

# PSO-Sub-ABLD-Based Parameter Optimization for Motor-Imagery BCI



Feiyu Yin, Yangyang Miao, Xingyu Wang, and Jing Jin

**Abstract** Common spatial pattern (CSP) is one of effective feature extraction algorithms, which is widely applied to motor imagery (MI)-based brain–computer interface (BCI). However, its performance is susceptible to artifacts and noise. Therefore, some researchers proposed Sub-Alpha-Beta Log-Det Divergences (Sub-ABLD) algorithm to improve the performance of BCI systems. The performance of Sub-ABLD algorithm depends on the values of hyperparameters  $\alpha$ ,  $\beta$  and  $\eta$ . In this study, a strategy named PSO-Sub-ABLD was proposed to select three hyperparameters with particle swarm optimization (PSO). Two public BCI competition datasets were used to validate the effectiveness of the proposed strategy. The results show that compared with CSP and Sub-ABLD with default hyperparameters, PSO-Sub-ABLD method gains better classification accuracy.

## 1 Introduction

Brain–computer interface (BCI) system is a new technology designed to create a pathway that connects the human brain and external devices without peripheral nerves and muscles (Xu et al., 2018). The BCI system provides a new means of communication for people with severe neuromuscular disorders by decoding task-related electroencephalogram (EEG) recordings and translating them into computer instructions for control and communication with external devices (Wolpaw et al., 2002; Jin et al., 2011).

So far steady state visually evoked potentials (da Cruz et al., 2015), P300 evoked potentials (Jin et al., 2015), slow cortical potentials (Mensch et al., 2004), and event-related desynchronization (ERD) (Pfurtscheller, 1977)/event-related synchronization (ERS) (Pfurtscheller, 1992) are neural response patterns commonly used in BCI systems. Motor imagery (MI)-based BCI systems, which are based on the ERD and

---

F. Yin · Y. Miao · X. Wang · J. Jin (✉)

Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai, China

© Springer Nature Singapore Pte Ltd. 2021

A. Lintas et al. (eds.), *Advances in Cognitive Neurodynamics (VII)*, Advances in Cognitive Neurodynamics,

[https://doi.org/10.1007/978-981-16-0317-4\\_23](https://doi.org/10.1007/978-981-16-0317-4_23)

219

ERS phenomena, are widely used as it is easier to operate than other systems based on external stimulus (Qiu et al., 2016).

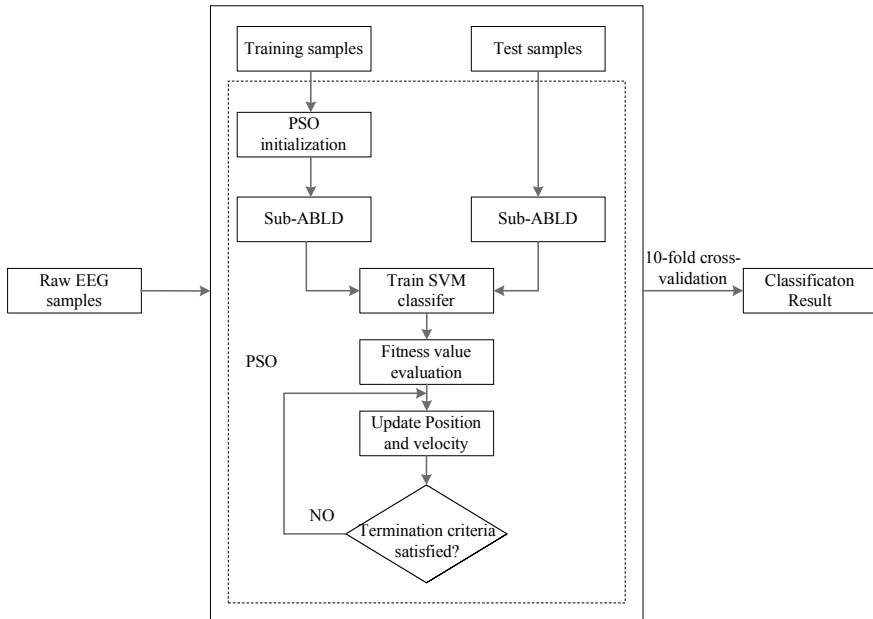
Common spatial pattern (CSP) has proven to be a very effective feature extraction method (Ramoser et al., 1998), and its principle is to find spatial filters, to effectively evaluate discriminant information of MI by maximizing the variances of the projected signal of one class while minimizing another (Zhang et al., 2015; Blankertz et al., 2008). Since the EEG signal is very sensitive to noise, outliers caused by noise will cause poor computation of spatial filters based on the spatial covariance matrix, which leads to poor classification accuracy (Lotte & Guan, 2011; Thiyam et al., 2017). To address this problem, a large number of improved algorithms based on CSP are proposed. Sub-ABLD algorithm is a modified algorithm of CSP, and its principle is to overcome the problem caused by the non-stationary nature of EEG data by appropriately scaling the conditional covariance matrix and using different filter selection strategies (Thiyam et al., 2017). Sub-ABLD algorithm shows a certain degree of robustness to outliers trials in EEG data (Feng et al., 2018). Three real hyperparameters  $\alpha$ ,  $\beta$ , and  $\eta$  affect the performance of Sub-ABLD (Thiyam et al., 2017), so how to choose better hyperparameters has a greater impact on improving the performance of Sub-ABLD algorithm. For the selection of hyperparameters, evolutionary algorithms (EA) such as genetic algorithm (GA) (Garrett et al., 2003) and particle swarm optimization (PSO) often have better effects. Compared with GA, PSO is widely used because of its advantages of simple programming, few parameters, and global search. Thus, this present study proposed the method of optimization hyperparameters of Sub-ABLD with PSO (PSO-Sub-ABLD) and compares it with CSP and Sub-ABLD with default hyperparameters. Two BCI competition datasets are selected to evaluate the performance.

The remainder content of this article is as follows: Sect. 2 describes the competition datasets used in this paper. Section 3 introduces the proposed method. Section 4 shows results. Finally, Sect. 5 concludes this study.

## 2 Description of the Data

In this paper, two competition datasets are used to evaluate the effectiveness of optimization parameters with PSO in the MI classification (Fig. 1).

- (1) Dataset1 (BCI competition IV datasets I): The dataset was recorded from 4 healthy subjects (named as a, b, f, and g) at 59 electrodes with sampling rate 100 Hz, during right hand, left hand, and foot MI tasks with a total of 200 trials that two classes of three tasks would be selected (Zhang et al., 2012). Each trial started with left, right, or bottom visual cues on the screen for a duration of 4 s and the subject completed the motor imagery tasks. More details about this dataset can be seen in the following website: <http://www.bbci.de/competition/IV/>.



**Fig. 1** Flow diagram of optimizing hyperparameters of Sub-ABLD with PSO

(2) Dataset2 (BCI competition III datasets IVa): The dataset was recorded from 5 healthy subjects (aa, al, av, aw, and ay) at 118 electrodes with down-sampling rate 100Hz, during right hand and foot MI tasks with a total of 280 trials. The visual cue for each trial lasts for 3.5 s, in which only the right hand and right foot cues were displayed in the competition (Novi et al., 2007). More information about this dataset can be found from the following website: <http://www.bci.de/competition/iii/>.

### 3 Method

#### 3.1 Data Processing

For two datasets, starting from 0.5 to 2.5 s of EEG segment was selected from each trial (Song & Epps, 2007). The EEG data were third-order band-pass filtered with Butterworth band-pass filter of 8–30 Hz in this study (Sun et al., 2010).

### 3.2 Sub-Alpha-Beta Log-Det Divergences

Sub-ABLD is an improved method of CSP algorithm. Its main purpose is to extract  $d$  expected spatial filters to reduce the impact of outliers contained in the EEG data on feature extraction. It is mainly divided into the following two steps. First, the discriminant subspace of the spatial filters is obtained by a robust method, and then, the EEG signal is filtered by the spatial filter obtained by discriminating the subspace. In the second step, the features are extracted with CSP algorithm. The input parameters of Sub-ABLD algorithm are the covariance matrices  $M_j, N_j$  of two types of samples, hyperparameters  $\alpha, \beta, \eta$  and the number of filters. The process of obtaining the final spatial filter matrix is as follows:

- (1) Compute the prior probability  $p(c_1), p(c_2)$  and the average covariance matrix  $M, N$  of each class.
- (2) Compute the average covariance matrix  $\text{Cov}(x)$  of the population (all types of stimuli) and perform eigenvalue decomposition on it.

$$\text{Cov}(x) = p(c_1)M + p(c_2)N \quad (1)$$

$$\text{Cov}(x) = U_1 \Delta U_1^T \quad (2)$$

- (3) Compute the whitening matrix  $T$ :

$$T = \Delta^{-\frac{1}{2}} U_1^T \quad (3)$$

- (4) The whitening conversion process is performed on the covariance matrix of two types of samples and the average covariance matrix of each class to obtain  $\hat{M}_j, \hat{N}_j, \hat{M}, \hat{N}$ .
- (5) Compute scaling parameters  $k$ :

$$k = \begin{cases} k_{\text{inf}} + \varepsilon & \text{for } k_{\text{inf}} \geq 1 \\ 1 & \text{for } 1 \in (k_{\text{inf}}, k_{\text{sup}}) \\ k_{\text{sup}} - \varepsilon & \text{for } k_{\text{sup}} \leq 1 \end{cases} \quad (4)$$

- (6) Initialize the iteration counter,  $i = 0$ ;
- (7) Initialize the semi-orthogonalization matrix  $\Omega^{(i)} = I_{n \times d}$ , where  $n$  refers to the size of the average covariance matrix of each class, and  $d$  is the number of filters;
- (8) Compute robust criterion:

$$\begin{aligned}
f(\Omega^{(i)}) &= \eta(p(c_1)) \frac{1}{N_1} \sum_{j=1}^{N_1} D_{AB}^{(\alpha, \beta)} ((\Omega^{(i)})^T \hat{N}_j(\Omega^{(i)}) \| k(\Omega^{(i)})^T \hat{N}(\Omega^{(i)}) \\
&\quad - \eta(p(c_1)) \frac{1}{N_1} \sum_{j=1}^{N_1} D_{AB}^{(\alpha, \beta)} ((\Omega^{(i)})^T \hat{M}_j(\Omega^{(i)}) \| k(\Omega^{(i)})^T \hat{M}(\Omega^{(i)})) \\
&\quad + D_{AB}^{(\alpha, \beta)} ((\Omega^{(i)})^T \hat{M}(\Omega^{(i)}) \| k(\Omega^{(i)})^T \hat{N}(\Omega^{(i)}))
\end{aligned} \tag{5}$$

- (9) Compute gradients  $\nabla f(\Omega^{(i)})$ , tangent matrices  $\Omega_{tg}^{i+1}$ , and projection matrices  $\Omega^{i+1}$ .

$$\nabla f(\Omega^{(i)}) = \frac{\partial f(\Omega^{(i)})}{\partial \Omega^{(i)}} - \Omega^{(i)} \frac{\partial f(\Omega^{(i)})}{\partial \Omega^{(i)}} \Omega^{(i)} \tag{6}$$

$$1. \Omega_{tg}^{i+1} = \Omega^i + \mu^{(i)} \nabla f(\Omega^{(i)}) \tag{7}$$

$$\Omega^{i+1} = N_L N_R^T \tag{8}$$

$$[Q_L, D, Q_R] = \text{svd}(\Omega_{tg}^{i+1}, 0) \tag{9}$$

- (10) Increase the iteration counter and determine if it converges, otherwise return to step 8.  
(11) Select the maximum/minimum eigenvector of  $((\Omega^{(i_{\max})})^T \hat{M} \Omega^{(i_{\max})}, (\Omega^{(i_{\max})})^T \hat{N} \Omega^{(i_{\max})})$ .  
(12) Obtain the final spatial filter matrix  $W^T$ :

$$W^T = V^T (\Omega^{(i_{\max})})^T T \tag{10}$$

After obtaining the final spatial filter matrix  $W^T$ , the raw EEG data obtained by the time window interception is projected through the spatial filter matrix, and two new types of EEG data can be obtained by constructing corresponding features.

The total number  $d$  of expected spatial filters is 6 in this study.

### 3.3 PSO

PSO is a group intelligent optimization algorithm proposed by Kennedy and Eberhart to imitate bird foraging, which has been successfully applied to various optimization problems (Poli et al., 2007). The basic principle is to randomly initialize a group containing  $N$  particles in a three-dimensional space, each of which is a feasible solution to the optimization parameters. Mark the  $i$ th particle as  $x_i = (x_{i1}, x_{i2}, x_{i3})$ , bring it into the evaluation function, and determine the individual optimal position  $p_i^t = (x_{i1}^t, x_{i2}^t, x_{i3}^t)$  and global optimal position  $p_g^t = (x_{g1}^t, x_{g2}^t, x_{g3}^t)$ . At the same time recording the velocity of the  $i$ th particle  $v_i = (v_{i1}, v_{i2}, v_{i3})$ , its position and velocity update formula is as follows:

$$v_{id}^{t+1} = wv_{id}^t + c_1r_1(p_{id}^t - x_{id}^t) + c_2r_2(p_{gd}^t - x_{id}^t) \tag{11}$$

$$x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \tag{12}$$

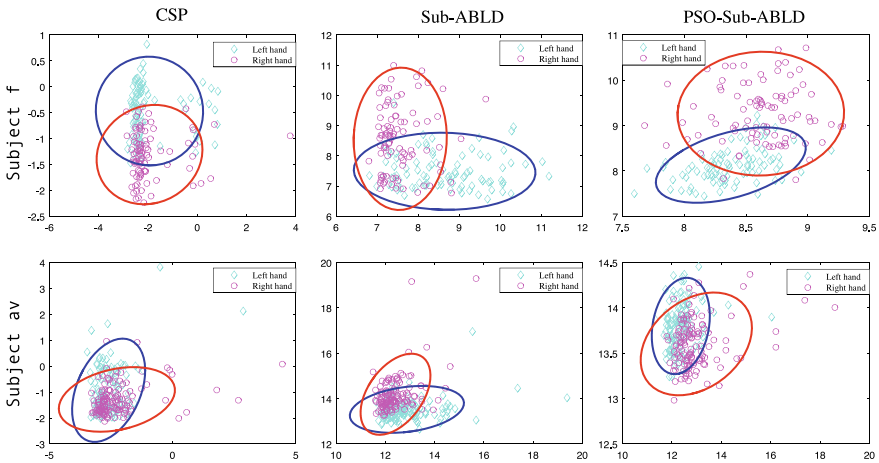
where  $w$  ( $w = 0.1$ ) is the inertia weight proposed by Shi and Eberhart; the size of  $w$  affects the ability of the particle global and local search;  $r_1$  and  $r_2$  are random constants in  $(0, 1)$ ; and  $c_1$  ( $c_1 = 1.2$ ) and  $c_2$  ( $c_2 = 1.2$ ) are learning factors. The velocity of particles range from  $-1$  to  $1$  and the range of position is  $(-2, 2)$ .

### 4 Results

Subject  $f$  of dataset 1 and subject  $av$  of dataset 2 are respectively selected, and feature distributions of CSP, Sub-ABLD, and PSO-Sub-ABLD are compared as shown in Fig. 2. From the figure, it can be observed clearly that features extracted by PSO-Sub-ABLD are easier to classify than two other methods for two subjects.

Tables 1 and 2 present classification accuracy derived by CSP, Sub-ABLD, and PSO-Sub-ABLD for all participants with test set of two datasets. For nine subjects selected in two datasets, PSO-Sub-ABLD improves the classification accuracy compared with two other algorithms and shows good generalization performance.

Table 1 summarizes that PSO-Sub-ABLD shows better classification accuracy for four subjects compared to two other methods. The average classification accuracy



**Fig. 2** Feature distribution (subject  $f$  of dataset 1 and subject  $av$  of dataset 2) of each class extracted by CSP, Sub-ABLD, and PSO-Sub-ABLD (cyan diamond represents the features of left hands, and magenta circle represents the features of right hands)

**Table 1** Classification accuracy comparison of CSP, Sub-ABLD ( $\alpha = \beta = 1.25, \eta = 0.25$ ), and PSO-Sub-ABLD ( $\alpha = 0.9, \beta = 1.7, \eta = 0.5$ ) using BCI competition IV dataset I

Subject	Dataset 1 (BCI computation IV dataset I)		
	CSP (%)	Sub-ABLD (%)	PSO-Sub-ABLD (%)
a	61.5	68	78
b	51	54.5	56.5
f	45	58	74.5
g	91	88	94.5
Mean	62.1	67.1	75.9
STd	17.7	15	13.6

**Table 2** Classification accuracy comparison of CSP, Sub-ABLD ( $\alpha = \beta = 2, \eta = 0.5$ ), and PSO-Sub-ABLD ( $\alpha = 0.6, \beta = 1.9, \eta = 1.4$ ) using BCI competition III dataset IVa

Subject	Dataset 2 (BCI competition III dataset IVa)		
	CSP (%)	Sub-ABLD (%)	PSO-Sub-ABLD (%)
aa	78.2	71.4	83.6
al	96.4	96.8	97.5
av	47.9	67.5	70.4
aw	86.0	82.9	90.4
ay	88.2	89.6	92.9
Mean	79.3	81.6	87.0
STd	16.8	12.2	10.5

obtained by PSO-Sub-ABLD ( $\alpha = 0.5, \beta = 0.9, \eta = 1.7$ ) is 75.9%, 13.8% higher than CSP and 8.8% higher than Sub-ABLD, respectively.

Table 2 shows classification accuracy for five subjects of dataset 2. The average classification accuracy obtained by PSO-Sub-ABLD ( $\alpha = 1.4, \beta = 0.6, \eta = 1.9$ ) is 87%, 7.7% higher than CSP and 5.4% higher than Sub-ABLD, respectively.

## 5 Conclusion

In this paper, PSO-Sub-ABLD has better robustness to outliers than CSP and Sub-ABLD to get better classification accuracy. Since default hyperparameters of Sub-ABLD result in poor generalization ability for different datasets, this study proposes PSO-Sub-ABLD algorithm, and it shows better generalization performance. In summary, compared with CSP and Sub-ABLD, PSO-Sub-ABLD obtains better classification performance based on the same classifier.

In the future work, one of the explorations is to combine channel selection and feature selection with PSO-Sub-ABLD to achieve better performance and to apply it to online systems.

## References

- Blankertz, B., Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., & Müller, K.-R. (2008). Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Processing Magazine*, 25, 581–607.
- da Cruz, J. N., Wan, F., Wong, C. M., & Cao, T. (2015). Adaptive time-window length based on online performance measurement in SSVEP-based BCIs. *Neurocomputing*, 149, 93–99. Advances in neural networks Advances in Extreme Learning Machines.
- Feng, J., Yin, E., Jin, J., Saab, R., Daly, I., Wang, X., et al. (2018). Towards correlation-based time window selection method for motor imagery BCIs. *Neural Networks*, 102, 87–95.
- Garrett, D., Peterson, D. A., Anderson, C. W., & Thaut, M. H. (2003). Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2), 141–144.
- Jin, J., Allison, B. Z., Sellers, E. W., Brunner, C., Horki, P., Wang, X., et al. (2011). An adaptive p300-based control system. *Journal of Neural Engineering*, 8(3), 036006.
- Jin, J., Sellers, E. W., Zhou, S., Zhang, Y., Wang, X., & Cichocki, A. (2015). A p300 brain–computer interface based on a modification of the mismatch negativity paradigm. *International Journal of Neural Systems*, 25(03), 1550011.
- Lotte, F., & Guan, C. (2011). Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms. *IEEE Transactions on Biomedical Engineering*, 58(2), 355–362.
- Mensh, B. D., Werfel, J., & Seung, H. S. (2004). BCI competition 2003-data set Ia: Combining gamma-band power with slow cortical potentials to improve single-trial classification of electroencephalographic signals. *IEEE Transactions on Biomedical Engineering*, 51(6), 1052–1056.
- Novi, Q., Guan, C., Dat, T. H., & Xue, P. (2007). Sub-band common spatial pattern (SBCSP) for brain–computer interface. In *2007 3rd International IEEE/EMBS Conference on Neural Engineering* (pp. 204–207).
- Pfurtscheller, G. (1977). Graphical display and statistical evaluation of event-related desynchronization (ERD). *Electroencephalography and Clinical Neurophysiology*, 43(5), 757–760.
- Pfurtscheller, G. (1992). Event-related synchronization (ERS): An electrophysiological correlate of cortical areas at rest. *Electroencephalography and Clinical Neurophysiology*, 83(1), 62–69.
- Poli, R., Kennedy, J., & Blackwell, T. (2007). Particle swarm optimization. *Swarm Intelligence*, 1(1), 33–57.
- Qiu, Z., Jin, J., Lam, H.-K., Zhang, Y., Wang, X., & Cichocki, A. (2016). Improved SFFS method for channel selection in motor imagery based BCI. *Neurocomputing*, 207, 519–527.
- Ramoser, H., Ramoser, H., Müller-Gerking, J., & Pfurtscheller, G. (1998). Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Transactions on Rehabilitation Engineering*, 8, 441–446.
- Song, L., & Epps, J. (2007). Classifying EEG for brain–computer interface: Learning optimal filters for dynamical system features. *Computational Intelligence and Neuroscience*, 57180.
- Sun, G., Hu, J., & Wu, G. (2010). A novel frequency band selection method for common spatial pattern in motor imagery based brain computer interface. In *The 2010 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–6). IEEE.
- Thiyam, D. B., Cruces, S., Olias, J., & Cichocki, A. (2017). Optimization of Alpha-Beta Log-Det divergences and their application in the spatial filtering of two class motor imagery movements. *Entropy*, 19(3), 89.



- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, *113*(6), 767–791.
- Xu, B., Zhang, L., Song, A., Wu, C., Li, W., Zhang, D., et al. (2018). Wavelet transform time-frequency image and convolutional network-based motor imagery EEG classification. *IEEE Access*, *7*, 6084–6093.
- Zhang, H., Guan, C., Ang, K. K., & Wang, C. (2012). BCI competition IV—Data set I: Learning discriminative patterns for self-paced EEG-based motor imagery detection. *Frontiers in Neuroscience*, *6*, 7.
- Zhang, Y., Zhou, G., Jin, J., Wang, X., & Cichocki, A. (2015). Optimizing spatial patterns with sparse filter bands for motor-imagery based brain–computer interface. *Journal of Neuroscience Methods*, *255*, 85–91.