Motor Imagery EEG Recognition Based on FBCSP and PCA

Banghua Yang[1], Jianzhen Tang[1(✉)], Cuntai Guan[2], and Bo Li[2]

[1] Department of Automation, School of Mechanical Engineering and Automation, Shanghai University, No. 149, Yanchang Road, Shanghai 200072, China
{yangbanghua,tjz3l52l}@shu.edu.cn
[2] School of Computer Science and Engineering, Nanyang Technological University, 50 Nanyang Avenue, Jurong West 639798, Singapore

Abstract. In motor imagery-based Brain Computer interfaces (BCIs), the classification accuracy of using the Common Spatial Pattern (CSP) algorithm to deal with the electroencephalogram (EEG) is closely related to the frequency range selected. Due to individual differences, the frequency range selected that reaches the best performance is different, which limits the generality and the actual use of the algorithm. To solve this problem, this paper proposes a motor imagery recognition method based on Filter Bank Common Spatial Pattern (FBCSP) and Principal Components Analysis (PCA), which is called FBCSP +PCA. The feasibility of the FBCSP+CSP is preliminary verified using the 2008 BCI competition data and further verified using data collected by our laboratory with wireless dry electrode device. The average classification accuracy of the data collected by our laboratory reaches 75.7% in the absence of individual band selection. That is also to say that the proposed method has good generality and and practical value because it can obtain high performance without the need of giving each individual a specific optimum frequency band.

Keywords: Motor imagery · FBCSP · PCA · Dry electrode

1 Introduction

Currently, a Brain-Computer Interfaces (BCI) system based on motor imagery is slowly approaching public view as a newly emerging rehabilitation method in the field of clinical rehabilitation. It achieves the human-computer interaction between the human brain and the computer by real-time measurement and analysis of the electroencephalogram (EEG) generated when motor imaging [1].

Effective and rapid feature extraction and classification methods are important in the practical application of BCI systems. The feature extraction method of Common Spatial Patterns (CSP) has good performance and is the most widely method used in the BCI system based on motor imagery [2]. However, the CSP method itself lacks frequency domain information, and the classification accuracy is closely related to the frequency range selection of EEG signals [3]. In actual use, because of the individual

differences of EEG signals, it is necessary to manually adjust a specific frequency range for each individual to obtain a higher correct rate, which limits its generality and practical application.

In order to solve the problem of manually selecting the subject-specific frequency range for the CSP algorithm, several methods have been proposed, including the Common Spatio-Spectral Pattern (CSSP) [4], the Common Sparse Spetral Spatial Pattern (CSSSP) [5], the Sub-band Common Spatial Pattern (SBCSP) and Filter Bank Common Spatial Pattern (FBCSP) [3, 6]. In the method of CSSP and CSSSP, due to the inherent nature of the optimization problem, the solution of filter coefficients is also strongly dependent on the choice of initial parameters. When using SBCSP, a comparative study of using different sub band score fusion techniques and classification algorithms are not available [6]. FBCSP algorithm comprises four stages: frequency filtering, spatial filtering, feature selection and classification. In the first stage, the EEG measurements are band-pass filtered into multiple frequency bands. In the second stage, CSP features are extracted from each of these bands. In the third stage, the Mutual Information (MI)-based feature selection algorithm is used to automatically select discriminative pairs of frequency bands and corresponding CSP features. In the fourth stage, a classification algorithm is used to classify the CSP features. And the FBCSP has been shown to yield superior classification accuracy compared against SBCSP on a publicly available dataset [3].

In this paper, a motor imagery recognition method based on FBCSP and Principal Components Analysis (PCA), which is called FBCSP+PCA, is proposed. The method uses a set of band-pass filters to decompose the EEG signal from 4–40 Hz into a plurality of frequency bands with a specific bandwidth, which enriches the frequency domain information. Meantime, intercepting a specific time period of EEG during imaging task through a time window eliminates the brain waves caused by the state conversion of thinking and visual evoked [7]. The CSP algorithm is used to extract the features of EEG from multiple frequency bands. The extracted multidimensional features are filtered by PCA for feature dimension reduction, which can solve the problem manually selecting a specific frequency caused by individual differences. Finally, the feature after dimension reduction are classified by Support Vector Machine (SVM) and K-NearestNeighbor (KNN). The method has achieved good experimental results, which lays the foundation for the practical application of the algorithm in BCIs.

2   Method

2.1   Introduction of CSP

The CSP technique has become a popular feature extraction approach in EEG-based BCI applications and it essentially finds spatial filters that maximize the variance for one class and simultaneously minimize the variance for the other class [8]. This method decomposes the raw EEG signals into spatial patterns, which are extracted from two classes of single trial EEG. Suppose the EEG signals of each trial are E_{D*N}, where D is the number of channels and N is the number of samples. The detailed computing process of the CSP can be seen in [9]. After the feature extraction by using CSP, a

feature vector $F = \{f_1, f_2, \cdots, f_{2m}\}$, where m is the number of feature pairs chosen, is obtained and used as an input for the feature dimension reduction.

2.2 Filter Bank Common Spatial Pattern

In this paper, the FBCSP algorithm comprises four stages: frequency filtering, spatial filtering, feature reduction and classification. In the first stage, the EEG signal is band-pass filtered into multiple frequency bands. In the second stage, CSP features are extracted from each of these bands. In the third stage, the PCA-based feature reduction algorithm is used to automatically get the most effective CSP features. In the fourth stage, the SVM algorithm and KNN algorithm are used to classify the CSP features. The specific description is as follows: (1) two groups of IIR (Infinite Impulse Response) band-pass filter are respectively used to divide EEG signals from 4–40 Hz with bandwidth 4 Hz and 2 Hz, which leads to a group of 9 bands signal and a group of 18 bands signal. According to the bandwidth selection in the reference [3] and that the smaller the bandwidth is, the more sub-bands are decomposed, the more frequency information is provided in principle. So the bandwidth of 4 Hz and 2 Hz are selected to decompose EEG. (2) The each band signal of two groups is respectively as input of the CSP filter to extract 2-dimensional features. The group of signals with a bandwidth of 4 Hz obtains a feature vector $F_1 = \{f_1, f_2, \cdots, f_{18}\}$ and the other group of signals with a bandwidth of 2 Hz obtains a feature vector $F_2 = \{f'_1, f'_2, \cdots, f'_{36}\}$. (3) The PCA algorithm is used for the feature reduction from two groups of CSP features. (4) The two groups of features after dimensionality reduction are used as input of SVM and KNN algorithm to obtain classification results. The processing of EEG signal is shown in Fig. 1.
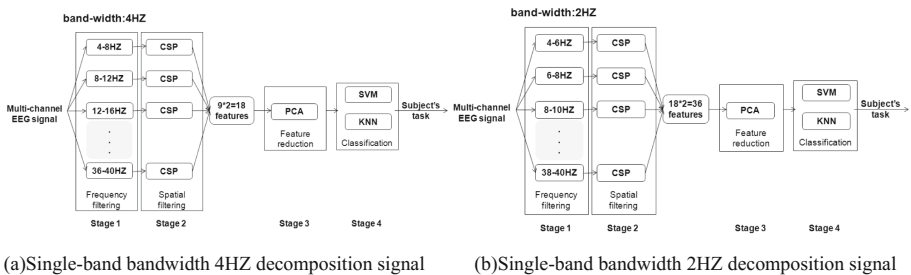

(a)Single-band bandwidth 4HZ decomposition signal   (b)Single-band bandwidth 2HZ decomposition signal

Fig. 1. The processing of EEG signal

2.3 PCA Algorithm

PCA is a multivariate statistical method that examines the correlations among multiple variables, which derives a small number of principal components that retain the information of the original variables as much as possible from the original variables. Using PCA to extract features of multidimensional features, we can extract the most

relevant signal features of EEG from motor imagery, and remove the noise signals and irrelevant components to improve classification accuracy and speed.

Suppose there is an extracted multidimensional feature matrix f_{n*M} where we can consider n is the total task number of left-handed imaging tasks and right-handed imaging tasks, and M is the dimension of the extracted features. The detailed steps of PCA can be described as follows:

- Step 1: Center features

$$A_{n*M} = f_{n*M} - \bar{f} \tag{1}$$

where $\bar{f}$ is the matrix formed by the average of each dimension in f_{n*M}.
- Step 2: Calculate the covariance matrix of the centralized matrix

$$B_{M*M} = \frac{A^T A}{n} \tag{2}$$

- Step 3: Calculate the eigenvector matrix U and the eigenvalue matrix $L(l_1, l_2, \cdots l_M)$ of the covariance matrix B, where $l_1, l_2 \cdots l_M$ is the eigenvalue of the B matrix and is expressed as descending order.
- Step 4: Select the eigenvectors corresponding to the first k largest eigenvalues according to the cumulative contribution rate G to form a transformation matrix T_{M*k}.

$$G = \frac{\sum_{i=1}^{k} l_i}{\sum_{i=1}^{M} l_i} \tag{3}$$

- Step 5: Reduce dimension

$$newf_{n*k} = f_{n*M} * T_{M*k} \tag{4}$$

By the feature dimensional reduction process of PCA, the few most effective features are automatically obtained, which avoids the need to manually adjust the optimal frequency range for different individuals and different time segments of the same individual. It improves the generality and practicability of the CSP algorithm.

3 Datasets

3.1 Dataset I

The first dataset used in this study is from "Data sets 1" of BCI Competition 4, which is launched on July 3rd 2008. The EEG data from four subjects (c, d, e, g) of the dataset are selected in this study, because the EEG data from these four subjects include two classes of motor imagery tasks, left hand and right hand. The main purpose of this

paper is to distinguish between motor imagery tasks. These EEG signals were recorded from 59 channels. The signals were band-pass filtered between 0.05 and 200 Hz and then sampled at 1000 Hz. More details were described in [10].

3.2   Dataset Ⅱ

The second dataset was from the authors' laboratory experiments. The experiment used g.Nautilus-8 wireless dry electrode of g.tec to collect EEG data. Sampling frequency is 250 Hz, cycle of each trial is 8 s. Each experimental cycle is divided into the following three links, as shown in Fig. 2:

- Link 1: 0–2 s black screen, subjects are in a relaxed state.
- Link 2: Cross appears on 2–4 s screen "+", subject is ready.
- Link 3: 4–8 s left or right arrow appears on the screen, the subjects carry on left or right hand motor imagery task according to the arrow instructions.

In the experiment, EEG data of eight subjects were collected. The EEG data of each subject contains four groups of left and right hand motor imagery EEG data. Each group of experimental data contains 40 times left-hand motor imagery tasks and 40 times right-hand motor imagery tasks. The electrodes distribution of EEG signals is shown in Fig. 3.
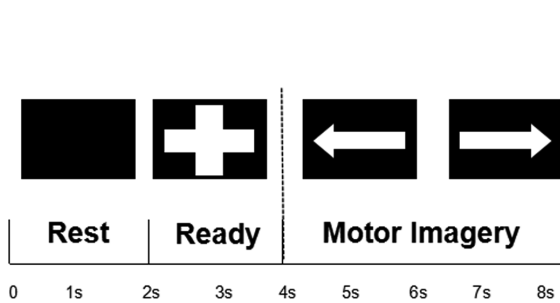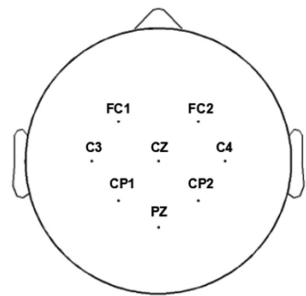

Fig. 2. Experimental flow chart

Fig. 3. The electrodes distribution

4   Results

4.1   Experimental Results Based on the First Dataset from BCI Competition 4

This paper first verifies the EEG data of four subjects (c, d, e, g) from Data sets 1 provided by the 4th BCI competition in 2008. In the competition data, each participant's data contains 100 left-handed imaging tasks and 100 right-handed imaging tasks, with a sampling frequency of 1000 Hz. Fifty left-handed and fifty-right-handed data were extracted from each set of data for training and the rest were used for testing. The

feature extraction method is FBCSP, then PCA is used for feature dimension reduction, the SVM and KNN methods are used to classify the left hand and right hand motor imagery. The detailed theory and Implementation of SVM and KNN can be seen in [11–13]. When the two feature vectors F_1 and F_2 are reduced by the PCA, the final feature dimension k needs to be determined in order to obtain the best classification result. In the experiment, the classification accuracy is closely related to the feature dimension k. And through the prior verification of the experiment, it is learned that when k is greater than or equal to 5, the classification effect is less than that k is small. So $k = 2, 3, 4$ is selected as the feature dimension for the final selection. The classification accuracy of the data set with bandwidths of 2 Hz and 4 Hz at different values of k is shown in Tables 1 and 2 respectively.

Table 1. The classification accuracy of different subjects at 2 Hz bandwidth from the first dataset

K	2		3		4	
Subjects	*SVM*	*KNN*	*SVM*	*KNN*	*SVM*	*KNN*
c	0.72	0.73	0.78	0.75	0.76	0.75
d	0.85	0.83	0.90	0.87	0.90	0.85
e	0.96	0.93	0.97	0.97	0.96	0.95
g	0.87	0.84	0.90	0.89	0.89	0.85
mean	0.85	0.832	0.887	0.87	0.877	0.85

Table 2. The classification accuracy of different subjects at 4 Hz bandwidth from the first dataset

K	2		3		4	
Subjects	*SVM*	*KNN*	*SVM*	*KNN*	*SVM*	*KNN*
c	0.67	0.62	0.70	0.67	0.68	0.63
d	0.83	0.80	0.85	0.84	0.86	0.80
e	0.96	0.94	0.97	0.95	0.93	0.92
g	0.88	0.84	0.92	0.92	0.91	0.90
mean	0.835	0.80	0.86	0.845	0.845	0.812

The Tables 1 and 2 shows that: (1) When the value of k is 3, the feature dimension after dimension reduction is 3 dimensions, the highest average classification accuracy rate can be achieved in both groups of features with bandwidth of 2 Hz and 4 Hz. (2) The accuracy from two groups features with a bandwidth of 2 Hz are all higher on average than that with a bandwidth of 4 Hz. The results can also be interpreted as a data set with a bandwidth of 2 Hz decomposes more frequency bands than a set of signals with a bandwidth of 4 Hz, and the extracted frequency domain information is more abundant.

4.2 Experimental Results Based on the Second Dataset from Laboratory Data

Twenty left hand and twenty right hand motor imagery data from each set of 8 subjects collected in our laboratory were adopted for training and the rest for testing. After the same processing as the above competition data, the average classification accuracy with bandwidths of 2 Hz and 4 Hz under different k values are shown in Tables 3 and 4 respectively.

Table 3. The classification accuracy of different subjects at 2 Hz bandwidth from the second dataset

K	2		3		4	
Subjects	SVM	KNN	SVM	KNN	SVM	KNN
A	0.718	0.743	0.75	0.787	0.675	0.731
B	0.731	0.762	0.756	0.781	0.681	0.718
C	0.706	0.718	0.743	0.756	0.637	0.675
D	0.681	0.681	0.731	0.718	0.675	0.669
E	0.718	0.706	0.737	0.743	0.687	0.693
F	0.743	0.75	0.781	0.775	0.706	0.718
G	0.699	0.706	0.75	0.743	0.675	0.687
H	0.713	0.687	0.756	0.756	0.693	0.687
mean	0.714	0.719	0.751	0.757	0.678	0.697

Table 4. The classification accuracy of different subjects at 4 Hz bandwidth from the second dataset

K	2		3		4	
Subjects	SVM	KNN	SVM	KNN	SVM	KNN
A	0.706	0.681	0.693	0.718	0.637	0.625
B	0.669	0.706	0.7	0.725	0.637	0.675
C	0.681	0.675	0.693	0.706	0.643	0.657
D	0.637	0.625	0.681	0.675	0.625	0.613
E	0.681	0.657	0.706	0.7	0.663	0.675
F	0.718	0.7	0.743	0.718	0.687	0.663
G	0.675	0.687	0.706	0.699	0.675	0.657
H	0.663	0.637	0.687	0.7	0.643	0.663
mean	0.678	0.671	0.701	0.705	0.651	0.655

The Tables 3 and 4 shows that: (1) When the value of k is 3, the highest average classification accuracy can be achieved in both groups of features with bandwidths of 2 Hz and 4 Hz. After feature dimension reduction, 3 dimensions is the best feature dimension for each subject. (2) A set of signals with a bandwidth of 2 Hz is 4.3% higher on average than that with a bandwidth of 4 Hz, which is more obvious than that obtained with the competition data. It also shows the method has a good performance with the data collected from the authors' laboratory experiments by g.Nautilus-8 wireless dry electrode EEG acquisition equipment.

Figure 4 shows the most significant Spatial Patterns using the frequency range that is autonomously selected in FBCSP+PCA. The results show that the left hemisphere and right hemisphere discriminates the right hand action and left hand action respectively for eight subjects from the authors' laboratory experiments. These results verify the neurophysiological plausibility of the CSP projection matrix computed for these subjects. However, the results for subjects 'D' and 'E' do not show such patterns. This is a plausible reason for the relative inferior test accuracy obtained for subjects 'D' and 'E'.
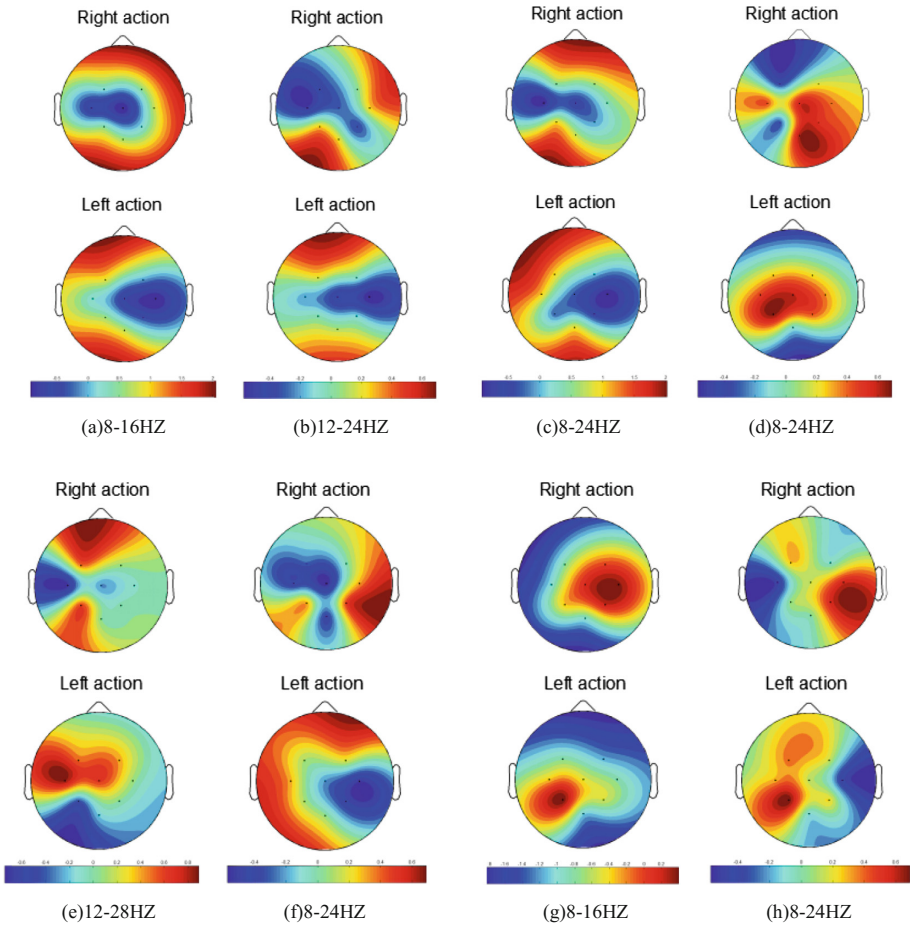

Fig. 4. Spatial patterns of using FBCSP on laboratory data for each subject

4.3 EEG Signal Processing Based on Time Window

The motor imagery time of each trial is four seconds. When subjects perform motor imagery tasks, there are the problems of visual evoked and delay of imaging especially in the beginning and end of 4 s, which would produce more volatility and reduce the signal-to-noise ratio, thus reducing the classification accuracy. To solve this problem, five time windows (0 to 4 s, 0.5 to 4 s, 0 to 3.5 s, 0.5 to 3.5 and 1 to 3 s) are chosen in the experiment.

Table 5 shows the average classification accuracy of different time windows for 8 subjects respectively under the bandwidth is 2 Hz and k takes 3. The result shows that when the time window is set to 0.5 to 4 s and 0 to 3.5 s, the average classification accuracy is higher than 0 to 4 s, and the classification accuracy of the time window (0.5 to 3.5 s) is the highest, which verifies it can eliminate the problems of visual evoked and delay of thinking and improves the classification accuracy by time window setting. When the time window is set to 1 to 3 s, the classification accuracy is down compared with 0.5 to 3.5 s. It shows that it can diminish the effect of EEG signal amplitude fluctuations, however, it also lose part of the useful information related to motor imagery under this time window.

Table 5. When bandwidth is 2 Hz and k takes 3, the average classification accuracy of different time windows for 8 subjects respectively

Time_window Subject	0-4s	0.5-4s	0-3.5s	0.5-3.5s	1-3s
A	0.675	0.718	0.706	0.75	0.725
B	0.669	0.693	0.7	0.756	0.731
C	0.657	0.693	0.706	0.743	0.706
D	0.643	0.681	0.693	0.731	0.7
E	0.65	0.675	0.681	0.737	0.718
F	0.687	0.706	0.731	0.781	0.743
G	0.669	0.681	0.699	0.75	0.718
H	0.657	0.687	0.699	0.756	0.725
mean	0.663	0.691	0.702	0.751	0.721

Figure 5 shows the average classification accuracies of the traditional CSP and the method proposed in this paper for 8 subjects from the authors' laboratory experiments respectively. The result shows the average accuracy of the method proposed in this paper is higher than the traditional CSP and reaches 75.7%. It verifies the method FBCSP+PCA can effectively eliminate the individual differences of EEG signals and gets better classification results.
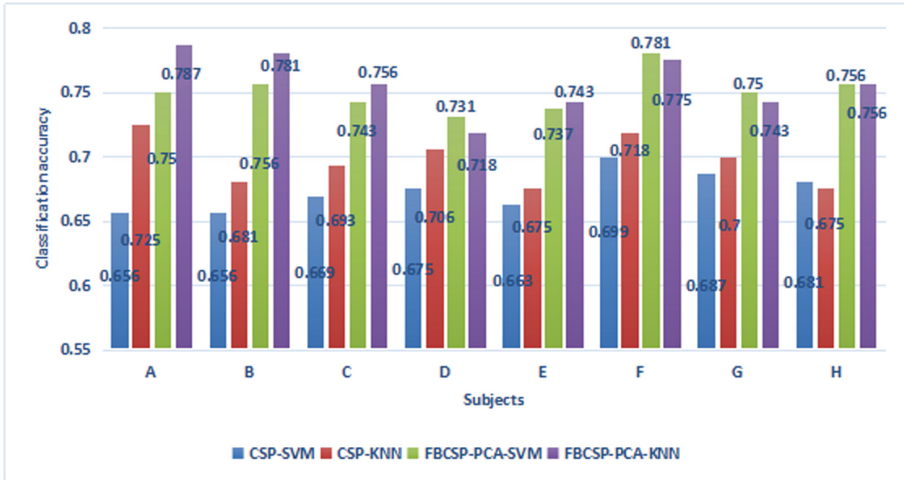
Fig. 5. The average classification accuracy of the traditional CSP and the method proposed in this paper for 8 subjects respectively

5   Discussion

Furthermore, the reason for using PCA as the feature reduction method is the PCA can reduce the correlation of the components in multidimensional features and so it increases the significance of each component. However, there is a drawback of the PCA method. The drawback is it can cause the presence of unexpected observations within the data to be processed by the PCA, which results in an incorrect de-correlation of the different features and so it may lacks part of information of motor imagery when feature reducing. There is a method called Robust PCA (RPCA) based on the Minimum Covariance Determinant (MCD) estimator, which has a good performance on this problem [14, 15]. Using this method to replace the conventional PCA method in FBCSP+PCA may has a better performance in getting the most effective features from different individuals and can improve the practical use of the method proposed in the paper.

6   Conclusion

In this paper, a motor imagery recognition algorithm called FBCSP+PCA is presented. The FBCSP algorithm can enrich the frequency domain information of feature extracted. The PCA algorithm can get the few most effective motor imagery features by feature reduction. The combination of these two algorithm effectively avoid the problem that need to manually adjust the certain frequency range selected caused by individual differences when using traditional CSP algorithm. By verifying two EEG data set from BCI Competition 4 and the authors' laboratory experiments, the method showed a good performance, which lays a foundation for in-depth study of practical application of portable BCI system based on motor imagery.

References

1. Yang, B., Li, H., Wang, Q., et al.: Subject-based feature extraction by using fisher WPD-CSP in brain-computer interfaces. Comput. Methods Prog. Biomed. **129**(C), 21–28 (2016)
2. Blankertz, B., Kawanabe, M., Hohlefeld, F.U., et al.: Invariant common spatial patterns: alleviating nonstationarities in brain-computer interfacing. In: International Conference on Neural Information Processing Systems, pp. 113–120. Curran Associates Inc. (2007)
3. Kai, K.A., Zheng, Y.C., Zhang, H., et al.: Filter bank common spatial pattern (FBCSP) in brain-computer interface. In: IEEE International Joint Conference on Neural Networks, pp. 2390–2397. IEEE (2008)
4. Lemm, S., Blankertz, B., Curio, G., et al.: Spatio-spectral filters for improving the classification of single trial EEG. IEEE Trans. Biomed. Eng. **52**(9), 1541–1548 (2005)
5. Dornhege, G., Blankertz, B., Krauledat, M., et al.: Combined optimization of spatial and temporal filters for improving brain-computer interfacing. IEEE Trans. Bio-Med. Eng. **53** (11), 2274 (2006)
6. Novi, Q., Guan, C., Dat, T.H., et al.: Sub-band common spatial pattern (SBCSP) for brain-computer interface. In: International IEEE/EMBS Conference on Neural Engineering, pp. 204–207. IEEE (2007)
7. Xu, Y., Haykin, S., Racine, R.J.: Multiple window time-frequency distribution and coherence of EEG using Slepian sequences and hermite functions. IEEE Trans. Biomed. Eng. **46**(7), 861–866 (1999)
8. Wang, Z., Logothetis, N.K., Liang, H.: Extraction of percept-related induced local field potential during spontaneously reversing perception. Neural Netw. **22**(5), 720–727 (2009)
9. Yang, B.H., Wu, T., Wang, Q., et al.: Motor imagery EEG recognition based on WPD-CSP and KF-SVM in brain computer interfaces. Appl. Mech. Mater. **556–562**, 2829–2833 (2014)
10. Blankertz, B., Dornhege, G., Krauledat, M., et al.: The non-invasive Berlin brain-computer interface: fast acquisition of effective performance in untrained subjects. Neuroimage **37**(2), 539 (2007)
11. Nishino, K., Nayar, S.K., Jebara, T.: Clustered blockwise PCA for representing visual data. IEEE Trans. Pattern Anal. Mach. Intell. **27**(10), 1675 (2005)
12. Joachims, T.: Making large-scale SVM learning practical. Technische Universität Dortmund, Sonderforschungsbereich 475: Komplexitätsreduktion in multivariaten Datenstrukturen, pp. 499–526 (1998)
13. Yu, C., Ooi, B.C., Tan, K.L., et al.: Indexing the distance: an efficient method to KNN processing. In: VLDB (2001)
14. Zabalza, J., Clemente, C., Caterina, G.D., et al.: Robust PCA micro-doppler classification using SVM on embedded systems. IEEE Trans. Aerosp. Electron. Syst. **50**(3), 2304–2310 (2014)
15. Zabalza, J., Ren, J., Yang, M., et al.: Novel Folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and SAR in remote sensing. ISPRS J. Photogramm. Remote Sens. **93**(7), 112–122 (2014)