Bio-Medical Materials and Engineering 26 (2015) S1523–S1532 S1523
DOI 10.3233/BME-151451
IOS Press

Binary particle swarm optimization for frequency band selection in motor imagery based brain-computer interfaces

Qingguo Wei*and Zhonghai Wei
Deparment of Electronic Engineering, Nanchang University, Nanchang 330031, China

Abstract. A brain-computer interface (BCI) enables people suffering from affective neurological diseases to communicate with the external world. Common spatial pattern (CSP) is an effective algorithm for feature extraction in motor imagery based BCI systems. However, many studies have proved that the performance of CSP depends heavily on the frequency band of EEG signals used for the construction of covariance matrices. The use of different frequency bands to extract signal features may lead to different classification performances, which are determined by the discriminative and complementary information they contain. In this study, the broad frequency band (8-30 Hz) is divided into 10 sub-bands of band width 4 Hz and overlapping 2 Hz. Binary particle swarm optimization (BPSO) is used to find the best sub-band set to improve the performance of CSP and subsequent classification. Experimental results demonstrate that the proposed method achieved an average improvement of 6.91% in cross-validation accuracy when compared to broad band CSP.

Keywords: Brain-computer interface, motor imagery, common spatial pattern, binary particle swarm optimization, frequency band selection

1. Introduction

Motor imagery (MI) is an important paradigm for the building of a brain-computer interface (BCI) [1]. When a subject conducts motor imagery, the power of EEG signals in specific areas of the brain decreases, a phenomenon called event related desynchronization (ERD). At the conclusion of motor imagery, the power of EEG signals in the same area of the brain increases, called event related synchronization (ERS) [2]. ERD and ERS are two physiological phenomena that are closely related to two sensory motor rhythms, i.e. mu and beta rhythms, which can be detected by relevant signal processing algorithms. It is widely believed that mu and beta rhythms are good signal features for MI based communication. Generally, mu and beta rhythms are located in the frequency ranges of 8-12 Hz and 18-25 Hz, respectively, but those frequency bands can vary across subjects and with varying mental states of the subjects [3, 4].

Common spatial pattern (CSP) has been reported as an effective and highly successful algorithm for the devising of spatial filters in ERD/ERS detection [5]. While CSP may perform well in extraction of

*Address for correspondence: Qingguo Wei, Department of Electronic Engineering, Nanchang University, 331 College Road, Nanchang 330031, China. Tel.: 86-791-83969680; Fax: 86-791-3969679; E-mail: wqg07@163.com.

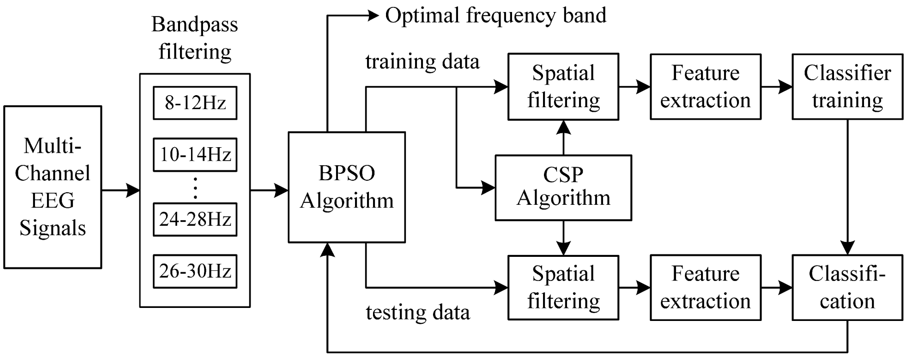

Fig. 1. Flow chart of frequency band selection based on BPSO.

spatial features, its performance is limited by the filtering of signals in the time domain. Classification of the CSP features generally yields poor accuracy when the EEG measurements are unfiltered or filtered with an inappropriate frequency range [6]. Hence, setting a broad frequency band (8-30 Hz) or manually selecting a subject-specific frequency band is common practice when using the CSP algorithm [7] because a broad band incorporates both mu and beta rhythms.

Although it is convenient to filter the EEG signals with a broad band (i.e. 8-30 Hz) filter, the EEG signals outside the mu and beta bands are useless for motor imagery recognition. Thus, the inclusion of extraneous signals degrades classification performance. Furthermore, the frequency ranges of mu and beta rhythms vary among subjects due to their individual distinctions in physiology, anatomy and brain state. Traditionally, an exhaustive search is performed with manual adjustments made for each subject [8], but the practice is highly time-consuming. Therefore, it is necessary to find a general purpose method that is capable of selecting the most responsive frequency bands for specific subjects in order to improve classification performance.

2. BPSO based CSP

To accurately locate the responsive frequency band of each subject to motor imagery tasks, the 8-30Hz broad band is divided into 10 sub-bands with widths of 4 Hz, and overlapping by 2 Hz. The frequency band selection method consists of four parts: sub-band filtering, sub-band selection with BPSO algorithm, spatial filtering and feature extraction in each sub-band based on CSP algorithm, and classification of the feature signals by a linear discriminant analysis (LDA) classifier [9]. In each sub-band, the classification accuracy of 5×10-fold cross-validation is used as the measure of fitness of the BPSO algorithm, which serves as the evaluation criterion of the frequency band selection. Figure 1 illustrates the flow chart of frequency band selection based on BPSO.

2.1. Temporal filtering

Before using the CSP algorithm for feature extraction, the band-pass filtering of the original EEG signals is integral; the performance of CSP depends largely on the frequency band responsive to the mental tasks of motor imagery. Conventionally, the original EEG signals are filtered in the broad frequency band 8-30 Hz because it encompasses both the mu and beta rhythms. However, this is a simple and rough method for data preprocessing, which does not suit every subject in order to achieve accurate information extraction. Considering that different people have different active frequency bands with different resolutions, we divided the broad band into 10 sub-bands of width 4 Hz and overlapping by 2 Hz (i.e. $8-12\text{Hz}, 10-14\text{Hz}, \cdots, 26-30\text{Hz}$). Infinite impulse response (IIR) filters of Chebeshev Type I is used as the band-pass filters for frequency band partition. By applying BPSO based sub-band selection to each subject, we can more precisely determine the subject-specific responsive frequency band. The overlapping frequency band division was employed in order to improve adaptability of the algorithm and prevent the occurrence of an optimal frequency range which crosses the border of two sub-bands.

2.2. Common spatial pattern (CSP)

The common spatial pattern (CSP) algorithm is effective in discriminating between two classes of EEG data by maximizing the variance of one class while minimizing the variance of the other class [5]. The multichannel EEG evoked by two mental tasks, A and B, can be denoted as spatio-temporal signal matrices X_A and X_B with dimensions $N(channels)\times T(samples)$. Then, the normalized spatial covariance matrices of the EEG signals for these two tasks can be estimated by

$$R_A = \frac{X_A X_A^T}{trace(X_A X_A^T)}, \quad R_B = \frac{X_B X_B^T}{trace(X_B X_B^T)} \tag{1}$$

where superscript T denotes the transpose operator and $trace(M)$ is the sum of diagonal elements of matrix M. The terms X_A and X_B are recorded for different mental tasks under the same conditions, and as such can be modeled by source components as follows:

$$X_\text{A} = [C_\text{A} \;\; C_\text{C}]\begin{bmatrix} S_\text{A} \\ S_\text{C} \end{bmatrix}, \quad X_\text{B} = [C_\text{B} \;\; C_\text{C}]\begin{bmatrix} S_\text{B} \\ S_\text{C} \end{bmatrix} \tag{2}$$

where S_A and S_B represent the source components specific to tasks A and B, respectively; C_A and C_B represent their corresponding spatial patterns, respectively; and S_C and C_C are the source component and its corresponding spatial pattern related to the common condition, respectively.

The purpose of the CSP algorithm is to design two spatial filters so that the source components S_A and S_B can be extracted by the following formula:

$$S_\text{A} = F_\text{A} X_\text{A}, \quad S_\text{B} = F_\text{B} X_\text{B} \tag{3}$$

where F_A and F_B are the respective spatial filters corresponding to tasks A and B. S_A and S_B contain important information for discrimination between the two tasks. CSP is based on the simultaneous diagonalization of the two spatial covariance matrices X_A and X_B. Principal component analysis (PCA) [10] and spatial subspace analysis are applied to these two diagonalized covariance matrices using training data to estimate the two spatial filters. The two spatial filters are optimal in that they extract task-related components and eliminate common components. Concrete calculation steps and a detailed description of the CSP algorithm can be found in previous research [5].

2.3. Binary particle swarm optimization (BPSO)

Particle swarm optimization (PSO) is a population based search algorithm simulated from the social behavior of birds within a flock [11]. PSO is an evolutionary optimization algorithm that has been applied in many scientific and engineering fields in recent years [12, 13]. PSO is initialized with a group of particles placed randomly on the search space, and seeks an optimal solution by updating particle generations. For every generation, each particle's velocity and position are updated according to its previous best position and the best position of all particles. This evolutional process can be described by Eq. (4).

$$v_i^{n+1} = w \cdot v_i^n + c_1 \cdot r_1 \cdot (p_i^n - x_i^n) + c_2 \cdot r_2 \cdot (p_g^n - x_i^n), \quad x_i^{n+1} = x_i^n + v_i^{n+1} \tag{4}$$

where x_i^n, v_i^n, and p_i^n represent the position, velocity, and previous best position of the ith particle in the nth generation; p_g^n denotes the best position of all particles in the nth generation; w presents the inertia weight; c_1 and c_2 represent cognitive and social components, respectively, and control the distance that a particle will travel in a single trial; r_1 and r_2 are random numbers uniformly distributed between 0 and 1.

Binary particle swarm optimization (BPSO) is a discrete binary version of PSO [14]. Velocity is updated in the same way as in PSO; the difference between PSO and BPSO is that in BPSO, each component of one particle adopts a binary value of "0" or "1", and the update rule for components is adjusted according to the Eq. (5).

$$s(v) = (1 + e^{-v})^{-1}, \quad \begin{cases} x_{id}^{n+1} = 1 & if \ \varphi < s(v_{id}^{n+1}) \\ x_{id}^{n+1} = 0 & if \ \varphi \geq s(v_{id}^{n+1}) \end{cases} \tag{5}$$

where x_{id}^{n+1} and v_{id}^{n+1} represent the dth component of x_i^{n+1} and v_i^{n+1}, respectively; φ denotes random numbers uniformly distributed between 0 and 1; and $s(v)$ is a sigmoid limiting transformation.

2.4. BPSO-CSP algorithm for frequency band selection

In this study, CSP and BPSO are used in combination to select the best frequency band, because of their high performances in feature extraction of EEG signals and evolutional searching, respectively. In the BPSO-CSP algorithm, each particle consists of 10 components, each corresponding to one sub-band of the broad frequency band. Thereby, each particle represents a combination of selected sub-bands and thus is a potential solution to the frequency band selection. Depending on the number of binary values of "1" at 10 components, one or several sub-bands may be selected by each particle. The covariance matrices of filtered EEG data are first calculated in each chosen sub-band, and summed. Then, the CSP algorithm is applied to the summed covariance matrix in order to extract spatial features. Finally, LDA classifier is used to classify the extracted features. The processing procedure of the BPSO-CSP algorithm for optimal frequency band selection is summarized as follows:

- **Initialization:** Each particle is initialized as a 10-dimensional vector. Each component of the vector corresponds to a frequency sub-band and is composed of binary number "0" or "1". A "1" represents a selected sub-band, while "0" represents a rejected sub-band.
- **Computing Fitness:** Features are extracted by CSP algorithm and classified by LDA based on the combination of the chosen frequency sub-bands. The classification accuracy of 10×5-fold cross-validation is defined as the fitness value of each particle.
- **Updating:** After the *n*th iteration, p_i^n will be updated if the fitness of x_i^n is the greatest value achieved (representing best fitness), and p_g^n will be updated if the highest fitness of all the particles is the greatest value achieved. The velocity v_i^n and position x_i^n of each particle will be adjusted according to Eqs. (4)-(6).
- **Mutating:** In order for the BPSO algorithm to get out of local optimal points, a mutation operator is required. The mutation rate will be reduced with increasing iteration number, defined as $p-mutation = 1-currentgen/t_{\max}$, where currentgen refers to the current iteration number and $t_{\max}$ denotes the maximal iteration number. If $m\_random < p\text{-}mutation$, the mutation is applied to each particle, which is selected randomly as follows:

$$Pop(i) = \theta \cdot \mu \cdot (1 - m\_random) \cdot v(i) + Pop(i) \tag{6}$$

  where $m\_random = random(0,1)$, and $\theta = \pm 1$ represents a particle's direction after mutation as consistent with or opposite to the original direction. The variable μ represents for the changing range of velocity; $\mu = 3$ in this study.
- **Results:** Repeat the above step Updating until the number of iterations exceeds a predefined number, $t_{\max}$. Finally, the particle's best position $p_g^{t_{\max}}$ is obtained, and thus the combination of optimal frequency sub-bands is determined.

3. Experiment and discussion

3.1. Data recording and pre-processing

The data set used in the study is the publicly available IVa of BCI Competition III [15]. It was recorded from five subjects (i.e. aa, al, av, aw, and ay) during either right foot or hand movement imagination. EEG signals were collected from 118 electrodes on the scalp. Each subject performed a total of 280 trials with the same number of trials for each motor imagery task. From a visual cue, test subjects were required to carry out the given motor imagery task for 3.5 seconds. Prior to temporal filtering, the continuous experimental data were intercepted into single-trial data, and common average reference (CAR) was adopted to re-reference them.

In the motor imagery based BCI experiment, as many as 118 electrode channels were used to record experimental data, resulting in a large computational load for the CSP algorithm. To address this

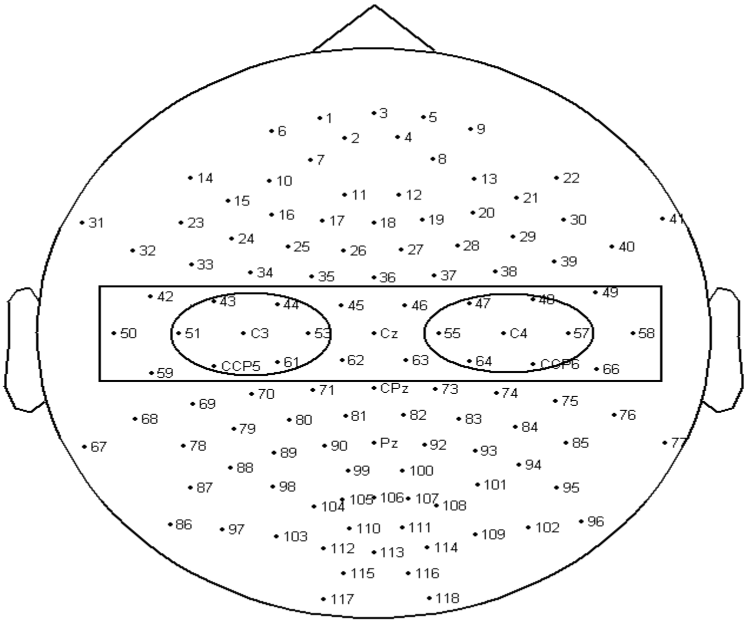

Fig. 2. Placement of the 118 EEG electrodes and selection of two channel subsets: one subset includes the 25 channels inside the rectangular box, whereas the second contains the 14 channels inside the two ellipses around electrodes C3 and C4.

disadvantage, we manually selected two channel subsets from the 118 channels, which are important for the neurophysiological discrimination between two mental tasks. As shown in Figure 2, the first channel subset includes the 25 channels in the rectangular box, while the second channel subset contains the 14 channels inside the two ellipses around electrodes C3 and C4.

3.2. Cross-validated classification

To assess the performance of frequency band optimization, two methods were used to extract classification features: a) original CSP was applied to the broad band 8-30 Hz; b) BPSO-CSP was applied to the 10 sub-bands. The feature vectors were classified by LDA, adopted because there is no regulation parameter to adjust. Three classification experiments were conducted to test the proposed algorithm. We respectively used the entire 118 channels, the selected 25 channels and 14 channels to discriminate between the two mental tasks. Both original CSP and BPSO-CSP were applied to the two channel subsets. Only CSP was applied to the entire 118 channels because the computation load of the BPSO-CSP algorithm is so large that the operation for optimal frequency band searching of a single subject was not concluded in several days.

The parameters of BPSO were set as follows: swarm size $pnum$=40; initial velocities are random numbers between -6 and 6; acceleration constants are c_1,c_2=$random$(1.5,2.0); inertia weight w=$random$(0.5,1.5); and iteration number $t_{\max}$=50.

The classification results of 10×5-fold cross-validation, achieved by the two methods with three different channel sets (25 channels, 14 channels and all 118 channels) are shown in Table 1. From the table, we observe that for all the subjects except "ay", the classification accuracies of the BPSO-CSP method are higher than those of CSP alone, for all channel sets. The classification accuracies of the sub-band BPSO-CSP with 25 channels are greater than the broad band CSP performed on all 118 channels for the first four subjects. The average classification accuracy of sub-band BPSO-CSP is

6.91% and 3.98% higher than those determined by broad band CSP using 25 and 14 channels, respectively. This is a significant improvement for a two-class motor imagery based BCI.

3.3. Frequency band selection

The frequency sub-bands selected by BPSO-CSP with respect to 25 and 14 channels for the five subjects are displayed in Table 2. When using 25 channels, the proposed algorithm selected only one sub-band for each of the five subjects, whereas when using 14 channels, the proposed algorithm selected two sub-bands for the first two subjects “aa” and “al”. The chosen sub-bands, however, are irregular for these five subjects, though the majority lies in the frequency range of 8-16 Hz with the exception of subject “av”, for whom the chosen sub-band is 20-24 Hz. This indicates that frequency band optimization in motor imagery based classification is closely related to the selection of channel subsets. However, the proposed BPSO-CSP algorithm could determine the optimal sub-band(s) for the classification task regardless of channel subsets.

3.4. Discussion

To verify the performance of the proposed algorithm in choosing optimal sub-band(s) for the classification task, a power spectral density analysis (PSDA) of the two-class motor imagery EEG data was conducted. Figure 3 depicts the average power spectrum of the EEG data set over trials derived from 25 channels for the five subjects. Based on the 25-channel EEG data, the BPSO-CSP algorithm selected only one sub-band for each subject, as shown in Table 2, because the two mental tasks are more easily discriminated in a single sub-band than in multiple sub-bands. It can be concluded from Figure 3 that the best frequency band, in which the power difference of the two classes of EEG signals is largest, is basically consistent with the sub-band selected by BPSO-CSP. This proves that the proposed BPSO-CSP is effective in selecting the optimal frequency band for an EEG data set with 25

Table 1

Classification accuracies of 10×5-fold cross-validation achieved by two methods with three different channel sets: 25 channels, 14 channels and all 118 channels

Subject	Sub-band BPSO-CSP		Wide band original CSP		
	25 channels	14 channels	25 channels	14 channels	118 channels
aa	89.57	88.64	75.64	81.46	79.57
al	97.57	95.04	95.29	91.82	95.39
av	77.32	74.11	68.21	69.11	76.43
aw	93.75	85.32	83.89	80.07	88.89
ay	94.25	92.39	94.86	93.14	95.39
Mean	90.49	87.10	83.58	83.12	87.13

Table 2

Frequency sub-band(s) selected by BPSO-CSP for five subjects with 25 channels and 14 channels

25 channels					14 channels				
aa	al	av	aw	ay	aa	al	av	aw	ay
12-16Hz	10-14Hz	20-24Hz	10-14Hz	8-12Hz	8-12Hz 10-14Hz	8-12Hz 10-14Hz	20-24Hz	10-14Hz	8-12Hz

channels.

A power spectrum analysis of the EEG data set derived from 14 channels was also conducted. There is no significant difference between the power spectra of EEG data with 25 and 14 channels, with the exception of subject “ay.” Based on the 14-channel EEG data, the proposed BPSO-CSP algorithm determined two sub-bands for the first two subjects and one sub-band for the other three subjects, as shown in Table 2. For subjects “aa” and “al”, a combination of sub-bands 8-12 Hz and 10-14 Hz could improve classification accuracy, although the power difference between the two-class EEG signals is small in sub-band 8-12 Hz. It is puzzling that the results for subject “av” demonstrate a significant power difference between the two mental tasks in the two sub-bands 10-14 Hz and 20-24 Hz, but the proposed algorithm determined only one sub-band. The reason may be that the two sub-bands provided similar information for classification due to their harmonic relation, and their combination did not contribute to recognition of the two mental tasks. Therefore, a conclusion can be drawn that the proposed BPSO-CSP algorithm is successful for the selection of optimal sub-bands in EEG data with14 channels.

It is well known that mu and beta rhythms are primarily found in the frequency range of 8-12 Hz and 18-26 Hz, respectively. However, this study reveals that the most responsive frequency bands in most subjects do not conform to these two rhythms, and the location of optimal frequency band varies from one subject to another. Thereby, it is necessary to develop an automatic method for optimal frequency band selection based on specific subjects, in order to improve classification performance in motor imagery based BCI systems. The classification results suggest that the proposed BPSO-CSP algorithm performed well in the selection of frequency bands.

4. Future work

In previous BCI research, the CSP algorithm was very successful in discriminating different classes of motor imagery. However, the correct selection of frequency bands is an important aspect affecting

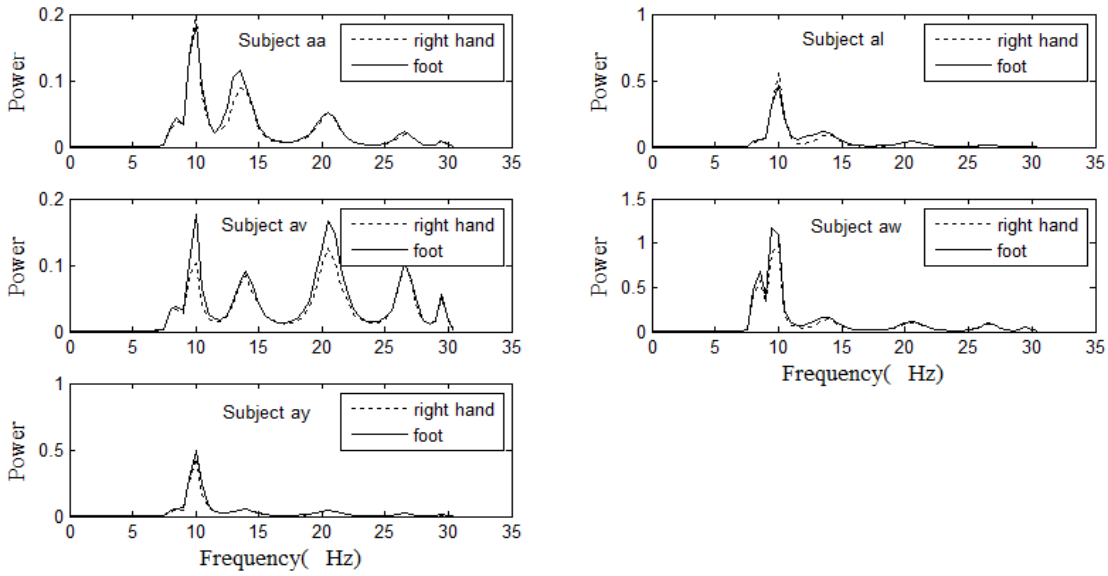

Fig. 3. Average power spectrum of the EEG data set with 25 channels for the five subjects.

the performance of CSP-based BCIs. In this paper, a BPSO-CSP frequency band selection method named is proposed which automatically and flexibly optimizes the frequency components before CSP is applied to the multi-channel EEG signals. Experimental results demonstrate that the proposed method achieved considerable improvement over CSP applied to a broad frequency band.

However, when using the BPSO method to select the optimal frequency band, the proposed method will requires a longer time to reach convergence. In searching for the optimal frequency band the particles will update by iterations, consuming a large amount of time. For this reason, we did not employ the BPSO-CSP method to select the optimal frequency bands for the data set with all 118 channels. This problem could be solved by improving the computer's hardware configuration or employing a better evolutionary algorithm which could search more quickly. In spite of this, the BPSO-CSP is a promising method; the efficiency and efficacy of the algorithm should be further verified by conducting more experiments on a large number of subjects, as well as comparing it with other evolutionary optimization algorithms.

This study also reveals that the selection of frequency bands depends largely on the selection of channels. In the study, we applied BPSO to two manually-chosen channel subsets based on physiological principle. This is not an accurate method for channel selection. Recent studies have shown that the informative features are located not only in a specific frequency band, but also in specific time segments and channel subsets. Future work analyzes the combined selection of channels, frequency bands and time segments using PSO optimization.

Acknowledgment

This project was funded by National Natural Science Foundation of China (# 61365013, # 60965004). The authors would like to thank the organizers of BCI Competition III for providing data.

References

[1] L.F. Nicolas-Alonso and J. Gomez-Gil, Brain computer interfaces: A review, Sensors **12** (2012), 1211-1279.
[2] G. Pfurtscheller and F.H. Lopes da Silva, Event-related EEG/MEG synchronization and desynchronization: Basic principles, Clinical Neurophysiology **110** (1999), 1842-1857.
[3] C. Andrew and G. Pfurtscheller, On the existence of different alpha band rhythms in the hand area of man, Neuroscience Letters **222** (1997), 103-106.
[4] G. Pfurtscheller, A. Stancak Jr and G. Edlinger, On the existence of different types of central beta rhythms below 30 Hz, Electroencephalography and Clinical Neurophysiology **102** (1997), 316-325.
[5] H. Ramoser, J. Muller-Gerking and G. Pfurtscheller, Optimal spatial filtering of single trial EEG during imagined hand movement, IEEE Transactions on Rehabilitation Engineering **8** (2000), 441-446.
[6] Q. Novi, C. Guan, T.H. Dat and P. Xue, Sub-band common spatial pattern (SBCSP) for brain-computer interface, 3rd International IEEE/EMBS Conference on Neural Engineering, CNE'07, 2007, pp. 204-207.
[7] S. Lemm, B. Blankertz, G. Curio and K. Muller, Spatio-spectral filters for improving the classification of single trial EEG, Biomedical Engineering **52** (2005), 1541-1548.
[8] K.K. Ang, Z.Y. Chin, H. Zhang and C. Guan, Filter bank common spatial pattern (FBCSP) in brain-computer interface, IEEE International Joint Conference on Neural Networks, 2008, pp. 2390-2397.
[9] D.F. Zhu, H. Wang, Y.H. Zhu, et al, Palmprint recognition combining LDA and the center band of Fourier magnitude, Emerging Intelligent Computing Technology and Applications, Springer Berlin Heidelberg, 2009, pp. 378-385.
[10] H. Abdi and L.J. Williams, Principal component analysis, Wiley Interdisciplinary Reviews: Computational Statistics **2** (2010), 433-459.
[11] J. Kennedy, Particle swarm optimization, Encyclopedia of Machine Learning, Springer US, 2010, 760-766.
[12] H. Gao, S. Kwong, J. Yang and J. Cao, Particle swarm optimization based on intermediate disturbance strategy algorithm

and its application in multi-threshold image segmentation, Information Sciences **250** (2013), 82-112.
[13] S. Liu, X. Zhao and C. Pang, Research on fast integer ambiguity resolution using pheromone-based particle swarm optimization algorithm, Journal of Information & Computational Science **12** (2015), 1081-1088.
[14] X. Yuan, H. Nie, A. Su, et al, An improved binary particle swarm optimization for unit commitment problem, Expert Systems with Applications **36** (2009), 8049-8055.
[15] K.K. Ang, Z.Y. Chin, C. Wang, et al, Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b, Frontiers in Neuroscience **6** (2012), 1-9.