Task Related & Spatially Regularized Common Spatial Patterns for Brain Computer Interfaces

Ayhan Yuksel and Tamer Olmez

Department of Electronics and Communication Engineering,
Istanbul Technical University, 34469 Istanbul, Turkey
ayhanyks@yahoo.com,olmezt@itu.edu.tr

Abstract. In this study, a novel regularized common spatial pattern method is introduced. Spatial filtering is an important processing step for feature extraction in motor imagery based brain computer interfaces. Common Spatial Patterns (CSP) method is an effective spatial filter for discriminating different motor imagery signals acquired using large number of EEG electrodes. Unfortunately, CSP method is sensitive to nonstationery sources like artefacts and noise, which cause overfitting. In the literature, some regularization methods developed in order to avert overfitting and generate filters that are less sensitive to noise. In this study, we present a method that regularizes CSP filters by taking care of physiological sources of executed motor imagery tasks and spatial relations between electrodes. We compared our method to well known CSP methods on a publicly available EEG dataset by calculating classifying performances and analyzing the effect of regularizing CSP visually. Results show that proposed method gives the best overall performance among six CSP methods.

Keywords: Brain Computer Interfaces (BCI), Motor Imagery (MI), EEG, Common Spatial Patterns (CSP), Regularization

1 Introduction

Brain Computer Interface (BCI) is a communication alternative between user and system where user does not need to use his brain muscular pathways for controlling an external device [1]. Since it is a direct communication method with brain and outer world, BCI system emerges a useful communication and control method for severely paralyzed people. In such a system, user should generate different signal (usually EEG) patterns with his brain for different commands. Moreover, discriminating these brain patterns and translating them to control commands for an electronic device is the most important part of the BCI system.

In motor imagery based BCIs, users motor intentions are extracted by analyzing multi channel EEG signal. In such system, users imagery of moving a limb is decoded into device control commands. The first motor imagery based

BCI design developed by Pfurtscheller et al. [2] analyzed EEG power changes during imaging of left and right arm movements.

A focal power decrease in a specific frequency band, called event related desynchronization (ERD), can be observed over the motor cortex during motor imagery [3, 4]. In 1937, Penfield and Boldrey studied on organization of motor cortex in human brain [5]. According to this study, any muscle group has a specific area on the motor cortex. By comparing the size of each of the areas on the motor cortex, homunculus figure was created (See: Figure 1).

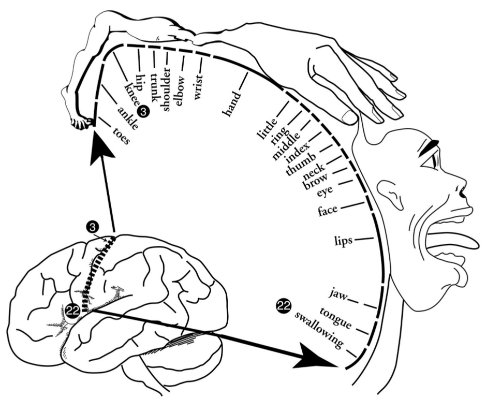

Fig. 1. Human motor cortex and homunculus [6]

Due to the topographical organization in motor cortex, different motor imagery tasks can be identified by their specific spatial location of related ERD rhythms [7]. However, due to the volume conduction effect, scalp EEG signal recorded from a specific area involves a mixture of several cortical sources located in different areas. Thus, raw scalp EEG potentials have poor spatial resolution [7]. For the purpose of eliminating the volume conduction effect and reaching the actual underlying signal sources, spatial filtering step is an indispensable technique [8].

Common Spatial Pattern (CSP) is a very popular and powerful spatial filtering method used in motor imagery EEG classification [9]. When using band power features, CSP computes spatial filters by aiming optimal discrimination between two classes [10]. CSP finds optimal spatial filters which maximize the variance ratio of two different classes. A computed CSP spatial filter projects multi dimensional EEG time domain signal to one dimensional signal in which the power of one class is maximized while power other class is minimized. Unlike PCA, CSP takes care of two classes at the same time and simultaneously diagonalize the covariance matrix of each class [11]. Moreover, CSP algorithm was proven to be efficient in BCI competitions [12, 13]. Although CSP is a powerful technique, with simplicity, it has some drawbacks. Due to its optimization function which tries to maximize the ratio of variances, CSP algorithm is very sensitive to outliers which cause over fitting [14]. Recently, some methods called regularized CSP (RCSP) have been proposed which aims computing more robust spatial patterns by add a regularization term to the CSP formula [8, 15, 10].

In this paper, we present a regularization method for CSP, called "Task related & spatially regularized common spatial patterns (TR&SR-CSP). Proposed method unifies prior information about executed motor imagery tasks and spatial locations of EEG channels. We tested our algorithm along with existing (regularized) CSP algorithms. Moreover, we visualized generated filters on the head model and show the effect of regularization.

The remainder of this study is organized as follows: In Section II, standard CSP method is briefly reviewed. In Section III, regularization framework and proposed TR&SR-CSP method with existing regularization methods are described. Section IV informs about motor imagery data set used in the study and EEG preprocessing routine, summarizes the evaluations and results of the study. Finally, Section V concludes the paper.

2 Common Spatial Patterns

CSP is used widely technique in order to have a good spatial resolution and discriminate between different motor imagery signals. Generally, a motor imagery experiment consists of epochs, in which the user imagines one kind of motor imagery task requested on the screen. An epoch can be one of the two classes: C_1 and C_2. (i.e. left hand - right hand). Let $X_{C,i} \in \mathbb{R}^{NxT}$ represent an epoch, where C is the class of epoch and i is the epoch number belonging to class C, N is the number of EEG channels and T is the number of samples in the epoch. Note that $X_{C,i}$ should be a zero average signal (i.e. band pass filtered). Let $\boldsymbol{w} \in \mathbb{R}^{Nx1}$ be a vector in N dimensional space. A projection of an epoch onto this vector will be

$$\boldsymbol{y}_{C,i} = \boldsymbol{w}^T X_{C,i} \tag{1}$$

where $\boldsymbol{y}_{C,i} \in \mathbb{R}^{1xT}$ denotes the projection of epoch $X_{C,i}$ and T is the transpose operation. Projected signal power $P_{C,i}$ can be written,

$$P_{C,i} = \boldsymbol{y}_{C,i}\boldsymbol{y}_{C,i}^T = \boldsymbol{w}^T X_{C,i} X_{C,i}^T \boldsymbol{w} \tag{2}$$

let $R_{C,i} \in \mathbb{R}^{NxN}$ be the covariance matrix of band pass filtered signal $X_{C,i}$ and $\overline{R}_C \in \mathbb{R}^{NxN}$ be the average covariance matrix of class C:

$$R_{C,i} = \frac{X_{C,i}X_{C,i}^T}{tr(X_{C,i}X_{C,i}^T)} \qquad \overline{R}_C = \frac{1}{n_C}\sum_{i \in C}^{n_C} R_{C,i} \tag{3}$$

where tr is the trace function and n_C is the number of epochs in C. let average power of class C be $\overline{P}_C$. Then, $\overline{P}_C$ is calculated as below:

$$\overline{P}_C = \frac{1}{n_C}\sum_{i\in C}^{n_C} \boldsymbol{w}^T X_{C,i} X_{C,i}^T \boldsymbol{w} = \frac{1}{n_C}\sum_{i\in C}^{n_C} \boldsymbol{w}^T R_{C,i}\boldsymbol{w} = \boldsymbol{w}^T \overline{R}_C \boldsymbol{w} \tag{4}$$

For two classes ($C = 1, 2$) case, CSP seeks for the maximum power ratio of the classes on the projected w axis. Thus, average power of one class is maximized while that of other class is minimized. In other words, spatial filter should maximize the following Rayleigh quotient problem [10]:

$$\arg\max_{\boldsymbol{w}} \frac{\boldsymbol{w}^T \overline{R}_1 \boldsymbol{w}}{\boldsymbol{w}^T \overline{R}_2 \boldsymbol{w}} \tag{5}$$

for any $\boldsymbol{w}$ that maximizes equation (5), denominator can be set to a constant value c by a scalar coefficient without changing the ratio.Thus, maximization of the Rayleigh quotient can be retranslated into a constrained optimization problem:

$$maximize \quad \boldsymbol{w}^T \overline{R}_1 \boldsymbol{w}, \qquad subject\, to \quad \boldsymbol{w}^T \overline{R}_2 \boldsymbol{w} = c \tag{6}$$

above constrained optimization problem can be solved by Lagrange multiplier method [16]:

$$L(\lambda, \boldsymbol{w}) = \boldsymbol{w}^T \overline{R}_1 \boldsymbol{w} - \lambda(\boldsymbol{w}^T \overline{R}_2 \boldsymbol{w} - c) \tag{7}$$

$$\frac{\partial L(\lambda, \boldsymbol{w})}{\partial \boldsymbol{w}} = 2\boldsymbol{w}^T \overline{R}_1 - \lambda(2\boldsymbol{w}^T \overline{R}_2) = 0 \tag{8}$$

where λ is Lagrage multiplier. Since $\overline{R}_C$ is a symmetric matrix, above equation can be written as a standard eigenvalue problem:

$$(\overline{R}_2^{\,-1} \overline{R}_1)\boldsymbol{w} = \lambda \boldsymbol{w} \tag{9}$$

According to Equation (9), w, which maximizes the Rayleigh quotient is the eigenvector corresponding to the largest eigenvalue of $(\overline{R}_2^{\,-1}\overline{R}_1)$.

CSP spatial filter $W_{CSP} \in \mathbb{R}^{dxN}$ matrice is constructed by taking $d = 2m$ ($d \leq N$) eigenvectors corresponding to the m largest and m smallest eigenvalues:

$$W_{CSP} = [\boldsymbol{w}_{\lambda_1}, ...\boldsymbol{w}_{\lambda_m} ......\boldsymbol{w}_{\lambda_{N-m+1}}, ...\boldsymbol{w}_{\lambda_N}]^T \tag{10}$$

where w_{λ_i} is the eigenvector corresponding to the eigenvalue λ_i. Any epoch $X_{C,i}$ is spatially filtered by,

$$Z_{C,i} = W_{CSP} X_{C,i} \tag{11}$$

where $Z_{C,i} \in \mathbb{R}^{dxT}$ is the spatially filtered signal. Band power (variance) is used as feature for classifier. For an epoch i, CSP feature vector is given by

$$\boldsymbol{fcsp}_{C,i}^k = log(\frac{var(Z_{C,i}^k)}{\sum_{l=1}^{2m} var(Z_{C,i}^l)}) \quad k = 1, 2, ...d \tag{12}$$

where $\boldsymbol{fcsp}^k_{C,i}$ is the k^{th} feature of feature vector $\boldsymbol{fcsp}_{C,i} \in \mathbb{R}^{dx1}$ that belongs to epoch i and $Z^k_{C,i}$ is k^{th} row of $Z_{C,i}$. Here, logarithm of variance ratio is calculated in order to approximate the distribution of the features to a normal distribution [11]. Next, features are used for training a linear classifier.

3 Regularized Common Spatial Patterns

CSP should be regularized in order to overcome its sensitivity to noise and over fitting [15]. Regularization of Rayleigh quotient given in Equation (5) is done by adding a penalty term to denominator. In this case, one should maximize spatial filters for each class seperately,

$$\boldsymbol{w}_1 = \underset{\boldsymbol{w}}{\operatorname{argmax}} \frac{\boldsymbol{w}^T \overline{R}_1 \boldsymbol{w}}{(1-\alpha)\boldsymbol{w}^T \overline{R}_2 \boldsymbol{w} + \alpha \boldsymbol{w}^T K_1 \boldsymbol{w}} \tag{13}$$

$$\boldsymbol{w}_2 = \underset{\boldsymbol{w}}{\operatorname{argmax}} \frac{\boldsymbol{w}^T \overline{R}_2 \boldsymbol{w}}{(1-\alpha)\boldsymbol{w}^T \overline{R}_1 \boldsymbol{w} + \alpha \boldsymbol{w}^T K_2 \boldsymbol{w}} \tag{14}$$

where $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are the spatial filters maximizing the variances of class 1 and 2 respectively. α is user defined regularization parameter to set the effect of penalty term. K_1 and K_2 are NxN penalty matrices. There are many variants about computation of K matrices in the literature. In the following subsections, five regularized CSP variants with proposed regularized CSP method will be reviewed.

3.1 Tikhonov Regularized CSP

In Tikhonov regularized CSP (TRCSP) [15], K matrices are set to identity matrices so that penalty term is equal to $\boldsymbol{w}^T \boldsymbol{I} \boldsymbol{w} = \boldsymbol{w}^T \boldsymbol{w} = \|\boldsymbol{w}\|^2$ Thus, solutions with large weights are penalized. TRCSP is expected to generate filters with small norm hence reducing the effect of artifacts and outliers. Note that, TRCSP penalizes all channels equally, a channel with highly motor imagery related activity may be penalized. However, there should not be a penalty if it contains useful information.

3.2 Weighted Tikhonov Regularized CSP

In TRCSP, each channel are penalized equally. However, it is known that some channels is more important than others. In Weighted Tikhonov Regularized CSP (WTRCSP),

$$K = diag(\boldsymbol{u}) \tag{15}$$

where $\boldsymbol{u}$ is a coefficient vector and $diag(\boldsymbol{u})$ is the diagonal matrix of $\boldsymbol{u}$ which has the information of the penalty level of each channel [15]. Since manual selection of u is not easy, it is computed by using the data from other test subjects:

$$\boldsymbol{u} = \left(\frac{1}{2N_f|\Omega|} \sum_{i \in \Omega} \sum_{f=1}^{2N_f} \left| \frac{\boldsymbol{w}_f^i}{||\boldsymbol{w}_f^i||} \right| \right)^{-1} \tag{16}$$

where $\boldsymbol{w}_f^i$ is the f^{th} spatial filter of subject i and $2N_f$ is the number of spatial filters used for each subject. Shortly, in WTRCSP penalty of each channel is adjusted by looking at other subjects. If average absolute value of the normalized weight of a channel is large in the CSP filters of other subject, penalty term assigned to this channel is small. This method needs more than one subjects in order to create a penalty function. Also, spatial filters of one subject can be contaminated by the outliers and noise in other subjects.

3.3 Invariant CSP

Invariant CSP (iCSP) method uses non task related EEG signal for building filters invariant to a given noise source [8]. Blankertz et. al. used disturbance covariance matrices from fluctuations in visual processing, parietal α-activity and used these covariance matrices as regularization matrices. Note that, there should be additional EEG measurements to compute the covariance matrix. For example, Samek et.al [10] recorded an extra session consisting of different eye movements,namely *eyes open*, *look left*, *look right*, *look up* and *look down* to generate an average covariance matrix K. Then adding $\boldsymbol{w}^T K \boldsymbol{w}$ as penalty term results in spatial filters which are invariant against changes generated by eye movements.

3.4 Spatially Regulaized CSP

Spatially regulaized CSP (SRCSP) (see Lotte and Guan [15]) deals with spatial relations between EEG channels. Normally CSP ignores positions and spatial relations of EEG electrodes. SRCSP method utilizes the information that neighboring neurons have similar functions, so that neighboring electrodes should measure similar brain signals. SRCSP method penalizes solutions with non-smooth filters in which spatially close electrodes should be similar weights in CSP spatial filter.

SRSCP uses the following regularization matrix K:

$$G(i,j) = exp(-\frac{1}{2}\frac{||\boldsymbol{v}_i - \boldsymbol{v}_j||^2}{r^2}) \tag{17}$$

$$K = D_G - G \tag{18}$$

where $\boldsymbol{v}_i$ is the position vector of i^{th} electrode, D_G is a diagonal matrix such as $D_G = \sum_j G(i,j)$ and r is a hyper parameter representing the maximum distance between two electrodes. SRCSP uses spatial information about electrode channels but ignores about penalizing the channels with non-task-related information.

3.5 Stationary CSP

Stationary CSP (sCSP) extends CSP to be invariant to non-stationaries in the data and regularizes CSP towards stationary subspaces. sCSP aims reducing the variations of the extracted features since it is assumed that variations like non-stationeries are the results of non task-related processes such as eye movements and electrode artefacts. sCSP penalizes the solutions with large variations in spatially filtered EEG data by analyzing training epochs. Generally, sCSP tries to minimize the following term in each class c:

$$D_c(\boldsymbol{w}) = \sum_k |\boldsymbol{w}^T R_c^{(k)} \boldsymbol{w} - \boldsymbol{w}^T \overline{R}_c \boldsymbol{w}| \tag{19}$$

where $R_c^{(k)}$ is the $k^t h$ trial of class c and $\overline{R}_c$ is the average covariance matrix of class c. For more information see the work of Samek et. al. [10]

3.6 Task Related & Spatially Regularized CSP

In this study, Task Related & Spatially Regularized CSP (TR&SR-CSP) is proposed. As mentioned before, each muscle group has a special area on the motor cortex, which supports the idea that each motor imagery task activates a special area on the motor cortex. We aim a regularizing method which emphasizes the electrodes that are spatially close to the center of imaging task being executed. TR&SR-CSP method regularizes the CSP spatial filter by directing it towards to center electrodes, which is assumed to record task-related signals. This is realized by penalizing the filter weights which are assumed as non-task related components with their spatial location. TR&SR-CSP designs regularizing matrices as below:

$$K_C(i,j) = \begin{cases} 1 - exp(-\frac{||\boldsymbol{v}_i - \boldsymbol{v}_c||^2}{r^2}), & i = j \\ 0, & i \neq j \end{cases} \tag{20}$$

where K_C is the regularizing matrice of class C, $\boldsymbol{v}_i$ is the spatial location of i^{th} electrode, $\boldsymbol{v}_c$ is the electrode which is assumed to be the center electrode of task related signal and r is a hyper parameter representing the maximum distance between two electrodes. Note that in TR&SR-CSP, regularizing K matrix is different for each class. Also spatial smoothness of filter is ensured by exponential function. Figure 2 illustrates penalty matrices on scalp figures with topographic maps. Diagonal values of K matrices for the two tasks (*left hand* and *foot*) has been plotted.

Despite similarities to SRCSP and WTRCSP, TR&SR-CSP method differs from SRCSP by incorporation of task relevant information when computing K. WTRCSP ignores spatial relations between electrodes but obtains a general spatial filter by analyzing data from other subjects, while TR&SR-CSP builds its regularizing matrices by using the information in the literature about mapping motor cortex and taking account of spatial relations between electrodes.

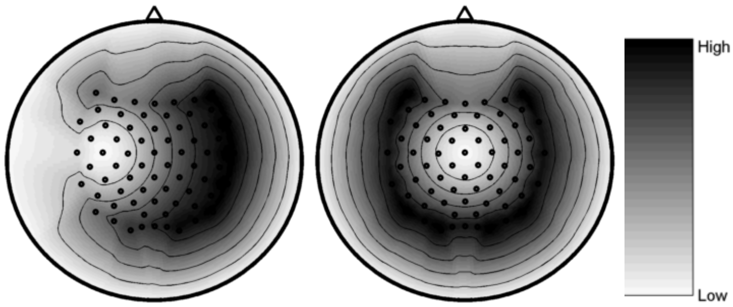

Fig. 2. Illustration of penalty values for *left hand* and *foot* ($r = 0.3$)

4 Evaluation

4.1 EEG Data Set Used

In this study, we used a publicly available EEG Motor Imagery data set, BCI competition III Data Set IVa [13]. This data set involves EEG recordings of 5 subjects who performed motor imagery of right hand and foot. 118 electrodes were used for recording EEG with the sample rate of 100Hz. There are 280 trials for each subject. However, number of training and test sets differ for each subject: 168, 224, 84, 56 and 28 trials are sizes of the training sets for subjects labelled as *aa*, *al*, *av*, *aw* and *ay* respectively and remaining trials form the test set. We used EEG electrodes which roughly cover the motor cortex, totally 68 electrodes were used (F*, FFC*, FC*, CFC*, C*, CCP*, CP*, PCP*, P*, PPO* and PO*, with numbers higher than 6 according to the International 10-20 system are discarded) as done in [10].

4.2 Preprocessing, Hyper Parameters and Classification Method

For each trial in the data set and for each CSP method used in evaluation, we applied the same preprocessing steps. i) EEG signal is band bass filtered with 0.5 - 30 Hz 5th order Butterworth filter. ii) For each trial, we used EEG signals in time segment between 0.5s - 2s after instruction cue. iii) Three spatial filter pairs were used ($m = 3$) as recommended in [7]. In TR&SR-CSP method, center electrodes are selected as C3 and CZ for right hand and foot motor imageries, respectively. For hyper parameters, values were selected from the sets [0.01, 0.025, 0.05, 0.075, 0.1, 0.25, 0.5, 0.75] for α and [0.1, 0.25, 0.5, 0.75, 1, 1.25, 1.5] for r. The parameter values which maximize the 2-fold cross-validation accuracy for the training set were selected as optimal hyper parameters. All covariance (R_C) and regularization (K_C) matrices were normalized by dividing them by their traces. For stationary CSP (sCSP) method, we compute the covariances for single trial period (i.e., chunk size (ν) was set to 1). We used Linear discriminant analysis (LDA) as classifier.

4.3 Performance Comparison

We tested TR&SR-CSP method along with conventional CSP method described in Section 2 and other regularizing methods in Section 3. Note that we couldn't test iCSP method because it requires an additional EEG measurement session. Table 1 shows the average error rates of CSP, TRCSP, WTRCSP, SRCSP, sCSP and TR&SR-CSP methods for the BCI Competition III Data Set IVa. According to the results, TR&SR-CSP method outperforms other algorithms. TR&SR-CSP improved standard CSP algorithm except subjects *al* and *aw*. For those subjects, TR&SR-CSP algorithm performs similar to CSP. The largest improvement achieved for subjects *av* and *ay*. Subject *av* has lower classification performance with all of the methods. This can be because of low signal-to-noise ratio and artefacts for *av*. Subject *ay* has small training set (56 trials) which causes CSP to over fit. Since TR&SR-CSP method leads spatial filters towards to motor cortex area defined by physiological facts, performances of *av* and *ay* are improved.

Table 1. Classification error rates for each subject in BCI competition III Data Set IVa for classical CSP and regularized CSP algorithms. The best result for each subject is given in bold.

	SUBJECTS					OVERALL		
Method	AA	AL	AV	AW	AY	Mean	Med.	Std
CSP	26.8	**1.8**	32.7	20.5	25.8	21.5	25.8	11.8
TRCSP	30.9	1.8	32.7	26.7	17.9	22.0	26.7	12.7
WTRCSP	22.1	1.8	33.0	23.5	16.2	19.3	22.1	11.5
SRCSP	31.9	4.6	34.5	29.6	19.0	23.9	29.6	12.3
sCSP	26.0	**1.8**	35.7	**18.2**	27.9	21.9	26.0	12.9
TR&SR-CSP	**20.7**	**1.8**	**26.9**	20.6	**13.0**	**16.6**	**20.6**	9.6

We also tested the performance of our method with Wilcoxon Signed-Rank Test. We ran each method 30 times for each subject. In Table 2, TR&SR-CSP method and other methods are compared one by one. The null hypothesis H_0 for this test is: There is no difference between the median of performance obtained by method A and the median of performance obtained by method B for same benchmark problem, To determine whether method A outperforms method B when the null hypothesis fails, the sizes of the ranks provided by the Wilcoxon Signed-Rank Test (i.e. T+ and T- as defined in [17]) were examined. In Table 2, '+' indicates TR&SR-CSP is better ($p < 0.05$ and $T^+ < T^-$), '=' is indicates there is no statistical difference between those two methods ($p >= 0.05$) and '-' indicates opponent algorithm is better than TR&SR-CSP ($p < 0.05$ and $T^+ > T^-$). The last column of the table shows the total count of (+/=/-) signs for each subject.

We also illustrate the effect of regularizing on the spatial filters. To this end, channel weights of the most important spatial filter of each class (i.e. first and

Table 2. Wilcoxon signed rank test at the %95 significance level for comparison of TR&SR-CSP and other CSP methods

TR&SR-CSP vs		AA	AL	AV	AW	AY	SCORES
CSP	p	0.000	1.000	0.000	0.577	0.000	
	$T+/T-$	0 / 465	0 / 0	0 / 465	183 / 142	0 / 465	
	winner	+	=	+	=	+	3/2/0
TRCSP	p	0.000	1.000	0.000	0.000	0.000	
	$T+/T-$	0 / 465	0 / 0	15 / 450	0 / 465	0 / 465	
	winner	+	=	+	+	+	4/1/0
WTRCSP	p	0.000	1.000	0.000	0.000	0.000	
	$T+/T-$	21 / 330	0 / 0	33 / 374	17 / 449	0 / 465	
	winner	+	=	+	+	+	4/1/0
SRCSP	p	0.000	0.000	0.000	0.000	0.000	
	$T+/T-$	0 / 465	0 / 136	3 / 404	0 / 465	0 / 465	
	winner	+	+	+	+	+	5/0/0
sCSP	p	0.000	1.000	0.000	0.000	0.000	
	$T+/T-$	0 / 465	0 / 0	3 / 462	438 / 28	0 / 465	
	winner	+	=	+	-	+	3/1/1

last rows of W_{CSP}) are topographically mapped onto a scalp and colorized (white to black). Figure 3 shows spatial filters obtained by different CSP algorithms by interpolating the whole scalp by using the weights at the corresponding electrode positions. Note that figures illustrate the absolute values of the weights, sign of each weight is ignored. CSP and other RCSP filters are complicated and have peaks at unexpected locations while spatial filters obtained by TR&SR-CSP method physiologically more relevant, with stronger weights over the motor cortex area relevant to task being executed [5]. This suggests proposed method has the ability to produce more generalized filters and avoids obtained spatial filters being over fitted.

5 CONCLUSION

In this paper, we proposed a new Regularized CSP method named TR&SR-CSP. We revisited conventional CSP algorithm and existing RCSP algorithms. We evaluated 6 different CSP methods (1 CSP method, 4 other RCSP methods and 1 proposed TR&SR-CSP method) on the BCI competition EEG data from 5 subjects [13]. Results show that TR&SR-CSP method clearly outperforms other methods. We also illustrated the spatial filters on a scalp map with color coded and topographically mapped. Proposed methods physiological plausibility can be clearly seen in these figures. Future works could test our method on more databases and new EEG motor imagery task experiments. Proposed methods generalizing capability suggests minimal user training, it could be also investigated with small training sets. Also we could adapt our algorithm to multi class cases.

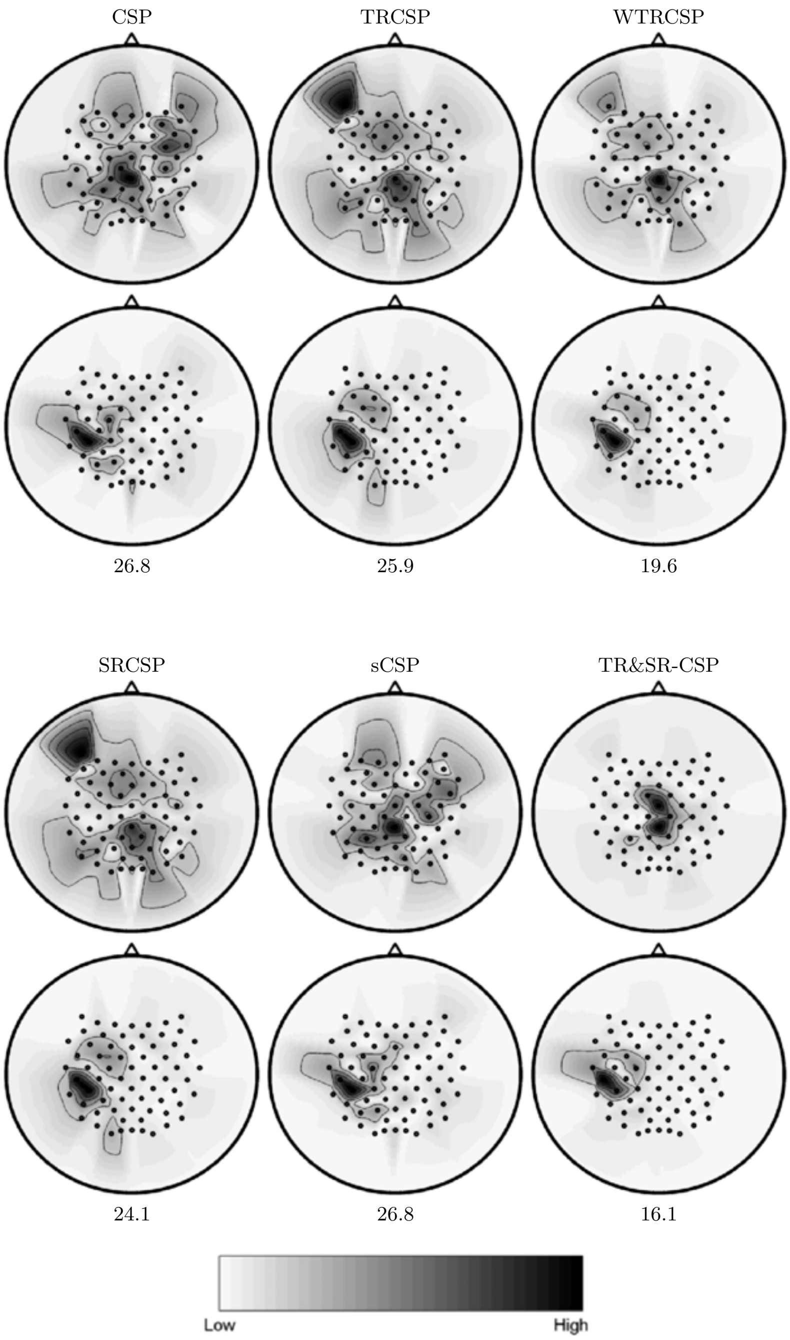

Fig. 3. Illustration of CSP filters for subject *aa*. Upper and lower rows represent the first and second MI tasks, respectively. Hyper parameters are $\alpha = 0.05$ and $r = 1$. Error rates for each method are given at the bottom.

Proceedings IWBBIO 2014. Granada 7-9 April, 2014 52

References

1. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain–computer interfaces for communication and control. Clinical neurophysiology **113**(6) (2002) 767–791
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**(6) (1997) 642–651
3. Pfurtscheller, G., Lopes da Silva, F.: Event-related eeg/meg synchronization and desynchronization: basic principles. Clinical neurophysiology **110**(11) (1999) 1842–1857
4. Pfurtscheller, G., Aranibar, A.: Event-related cortical desynchronization detected by power measurements of scalp eeg. Electroencephalography and clinical neurophysiology **42**(6) (1977) 817–826
5. Penfield, W., Boldrey, E.: Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain: A journal of neurology (1937)
6. Sage, G.: Introduction to motor behavior: a neuropsychological approach. Addison-Wesley series in physical education. Addison-Wesley Pub. Co. (1971)
7. Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., Muller, K.R.: Optimizing spatial filters for robust eeg single-trial analysis. Signal Processing Magazine, IEEE **25**(1) (2008) 41–56
8. Blankertz, B., Kawanabe, M., Tomioka, R., Hohlefeld, F., Müller, K.r., Nikulin, V.V.: Invariant common spatial patterns: Alleviating nonstationarities in brain-computer interfacing. In: Advances in neural information processing systems. (2007) 113–120
9. Ramoser, H., Muller-Gerking, J., Pfurtscheller, G.: Optimal spatial filtering of single trial eeg during imagined hand movement. Rehabilitation Engineering, IEEE Transactions on **8**(4) (2000) 441–446
10. Samek, W., Vidaurre, C., Müller, K.R., Kawanabe, M.: Stationary common spatial patterns for brain–computer interfacing. Journal of Neural Engineering **9**(2) (2012) 026013
11. Falzon, O., Camilleri, K.P., Muscat, J.: The analytic common spatial patterns method for eeg-based bci data. Journal of Neural Engineering **9**(4) (2012) 045009
12. Blankertz, B., Muller, K.R., Curio, G., Vaughan, T.M., Schalk, G., Wolpaw, J.R., Schlogl, A., Neuper, C., Pfurtscheller, G., Hinterberger, T., et al.: The bci competition 2003: progress and perspectives in detection and discrimination of eeg single trials. Biomedical Engineering, IEEE Transactions on **51**(6) (2004) 1044–1051
13. Blankertz, B., Muller, K.R., Krusienski, D.J., Schalk, G., Wolpaw, J.R., Schlogl, A., Pfurtscheller, G., Millan, J.R., Schroder, M., Birbaumer, N.: The bci competition iii: Validating alternative approaches to actual bci problems. Neural Systems and Rehabilitation Engineering, IEEE Transactions on **14**(2) (2006) 153–159
14. Reuderink, B., Poel, M.: Robustness of the common spatial patterns algorithm in the bci-pipeline. (2008)
15. Lotte, F., Guan, C.: Regularizing common spatial patterns to improve bci designs: unified theory and new algorithms. Biomedical Engineering, IEEE Transactions on **58**(2) (2011) 355–362
16. Bertsekas, D.: Constrained optimization and Lagrange multiplier methods. Optimization and neural computation series. Athena Scientific (1996)
17. Civicioglu, P.: Backtracking search optimization algorithm for numerical optimization problems. Applied Mathematics and Computation **219**(15) (2013) 8121–8144