

# Partitioned Common Spatial Pattern Method for single trial EEG Signal classification in Brain-Computer Interface System

DOI 10.7305/automatika.2016.07.1078  
UDK 004.5:[004.932'12:616.831-073-71]

Original scientific paper

Common spatial pattern (CSP) method is highly successful in calculating spatial filters for motor imagery-based brain-computer interfaces (BCIs). However, conventional CSP algorithm is based on a single wide frequency band with a poor frequency selectivity which will lead to poor recognition accuracy. To solve this problem, a novel Partitioned CSP (PCSP) algorithm is proposed to find the most relevant spatial frequency distribution with motor imaginary, so that the algorithm has flexible frequency selectivity. Firstly, we partition the dataset into frequency components using a constant-bandwidth filters bank. Then, a features selection method based on the Bhattacharyya distance is adopted for PCSP features ranking, selection and evaluation. Subsequently, the PCSP features are used to obtain scores which reflect the classification capability and being used for EEG signal classification. The experimental results on 4 subjects showed that the PCSP method significantly outperforms the other two existing approaches based on conventional CSP and Common Spatio-Spectral Pattern (CSSP).

**Key words:** Partitioned CSP, Motor Imagery, Brain-Computer Interface, Single trial classification

**Metoda raspodijeljenog zajedničkog prostornog uzorka za klasifikaciju EEG signala sučelja mozak-računalo u jednoj procjeni.** Metoda zajedničkog prostornog uzorka (eng. common spatial pattern, CSP) je vrlo uspješna u izračunu prostornih filtara za sučelja mozak-računalo zasnovana na motoričkoj predodžbi (eng. brain-computer interface, BCI). Međutim, konvencionalni CSP algoritam je zasnovan na jednom širokom pojasu frekvencija s lošom selektivnosti frekvencija što rezultira manjom točnošću prepoznavanja. Za rješavanje navedenog problema u ovom radu je predložen novi raspodijeljeni CSP algoritam za pronalaženje najznačajnije prostorno frekvencijske distribucije s motoričkom predodžbom, sa svojstvima fleksibilne selektivnosti frekvencije. Početna faza metode je podjela podataka na frekvencijske komponente korištenjem filtarskog sloga s konstantnom širinom pojasa. Potom, prilagođena je metoda odabira svojstava zasnovana na Bhattacharyya udaljenosti za rangiranje, odabir i evaluaciju PCSP svojstava. Zatim, PCSP svojstva se koriste za dobivanje ocjena koje reflektiraju mogućnosti klasifikacije te za klasifikaciju EEG signala. Eksperimentalni rezultati na 4 ispitanika pokazali su da PCSP metoda po performansama značajno nadmašuje druga dva postojeća pristupa zasnovana na konvencionalnom CSP-u i zajedničkom prostor-spektralnom uzorku (eng. common spatio-spectral pattern, CSSP).

**Ključne riječi:** podijeljeni CSP, motorička predodžba, sučelje mozak-računalo, klasifikacija u jednoj procjeni

## 1 INTRODUCTION

Nowadays, the most popular brain signal used for BCI is the scalp-recorded electroencephalogram (EEG), because it is a noninvasive measurement of brain electrical activities and has a high temporal resolution [1]. EEG-based BCI systems enable a subject, such as a disable person, to send commands for controlling a computer application or an electronic device such as a telephone only by means of brain activity [2]. To achieve this aim, using classification algorithms is the most popular approach [3–5]. These algorithms are used to identify patterns of brain activity [6]. Among various exiting patterns, the event-related de-synchronization/synchronization (ERD/ERS) patterns

during motor imagery are widely used. The CSP algorithm [7] is highly successful in calculating spatial filters for detecting ERD/ERS effects [8]. CSP is a decomposition method that finds a set of spatial patterns which are well suited to discriminate between different mental states induced by motor imagery as they focus on the synchronization and de-synchronization effects occurring over different locations of the sensorimotor cortex after performed motor imagery. It is designed to find a set of spatial patterns which maximize the power/variance ratios of the filtered signals between the two classes. It can be calculated by simultaneously diagonalize the covariance matrices corresponding to two classes of data. For the classifi-

cation of two classes of motor imageries, CSP can achieve the average accuracy above 90% on single trial EEG measurements [9]. However, this spatial filter is highly sensitive to noise and artifacts like eye blinks or loose electrodes [10], and it must only be applied to the informative frequency bands (mu and beta rhythms) i.e. in the 8-30 Hz range, which is specific to each subject because of the fact that neurophysiologically the discriminative band of ERD varies from one subject to another [8]. If utilizing CSP algorithm to the EEG data with a poor frequency bands selection will lead to a poor classify accuracy. Moreover, the performance of CSP severely depends on the preprocessing procedure of the temporal filtering, because CSP detects the changes of rhythmic activities based on the variances of signals. Only having the EEG signals band-pass filtered through the frequency domain of interest, high or low signal variances could reflect a strong or weak rhythmic activity respectively [11]. Additionally, The CSP method is prone to over-fitting resulting from simultaneously diagonalization of covariance matrices, which is a typical problem if there is only a small training set, and if there are a large number of channels [12].

To overcome the limitation of conventional CSP, several variants of CSP have been proposed to improve the robustness and discriminativity of the extracted features by applying regularization, incorporating data from other sessions/subjects, or using robust estimators [5]. For instance, the authors of [13, 14] regularized the covariance matrix to increase robustness, especially in small-sample settings. The authors of [15] present a way to robustify the popular common spatial patterns (CSP) algorithm under a max-min approach. They show that this kind of max-min optimization makes CSP robust to outliers and reduces its tendency to over-fit. The authors of [11] suggested an extension of CSP called Common Spatio-Spectral Pattern (CSSP) to the state space, which utilizes the method of time delay embedding which allows for individually tuned frequency filters at each electrode position and, thus, yields an improved and more robust machine learning procedure. Other authors [16, 17] improve the effectiveness of the solution by preserving the temporal relationship among samples of unlabeled trials. All these different methods were proposed for specific applications scenario with those own optimization strategy.

In this paper, we propose a novel Partitioned Common Spatial Pattern (PCSP) algorithm for single trial EEG classification. The proposed algorithm aims to find the most reliable spatial frequency distribution of the motor-imagery related neurophysiological phenomena thus achieving a higher level of subject-specific adaptation. For that we first partitioned the EEG dataset into frequency components covering the range represented by the 5-33 Hz band using a bank of constant-bandwidth Butterworth fil-

ters. Then, a features selection method based on the Bhattacharyya distance is adopted for PCSP features ranking, selection and evaluation. Subsequently, the PCSP features are used to obtain scores which reflect the classification capability and being used for EEG signal classification. The performance of the proposed algorithm is evaluated on data collected from 4 subjects performing motor imagery task. The classification accuracies of the proposed algorithm are compared with the results from two existing algorithms, namely, CSP and CSSP.

This paper is organized as follows. Section 2 introduces the mathematical background of CSP and its extension PCSP. In Section 3, we introduce the experiment setup and data acquisition. The results and data analysis are presented in Section 4. Finally, Section 5 draws the conclusion.

## 2 PARTITIONED COMMON SPATIAL PATTERN

### 2.1 The mathematical background of CSP

The purpose of Common Spatial Pattern is to design spatial filters that lead to new time series whose variances are optimal for the discrimination of two classes of EEG. Details of the algorithm will be described in the following with the example of discriminating left hand vs. right hand imaginary. The filtered signal corresponding to the desynchronization of the left hand motor cortex is characterized by a strong motor rhythm during imagination of right hand movements, and by an attenuated motor rhythm during left hand imagination. This criterion is exactly what the CSP algorithm optimizes: maximizing variance for the class of right hand trials and at the same time minimizing variance for left hand trials.

Let  $H_L$  and  $H_R$  denote the corresponding EEG matrices under the two conditions (left hand and right hand) with dimensions  $N \times M$ , where  $N$  is the number of selected channel, and  $M$  is the number of samples in each trial. The normalized spatial covariance of the EEG can be calculated as:

$$X_L = \frac{H_L H_L^T}{\text{tr}(H_L H_L^T)}, X_R = \frac{H_R H_R^T}{\text{tr}(H_R H_R^T)}, \quad (1)$$

where  $\text{tr}$  is the trace operator that sums up the diagonal elements of a matrix, and  $T$  denotes the transpose operator of a matrix. The final spatial covariances  $\overline{X}_L$  and  $\overline{X}_R$  are respectively computed by averaging over the trials under each condition. The composite spatial covariance matrix is defined as:

$$X = \overline{X}_L + \overline{X}_R. \quad (2)$$

As  $X$  is a symmetrical matrix, it can be factored into its eigenvectors by singular value decomposition:

$$X = \overline{X}_L + \overline{X}_R = R_0 \lambda_0 R_0^T, \quad (3)$$

where  $R_0$  is the matrix of eigenvectors and  $\lambda_0$  is the diagonal matrix of eigenvalue. Note that the eigenvalues are assumed to be sorted in a descending order. We have the whitening transformation matrix:

$$P = \sqrt{\lambda_0^{-1}} R_0^T. \quad (4)$$

The individual covariance matrices  $\overline{X}_L$  and  $\overline{X}_R$  are transformed to:

$$U_L = P \overline{X}_L P^T, U_R = P \overline{X}_R P^T. \quad (5)$$

$U_L$  and  $U_R$  share common eigenvectors and the sum of corresponding eigenvalues for the two matrices will always be one:

$$U_L = U \lambda_L U^T, U_R = U \lambda_R U^T, \lambda_L + \lambda_R = I, \quad (6)$$

where  $I$  is the identity matrix. Because the sum of two corresponding eigenvalues is always one, the eigenvector with the largest eigenvalue for  $U_L$  has the smallest eigenvalue for  $U_R$  and vice versa. This property makes the eigenvectors  $U$  useful for classification of the two distributions. The projection of whitened EEG onto the first and last eigenvectors in  $U$  will give feature vectors that are optimal for discriminating two populations of EEG in the least squares sense. With the projection matrix:

$$W = U^T P. \quad (7)$$

The decomposition (mapping) of a trial  $E$  can be transformed into the uncorrelated components:

$$Z = W E. \quad (8)$$

$Z$  can be thought as EEG source components including common and specific components of different tasks. The original EEG  $E$  can be reconstructed by:

$$E = W^{-1} Z, \quad (9)$$

where  $W^{-1}$  is the inverse matrix of  $W$ . The columns of  $W^{-1}$  are the common spatial patterns which can be regarded as the time-invariant vectors of EEG source distribution vectors. Fig.1 shows the four most significant spatial patterns extracted by CSP method for subject A.

## 2.2 Partitioned Common Spatial Pattern

In general, CSP algorithm was applied on a wide frequency band thus finding an optimal spatial filter Pattern based on the simultaneous diagonalization active employee of two covariance matrices, which frequency selectivity is not flexible and need more channels. Now we propose a novel PCSP (Partitioned Common Spatial Pattern), which

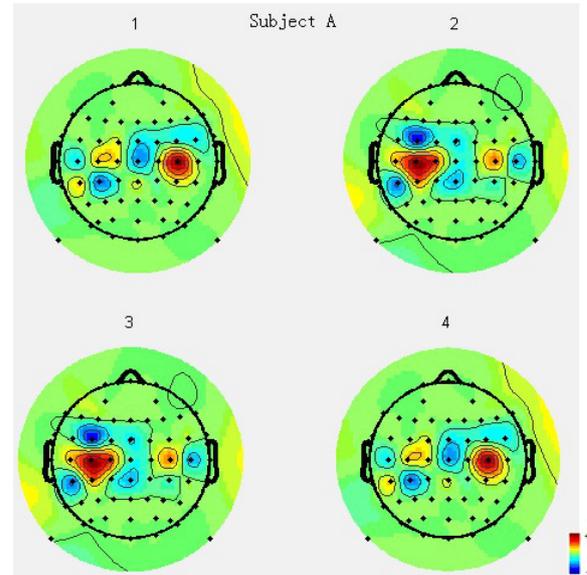


Fig. 1. Four most significant spatial patterns extracted by CSP.

applying the CSP method simultaneously in every partitioned frequency and find the most relevant spatial frequency distribution with motor imaginary, so that the algorithm has flexible frequency selectivity. Moreover, it can achieve good performance for a multi-class classification problem.

Details of the PCSP algorithm will be described in the following with the example of discriminating left hand vs. right hand vs. non imaginary.

### 2.2.1 Frequency Decomposition

The EEG data were digitized at a sampling frequency of 500 Hz and band-pass filtered the data on different and overlapped frequency bands between 5 Hz and 33 Hz using a bank of constant-bandwidth Butterworth filters. Butterworth filter is best suited for the closed loop gain to be as close to 1 as possible within the pass-band. Roll-offs become steeper, they approach the ideal filter more closely. Like all filters, the typical Butterworth filter is the low pass filter, which can be modified into a high-pass filter, or placed in series with others to form band-pass and band-stop filters, and higher order versions of these. Here we used a bank of fifth-order Butterworth filters with 6 Hz bandwidth which centred on the frequency of interest. The structure of the adopted filters bank is shown in the Fig.2. Each filter of the bank is identified by its central frequency  $k$  corresponding to a band-pass filter with cut-off frequencies  $[k - 3; k + 3]$ . In this way, the EEG data was decomposed into its frequency components  $Z_K(m, c)$ , Where  $K$

represents the filter identifier,  $m$  denotes the length of the data in samples and  $c$  is the number of data channel.

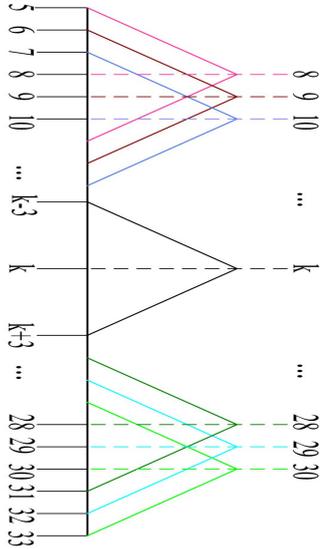


Fig. 2. The structure of the adopted filters bank

### 2.2.2 CSP Extension to Multi-class

The conventional CSP algorithm can handle only binary classification, Multi-class classification needs to extend CSP algorithm to multi-class CSP.

Multi-class extensions of CSP algorithm can be obtained from the following three strategies [18]:

a. One versus one (OVO): This algorithm reduces a multi-class classification problem to several binary problems. Calculated the spatial patterns extracted by CSP method, and then combined all the spatial patterns as the multi-class spatial patterns. This algorithm results in high dimension of feature extraction coefficient by translate N-class classification problem to  $N \times (N - 1)/2$  binary problems.

b. Simultaneous diagonalization (SIM): In the binary case, the CSP algorithm finds a simultaneous diagonalization of both covariance matrices whose eigenvalues sum to one. Thus a possible extension to many classes, i.e., many covariances  $(\sum_i)_{i=1, \dots, N}$  is to find a matrix  $R$  and diagonal matrices  $(D_i)_{i=1, \dots, N}$  with elements in  $[0; 1]$  and with  $R \sum_i R^T = D_i$ , for all  $i = 1, \dots, N$ ,  $\sum_{i=1}^N D^i = I$ . But this method can be done exactly for  $N = 2$ ; for  $N > 2$ , in general, only approximative solutions can be obtained.

c. One versus the rest (OVR): By computing spatial patterns for each class against all others, it translates N-class problem into N new two-class problems. The OVR approach appears rather similar to the OVO approach, but

there is in fact a large practical difference. OVR does multi-class classification on all projected signals whereas OVO does binary classification on the CSP patterns according to the binary choice.

In this work, we will use OVR and OVO methods to decompose multi-class classification problem (left hand imagination, right hand imagination and non imagination) into the combination of multiple binary sub-problems. Let  $H_L$ ,  $H_R$  and  $N$  denote the left hand imagination, right hand imagination and non imagination. Then the multi-class problem of identifying a state among  $H_L$ ,  $H_R$  and  $N$  are decomposed four binary sub-problems ( $N - H_L H_R$ ,  $H_L - H_R$ ,  $H_L - N$ ,  $N - H_R$ ).

### 2.2.3 Selection of the optimal PCSP filters

In this step, the extraction of brain rhythm topographic patterns by CSP is performed on each partitioned frequency component of EEG signal. Then specific CSP filters are calculated for each binary sub-problem and for each frequency component  $Z_K$ . Thus, the PCSP filters at the  $k - th$  frequency component is in the form of

$$W_k^{sp} = [w_{1,k}^{sp}, w_{2,k}^{sp}, \dots, w_{c,k}^{sp}], \quad (10)$$

where  $c$  denotes the number of data channel,  $k$  is the  $k - th$  frequency component,  $sp$  is the vector of the binary sub-problems,  $sp = [N - H_L H_R, H_L - H_R, H_L - N, N - H_R]$ . The eigenvectors of  $W_k^{sp}$  are sorted by decreasing corresponding eigenvalues.

At this point, we can get the optimal PCSP vector  $Y_k^{sp}$  by applying projection transforms  $W_k^{sp}$  to the frequency component  $Z_K$  and  $Y_k^{sp}$  should contains information related to the  $k - th$  frequency component useful at solving the spatial sub-problem  $sp$ .

### 2.2.4 Feature selection based on the Bhattacharyya distance

The Bhattacharyya distance has been used as a class separability measure for feature selection and is known to provide the upper and lower bounds of the Bayes error [19]. It is considered to be more reliable than the Mahalanobis distance, as the Mahalanobis distance is a particular case of the Bhattacharyya distance when the standard deviations of the two classes are the same. Therefore, when two classes have similar means but different standard deviations, the Mahalanobis distance would tend to zero, however, the Bhattacharyya distance would grow depending on the difference between the standard deviations.

For two normally distributed classes, the Bhat-

tacharyya distance is defined as follows:

$$D_B = \frac{1}{8}(\mu_2 - \mu_1)^T \left[ \frac{\sum_1 + \sum_2}{2} \right]^{-1} (\mu_2 - \mu_1) + \frac{1}{2} \ln \left( \frac{[\frac{\sum_1 + \sum_2}{2}]}{\sqrt{|\sum_1| |\sum_2|}} \right), \quad (11)$$

where  $D_B$  is the Bhattacharyya distance between classes,  $\mu_i$  and  $\sum_i$  are the mean vector and covariance matrix of class  $i$ , respectively. This equation gives the class separability due to the difference between class covariance matrices and thus guaranteeing the comparison of the two distributions shape rather than just their means. Furthermore, the optimal Bayes classification error between the two classes is bounded by the following expression:

$$\varepsilon \leq \sqrt{p_1 p_2} \exp(-D_B), \quad (12)$$

where  $p_i$  is a priori probability of class  $i$ . We will refer to the upper bound of the error probability evaluated from the inequality (12) with  $p_1 = p_2 = 0.5$ , as the Bhattacharyya error,  $\varepsilon_B$ . That is,

$$\varepsilon_B = 0.5 \cdot \exp(-D_B). \quad (13)$$

By setting the two prior probabilities equal, the two terms  $\varepsilon_B$  and  $D_B$  are equivalent in that both indicate the intrinsic separability of the two distributions, regardless of their prior probabilities. In summary, advantages of using the Bhattacharyya distance are that it is computationally very simple and that, since it is derived from an error bound rather than just from an exact solution, it provides a “smoothed” distance between the two classes in study, which is more appropriate since real-life data usually do not fit truly normal distributions.

In this paper, we applying the feature extraction algorithm in each  $Y_k^{sp}$  and calculating a binary class separability measure based on the Mahalanobis distance to identify which  $Y_k^{sp}$  contain the most useful information regarding each sub-problem  $sp$ . Then the optimal spatial frequency features were selected according to the highest separability measure. The spatial frequency confidence maps were utilized to indicate the distribution of information content. Let us take the example of the spatial frequency confidence map for the  $H_L - N$  binary sub-problem. The values of information content were coded as the gray levels (white for the highest and black for the lowest) in the spatial frequency confidence map (checkerboard). As shown in Fig.3, the color cluster of confidence map will fade to black when the gray level is close to zero which indicates the spatial frequency transformation does not lead to any useful information in discriminating the two classes. While the higher of gray level, the more discriminatory capability of the projection. Two red circles were used to mark the

highest separability measure namely the most relevant spatial frequency filters. We can see that the frequencies focus on mu rhythm (8-14Hz) and beta rhythm (15-30Hz). This phenomenon is coincident with both the event-related synchronization and de-synchronization phenomena in motor imagery.

In our feature selection scheme, we find the relevant frequency component in the mu and beta rhythms for each binary sub-problem. Also we are interested only in those spatial filters associated with the highest and lowest eigenvalues for each binary sub-problem.

Based on the aforementioned study, we can separate the whole search-space into 4 sub-spaces as shown in Fig. 4. The spatial frequency filters were sort according to the eigenvalues. For each of the identified sub-spaces only one cluster with the highest binary separability measure is selected. The green-highlighted clusters correspond to the most relevant spatial frequency filters.

In this way, we can find the optimal spatial frequency filters by applying the above selection scheme to each search-space  $sp$ .

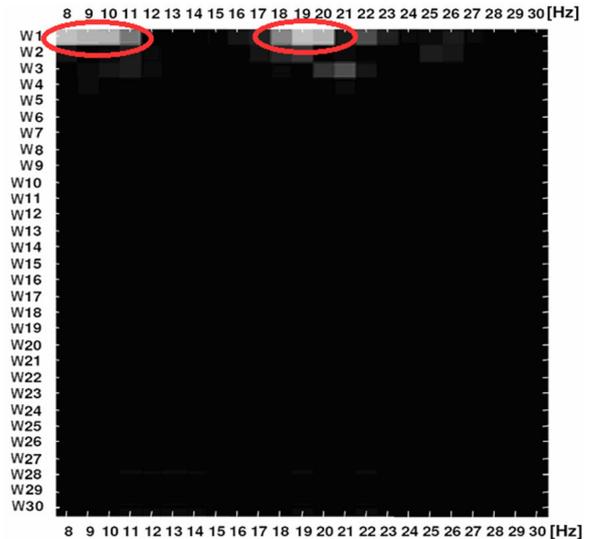


Fig. 3. The spatial frequency confidence map for the binary sub-problem.

### 3 EXPERIMENTAL SETUP AND DATA ACQUISITION

Subjects participated in the experimental study were four male students of Shandong University of Science and Technology and they were aged between 21 and 30, right-handed. All subjects had normal or corrected-to-normal vision. They all gave informed consent as approved by the Ethics Committee.

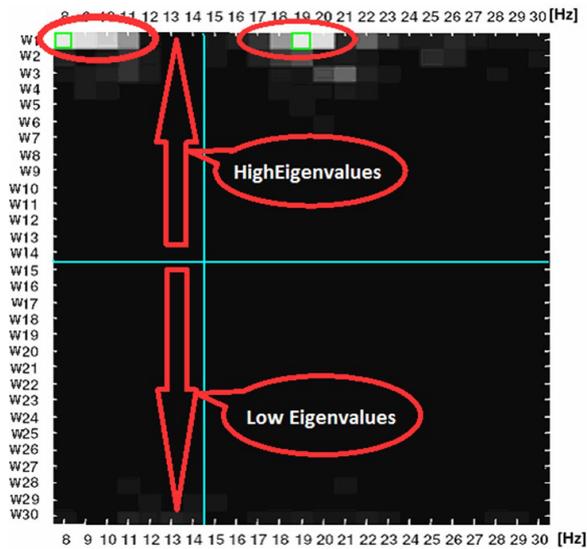


Fig. 4. The spatial frequency confidence map of search-space.

Subjects were asked to sit in an armchair with two hands relaxing, and looked at a 17" computer monitor approximately 1 m in front of the subject at their eye level. 62 channels of EEG signals were recorded in a shielded room by a 64 channel high-resolution EEG/ERP Systems (SynAmps2, Neuroscan) in our Lab using the following channels located at the positions of the 10-20 international electrode-positioning standard: FP1,FPZ, FP2, AF3, AF4, F7, F5, F3, F1, FZ, F2, F4, F6, F8, FT7,FC5, FC3, FC1, FCZ, FC2, FC4, FC6, FT8, T7, C5, C3,C1,CZ,C2, C4, C6, T8, TP7, CP5, CP3, CP1, CPZ, CP2, CP4, CP6,TP8, P7, P5, P3, P1, PZ, P2, P4, P6, P8, PO7, PO5, PO3, POZ,PO4, PO6, PO8, CB1, O1, OZ, O2 and CB2. Skin-electrode junction impedances were maintained below 5 k  $\Omega$ . Signals were digitized at a sampling frequency of 500 Hz and band-pass filtered between 5 Hz and 33 Hz. The data collection procedure has three stages: (1) Subject preparation, (2) Training data collection and (3) Test data collection. The paradigm required the subject to control a cursor moving on the monitor by imagining the movements of his right hand, left hand or null for 2 seconds with a 4 second break between trials. For each subject, the data were collected over two sessions with a 15 minute break. The first session was conducted without feedback, and 60 trials (20 trials for each class) obtained in this session were used for training and analysis. 150 trials (50trials for each class) in the next session were taken as testing data to give online feedbacks. Fig.5 shows the on-line feedback paradigm of the three-class motor imagery tasks. Three green arrows represent tasks corresponding to left hand, right hand, and non imagination, respectively. The progress bars provide

real-time visual feedback.

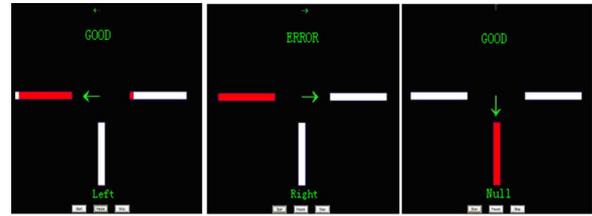


Fig. 5. The on-line feedback paradigm of motor imagery tasks

#### 4 RESULTS AND DATA ANALYSIS

In this section, the proposed PCSP-based scheme is applied to the datasets, and the experimental results are presented respectively. For performance evaluation, the error classification rate is used to measure the classification accuracy.

The raw EEG data were preprocessed by overlapping sliding window technology. Each EEG data segment with class labels was divided into many smaller data segments which have the same class labels. The sliding window keeps the same length during sliding. In this way, the EEG data segments with the length of 1 second, 2 seconds, 3 seconds and 4 seconds were obtained to verify the performance of the proposed algorithm in different length data segments.

Fig. 6 and Table 1 compare the 10-folds cross validation accuracies (mean error rate and standard deviation of the error) of four subjects obtained by CSP, CSSP and the proposed PCSP algorithms in binary classification (Left hand and Right hand imagination). The results show that the proposed PCSP algorithm yielded lowest test error rate of  $2.2 \pm 1.8\%$  in subject 1, whereas CSP and CSSP yielded  $3.1 \pm 2.8\%$ ,  $3.5 \pm 3.3\%$  respectively. This shows that the proposed scheme introduced in this paper is effective in improving the EEG-signal-classification accuracy. Furthermore, the figure also demonstrates that the classification results are subject dependant in motor imagery. For some subjects such as "sub.1," the classification error rate is generally lower, while for some other subjects such as "sub.4," the classification error rate is generally higher.

As a further analysis we applied the CSP, CSSP and the proposed PCSP method respectively on 1 second, 2 seconds, 3 seconds and 4 seconds length data segments.

In Figs. 7 - 14 are presented results of binary classification and triple classification for the four subjects separately. The figures illustrate that sub.1 gains the lowest error rate under 4 seconds imagination and PCSP method

Table 1. CLASSIFICATION ACCURACIES OF 10-FOLDS CROSS-VALIDATION PERFORMED USING CSP, CSSP AND PCSP

Subjects	CSP		CSSP		PCSP	
	Error rate (%)	Std (%)	Error rate (%)	Std (%)	Error rate (%)	Std (%)
sub.1	3.1	2.8	3.5	3.3	2.2	1.8
sub.2	8.5	5.4	14.6	6.2	4.6	2.2
sub.3	5.3	3.8	6.0	3.9	4.3	2.1
sub.4	29.1	8.2	32.6	7.6	16.4	3.5

has outperformed all the other methods on average. Moreover, we can observe that 4 subjects all obtain better performance under 4 seconds imagination than others which may give a hint of the performance continues to increase due to the subject learning and adaptation after some repeated experiments. We also compare the accuracy of binary classification and triple classification on same subject and the results show that the former has better performance in general.

### 5 CONCLUSION

In this paper, we have developed a novel Partitioned Common Spatial Pattern algorithm for single trial EEG classification. Unlike the conventional CSP method, The proposed algorithm finds the optimal spatial filters for each of the frequency components partitioned by a bank of fifth-order Butterworth filters, extracts features and selects relevant features by means of a class separability measure based on the Bhattacharyya distance. The spatial frequency confidence maps also were utilized to find the most relevant spatial frequency filters. The advantages of the proposed method were proved by its application to the

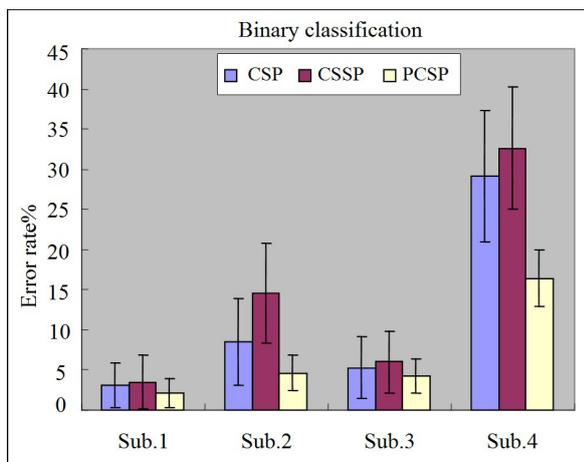


Fig. 6. Error rate and standard deviation of 10-folds cross validation

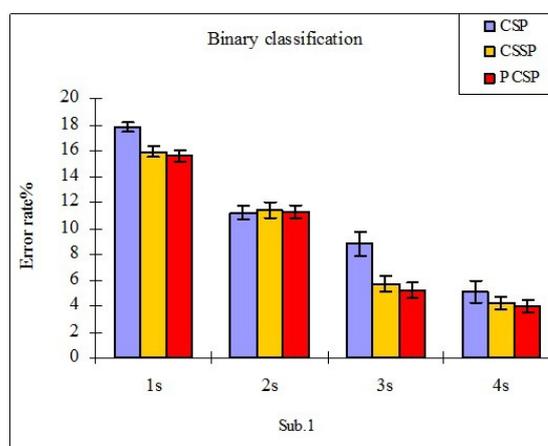


Fig. 7. Error rate and standard deviation of 10-folds cross validation for Sub.1's binary classification

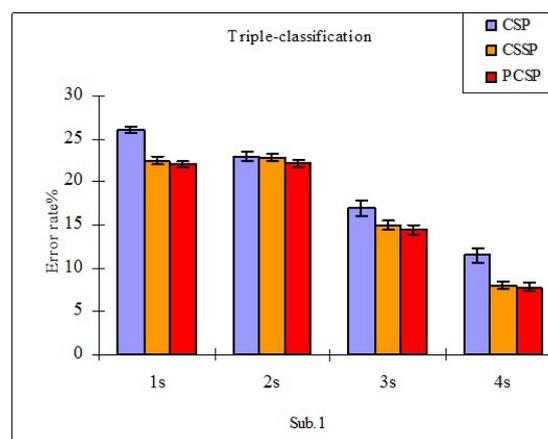


Fig. 8. Error rate and standard deviation of 10-folds cross validation for Sub.1's triple classification

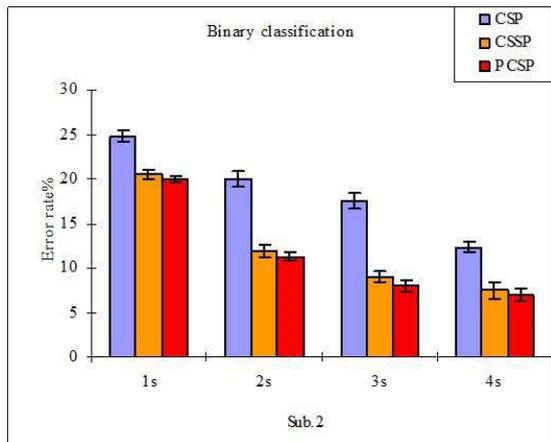


Fig. 9. Error rate and standard deviation of 10-folds cross validation for Sub.2's binary classification

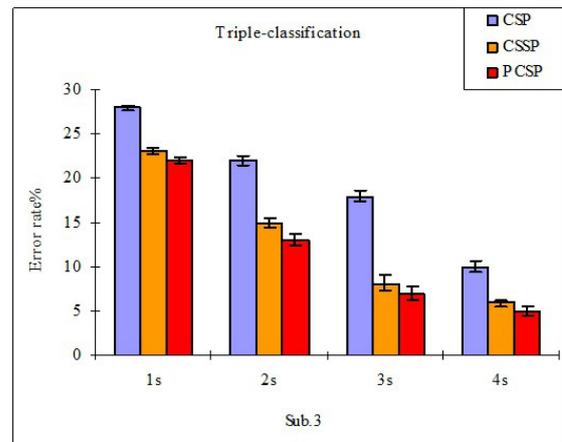


Fig. 12. Error rate and standard deviation of 10-folds cross validation for Sub.3's triple classification

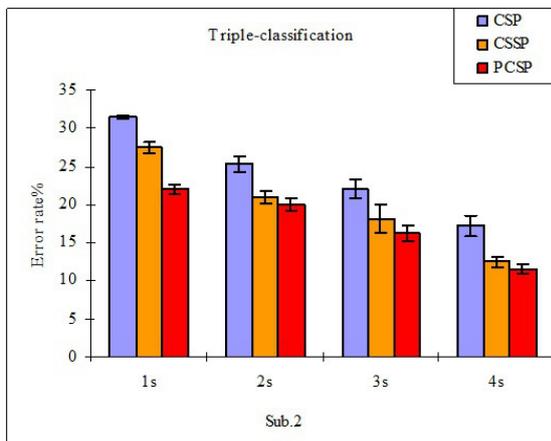


Fig. 10. Error rate and standard deviation of 10-folds cross validation for Sub.2's triple classification

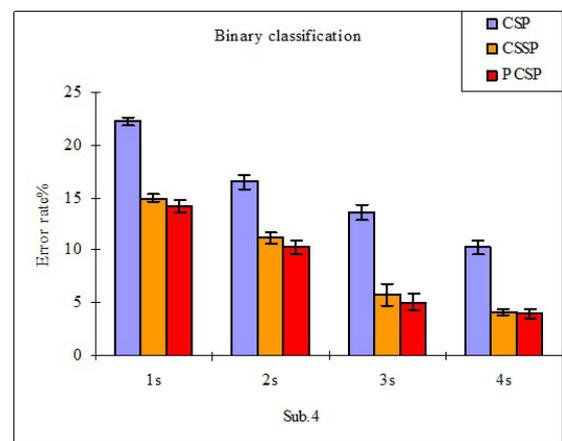


Fig. 13. Error rate and standard deviation of 10-folds cross validation for Sub.4's binary classification

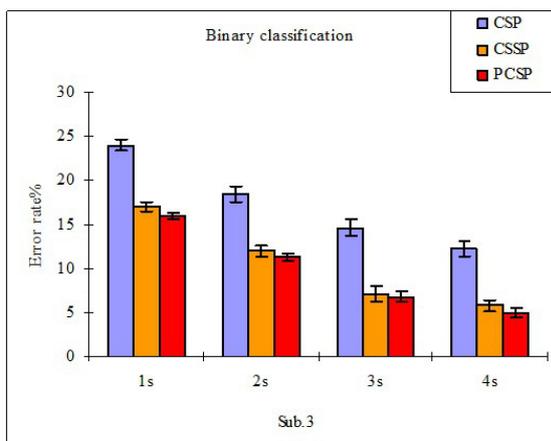


Fig. 11. Error rate and standard deviation of 10-folds cross validation for Sub.3's binary classification

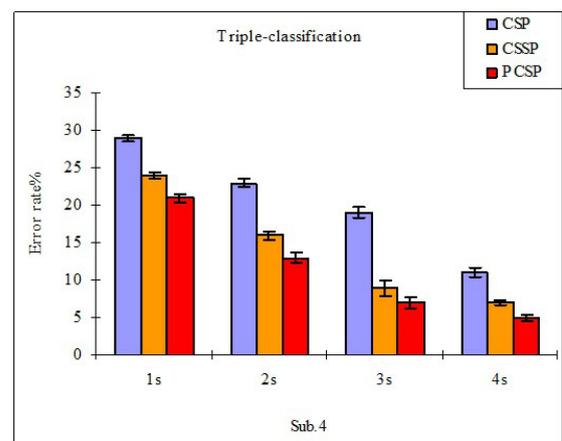


Fig. 14. Error rate and standard deviation of 10-folds cross validation for Sub.4's triple classification

classification of motor imagery. The experimental results on 4 subjects demonstrated that the PCSP algorithm introduced here outperforms the current state-of-the-art CSP and CSSP algorithm in terms of classification accuracy as well as robustness and will be a promising data exploratory tool for developing BCI system.

In future work we will focus on how to improve the efficiency of the proposed PCSP method in low-resolution EEG input and small-dataset conditions.

## ACKNOWLEDGMENT

The authors thank the editors and anonymous reviewers who have given many valuable comments. This work was supported in part by the Grants from the Research Award Fund for Outstanding Middle-aged and Young Scientists of Shandong Province (Grant No. BS2013DX016), Scientific Research Foundation for the Introduction of Talents of SDUST (Grant No. 2013RCJJ040), the Scholarship of Shandong Provincial Education Association for International Exchanges, the Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20133718110013) and The science and technology project of Huangdao district (Grant No. 2014-1-27).

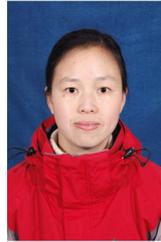
## REFERENCES

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.
- [2] G. R. Muller-Putzand and G. Pfurtscheller, "Control of an electrical prosthesis with an ssvp-based bci," *IEEE Trans. Biomed. Eng.*, vol. 17, no. 3, pp. 29–48, 2000.
- [3] M. Arvaneh, C. Guan, K. K. Ang, and C. Quek, "Optimizing the channel selection and classification accuracy in eeg based bci," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 6, pp. 1865–1873, 2011.
- [4] H. Higashi and T. Tanaka, "Simultaneous design of fir filter banks and spatial patterns for eeg signal classification," *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 4, pp. 1100–1110, 2013.
- [5] W. Samek, M. Kawanabe, and K. R. Muller, "Divergence-based framework for common spatial patterns algorithms," *IEEE Reviews in Biomedical Engineering*, vol. 7, pp. 50–72, 2014.
- [6] D. J. McFarland, C. W. Anderson, K. R. Muller, A. Schlogl, and D. J. Krusienski, "Bci meeting 2005-workshop on bci signal processing: feature extraction and translation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 135–138, 2006.
- [7] G. Muller-Gerking, J. Pfurtscheller and H. Flyvbjerg, "Designing optimal spatial filters for single-trial eeg classification in a movement task," *Clin. Neurophys.*, vol. 110, no. 5, pp. 787–798, 1999.
- [8] G. Pfurtscheller and F. L. d. Silva, "Event-related eeg/meg synchronization and desynchronization: Basic principles," *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, 1999.
- [9] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial eeg during imagined hand movement," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 4, pp. 441–446, 2000.
- [10] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve bci designs: Unified theory and new algorithms," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 2, pp. 355–362, 2010.
- [11] S. Lemm, B. Blankertz, G. Curio, and K. R. Muller, "Spatio-spectral filters for improving the classification of single trial eeg," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 9, pp. 1541–1548, 2005.
- [12] J. Farquhar, N. J. Hill, and B. Lal, T N andScholkopf, "Regularised csp for sensor selection in bci," in *Proc. of the Third International Brain-Computer Interface Workshop and Training Course*, (Graz, Austria), pp. 14–15, Sep. 2006.
- [13] H. Lu, K. Plataniotis, and A. Venetsanopoulos, "Regularized common spatial patterns with generic learning for eeg signal classification," in *Proc. IEEE Int. Conf. Eng. Med. Biol. Soc.*, (Minneapolis, MN), pp. 6599–6602, Sep. 2009.
- [14] H. Lu, H. L. Eng, C. Guan, K. Plataniotis, and A. Venetsanopoulos, "Regularized common spatial pattern with aggregation for eeg classification in small-sample setting," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 12, pp. 2936–2946, 2010.
- [15] M. Kawanabe, W. Samek, K. R. Muller, and C. Vidaurre, "Robust common spatial filters with a maxmin approach," *Neural Comput.*, vol. 26, no. 2, pp. 349–376, 2014.
- [16] H. Wang and D. Xu, "Comprehensive common spatial patterns with temporal structure information of eeg data: Minimizing nontask related eeg component," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2496–2505, 2012.
- [17] H. Wang, "Discriminant and adaptive extensions to local temporal common spatial patterns," *Pattern Recognit. Lett.*, vol. 34, no. 10, pp. 1125–1129, 2013.
- [18] G. Dornhege, B. Blankertz, G. Curio, and K.-R. Muller, "Increase information transfer rates in bci by csp extension to multi-class," in *Advances in Neural Information Processing Systems 16.*, (British Columbia, Canada), pp. 733–740, December. 2003.
- [19] K. Fukunaga, *Introduction to Statistical Pattern Recognition, 2nd Edition*. New York: Academic Press, 1990.



**Hongyu Sun** received his BS degree in Electronic and Information Engineering in 2003, MS degree in Signal and Information Processing from Shandong University of Science and Technology in 2008, and the PhD degree in Computer Science from Tongji University in 2012, respectively. Now, he is a lecturer in the College of electronic communication and physics of Shandong University of Science and Technology, a visiting lecturer in University of Colorado Colorado Springs, USA. His research interests are in

brain signal processing and brain-computer interface.



**Bisheng Chen** received her BS degree in Mechanical Manufacture and Automation from China University of Petroleum in 1999, MS degree in Systems Analysis and Integration from Huazhong University of Science and Technology in 2004, respectively. Now, she is a PHD candidate in Mechanical and Electronic Engineering and a lecturer in the College of Mechanical and Electronic Engineering of Shandong University of Science and Technology. Her research interests are in brain-computer interface

and robotics technology.



**Lijun Bi** received her BM degree in Medical imaging in 2002, MS degree in Detection technology and automation from Shandong University of Science and Technology in 2007. Now, she is a lecturer in the College of electronic communication and physics of Shandong University of Science and Technology. Her research interests are in Pattern recognition and intelligent system.

#### AUTHORS' ADDRESSES

**Hongyu Sun, Ph.D.**

**Lijun Bi, Ph.D.**

**Bisheng Chen, M.Sc.**

**Shandong University of Science and Technology,  
579 Qianwangang Road, 266590, Qingdao, China  
email: shyg@126.com, bilijun\_1@163.com,  
bishengch@163.com.**

Received: 2014-10-24

Accepted: 2015-12-09