ and φ in the following . We can merge together all the nonlinearly influencing variables into a function $g(\boldsymbol{r}, \boldsymbol{r}', \theta, \varphi) : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ and write (9) as

$$\sigma^2_{\tilde{x}_i[n]} = g(\boldsymbol{r}_i, \boldsymbol{r}', \theta, \varphi)^2\, \frac{1}{2N} \sum_{n=-N}^{N} s[n]^2 = g(\boldsymbol{r}_i, \boldsymbol{r}', \theta, \varphi)^2\, \sigma_s^2. \tag{10}$$

The second equation sign holds, as we assumed the signal variance to be stationary for one training condition (σ_s^2 denotes the variance of $s[n]$). Now, let's consider all available EEG channels. We can construct a M-dimensional variance vector $\boldsymbol{\sigma}^2_{\tilde{\boldsymbol{x}}[n]}$ by just stacking all M components on top of each other, which results in

$$\sigma^2_{\tilde{\boldsymbol{x}}[n]} = \begin{pmatrix} g(\boldsymbol{r}_1, \boldsymbol{r}', \theta, \varphi)^2\, \sigma_s^2 \\ \vdots \\ g(\boldsymbol{r}_i, \boldsymbol{r}', \theta, \varphi)^2\, \sigma_s^2 \\ \vdots \\ g(\boldsymbol{r}_M, \boldsymbol{r}', \theta, \varphi)^2\, \sigma_s^2 \end{pmatrix} = \mathrm{diag}\{\boldsymbol{g}\boldsymbol{g}^T\}\, \sigma_s^2 \tag{11}$$

with $\boldsymbol{g} = \boldsymbol{g}(\boldsymbol{r}', \theta, \varphi) : \mathbb{R}^3 \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}^M$. As the electrode positions $\boldsymbol{r}_1, ..., \boldsymbol{r}_M$ are well known, they are not mentioned as arguments of $\boldsymbol{g}$. Then, the final cost function yields to

$$J(\boldsymbol{r}', \boldsymbol{p}) = J(\boldsymbol{r}', \theta, \varphi, \sigma_s^2) = \left\| \frac{1}{2N}\, \mathrm{diag}\{XX^T\} \; - \mathrm{diag}\{\boldsymbol{g}\boldsymbol{g}^T\}\, \sigma_s^2 \right\|_2^2 \tag{12}$$

with $X \,\epsilon\, \mathbb{R}^{M\times 2N}$ the data sample matrix, containing $2N$ data samples of all M EEG channels. Now, an inequality constraint for this cost function is necessary, because we are only interested in sources located inside the brain (i.e., the innermost sphere of the head model). Hence, our optimization problem can be expressed by

$$\min_{\boldsymbol{r}', \theta, \varphi, \sigma_s^2} \{J(\boldsymbol{r}', \theta, \varphi, \sigma_s^2)\} \quad \text{s. t.} \;\; \|\boldsymbol{r}'\|_2 \leq R \tag{13}$$

with $R \,\epsilon\, \mathbb{R}$ the radius of the innermost head model sphere. If this optimization problem is numerically solved for a given training dataset, we retrieve an optimized dipole position $\boldsymbol{r}'_{opt}$ and optimized moment parameters $\theta_{opt}, \varphi_{opt}, \sigma^2_{s,opt}$.

The optimization problem (13) is solved using sequential quadratic programming as implemented in the Matlab optimization toolbox. To avoid getting stuck in local minima, an equally with 4 cm spaced, three-dimensional grid of start vectors $\boldsymbol{r}'_{0,k}$ ($k = 1 \ldots K$, with $K = 32$ the number of grid points) is processed to find the best global start vector. For each grid point k, optimal moment parameters are computed by minimizing the cost function $J(\boldsymbol{r}' = \boldsymbol{r}'_{0,k}, \theta, \varphi, \sigma_s^2)$. We also perform the optimization under the constraint of radially oriented dipole moments:

$$\min_{\boldsymbol{r}', \theta, \varphi, \sigma_s^2} \{J(\boldsymbol{r}', \theta, \varphi, \sigma_s^2)\} \quad \text{s. t.} \quad \|\boldsymbol{r}'\|_2 \leq R \tag{14}$$

$$\begin{pmatrix} \sin(\theta)\cos(\varphi) \\ \sin(\theta)\sin(\varphi) \\ \cos(\theta) \end{pmatrix} = c\,\boldsymbol{r}'$$

with $c \in \mathbb{R}$. This constraint is due to the fact that the cortical columns in the motor cortex are usually oriented radially to the surface of the cortex [9].

Based on the optimized position of the electric dipole, we create the ROI by placing a spherical dipole grid with radius 5 mm around the midpoint $\boldsymbol{r}'_{opt}$, while ensuring that no dipole lies outside the innermost sphere of the head model. The moment of each grid dipole is determined by either the optimized moment parameters in the case of unconstrained dipole orientation (13), or by the radial constraint in the case of the radially constrained optimization problem (14).

3 Results

In this chapter, we will briefly describe the setup of the experiment that delivered the EEG data used for the comparison of heuristically and numerically adapted spatial filters. More detailed information can be found in [5]. Afterwards, we will present the classification results achieved with each of the two different adaptation methods.

3.1 Experimental Setup

Four male subjects (S1, S2, S3, S4) took part in the experiment. They were 26, 30, 27 and 24 years old. Wearing an EEG cap with 128 electrodes, the subjects had to perform imaginary movements of the right and left hand, according to the direction of an arrow projected on a screen. The experiment was divided into 300 trials, each of 10 s length. The presentation of the arrow started at 3 s after the beginning of the actual trial and ended with the beginning of the next trial. The pointing direction of the arrow was selected from a list, created by randomizing 150 left directions and 150 right directions. In the first 3 s of each trial, a fixation cross replaced the arrow. The EEG data of each subject was recorded with 500 Hz sampling frequency using a common average reference.

3.2 Experimental Results

Per subject, we performed a heuristic as well as a numerical spatial filter adaptation as described in the methods sections. The resulting ASFs were then applied on two frequency bands of the EEG data. The first frequency band was found during the numerical optimization (see 2.2) while the second frequency band has been fixed between 20 Hz and 30 Hz, as ERD related to motor imagery can also be observed for these frequencies. Then, we computed the variance for each spatially filtered signal within 3.5 s and 10 s of each trial. This delivered a four-dimensional feature vector for each kind of filter and each trial, whereas two dimensions represent the activity in the ROI for the condition IR per spectral band and the other dimensions the activity in the ROI for the condition IL per spectral band. These feature vectors were then classified by using leave-one-out cross validation with Fisher Linear Discriminant Analysis [10].

The results of the classification are listed in Tab. 1. For the subjects S2 and S3, numerical optimization enhanced the classification rate according to our expectation from 81.0 % to 85.7 % and from 93.7 % up to 97.7 %. For the other subjects, a-priori classification still achieved the best classification results. The ROIs of the subjects S2 and S3 (radial constraint), for which numerical optimization improved the classification results, differ from the heuristically set ROIs with 2.38 cm (S2, IL), 1.56 cm (S2, IR), 1.65 cm (S3, IL) and 1.97 cm (S3, IR).

Table 1. Classification results achieved with heuristically and numerically optimized spatial filters for subjects 1-4

Subject	Frequency band 1	heuristic ASF	numerically optimized ASF	
			unconstrained	radial constraint
1	17 – 18 Hz	54.3 %	48.3 %	52.7 %
2	12 – 14 Hz	81.0 %	83.7 %	85.7 %
3	12 – 14 Hz	93.7 %	96.5 %	97.7 %
4	12 – 14 Hz	95.7 %	52.0 %	66.7 %

4 Discussion

Considering only numerically optimized filters, the adaptation using radially constrained dipoles delivered better classification results than adaptation with unconstrained dipole orientations. This may be due to a higher artifact resistance of the constrained optimization problem. This property becomes clearly visible by comparing the variance topographies of subject S3 shown in Fig. 2. Using unconstrained numerical adaptation (Fig. 2, [b]), the model built EEG variance reproduces both, the activity of the motor cortex and the artifacts observed in the region of the occipital lobe. This does not happen with a radially constrained dipole (Fig. 2, [c]).

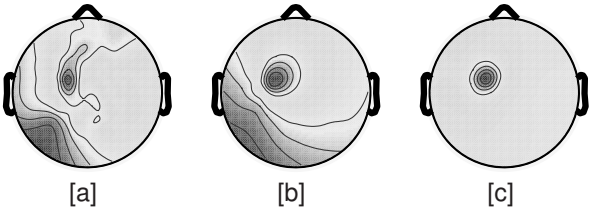

Fig. 2. Variance topographies of subject S3 and condition IR, computed from measured EEG data [a], from unconstrained dipole modeled data [b], and from radial constrained dipole modeled data [c]

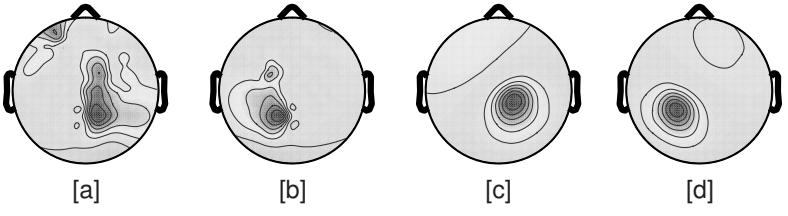

Fig. 3. Variance topographies of subject S4, computed from measured EEG data for IL [a], IR [b] and from model built data variance for IL [c], IR [d]

The collapse of the numerical methods for the data of subject S1 can be explained by the in principle poorly classifiable data set. It shows up many artifacts. Hence, a data driven approach to feature extraction must lead to worse results than one which is strictly independent of training data. The heuristic approach ignores artifacts and extracts the existing classifiable information that corresponds to the assumptions made. The poor classification rate in the case of numerically optimized filter adaptation for subject S4 might be surprising, as the result with a heuristically adapted spatial filter constitutes 95.7 %. But if we take a look at the measured EEG variance of this subject (Fig. 3), the reason becomes evident: There exists a second significant activation region for each condition (IL, IR), which is situated over the parietal cortex. This observation suggests that, contrary to most subjects, subject S4 recruits motor as well as parietal areas for motor imagery. The numerical filter adaptation method introduced in this paper is not able to handle such an occurrence, as we assumed only one activation region for each trial condition. Future work in this context should extend this method, to enable it to adapt to more than one activation region.

In summary, it was shown that numerical optimization of the ROI of an ASF can enhance the quality of the extracted features and thus lead to better classificaton results. This increase in classification accuracy however is not significant and requires an almost artifact-free data set. Furthermore, for those two subjects that showed an improvement in classification accuracy, the centers of the

heuristically chosen and numerically optimized ROIs differed on average by 1.89 cm with an average improvement in classification accurcay of 4.35 %. This suggests that mispositioning of the ROI by several centimeters only results in moderate decreases of classification accuracy.

We thus conclude that numerical optimization of the ROI of ASFs is a viable option if artifact-free training data is available, and very high classification accuracies are desired. However, heuristic positioning of the ROI achieves classification accuracies close to those obtainable by numerical optimization of the ROI without the lack of robustness associated with data driven optimization techniques.

References

1. Pfurtscheller, G., Lopes, F.H.: Event-related EEG/MEG synchronization and desynchronization: basic principles. Clinical Neurophysiology 110, 1842–1857 (1999)
2. Pfurtscheller, G., Neuper, C., Flotzinger, D., Pregenzer, M.: EEG-based discrimination between imagination of right and left hand movement. Electroencephalography and Clinical Neurophysiology 103, 642–651 (1997)
3. Ramoser, H., Mueller-Gerking, J., Pfurtscheller, G.: Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Transactions on Rehabilitation Engineering 8(4), 441–446 (2000)
4. Blanchard, G., Blankertz, G.: BCI competition 2003 - data set IIa: Spatial patterns of self-controlled brain rythm modulations. IEEE Transactions on Biomedical Engineering 51(6), 1062–1066 (2004)
5. Grosse-Wentrup, M., Gramann, K., Buss, M.: Adaptive spatial filters with predefined region of interest for EEG based brain-computer-interfaces. In: Schölkopf, B., Platt, J., Hoffman, T. (eds.) Advances in Neural Information Processing Systems 19, MIT Press, Cambridge, MA (2007)
6. Cuffin, B.N., Cohen, D.: Comparison of the magnetoencephalogram and electroencephalogram. Electroencephalography and Clinical Neurophysiology 47(2), 132–146 (1979)
7. Lee, T.W., Girolami, M., Sejnowski, T.J.: Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. Neural Computation 11, 417–441 (1999)
8. Delorme, A., Makeig, S.: EEGlab: an open source toolbox for analysis of singletrial EEG dynamics including independent component analysis. Journal of Neuroscience Methods 134(1), 9–21 (2004)
9. Nunez, P.L., Shrinivasan, R.: Electric Fields of the Brain. In: The Neurophysics of EEG, 2nd edn., Oxford University Press, Oxford (2006)
10. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification, 2nd edn. Wiley, Chichester (2000)