Journal of Neuroscience Methods 278 (2017) 13–24

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

# Discriminative spatial-frequency-temporal feature extraction and classification of motor imagery EEG: An sparse regression and Weighted Naïve Bayesian Classifier-based approach

Minmin Miao, Hong Zeng, Aimin Wang*, Changsen Zhao, Feixiang Liu

*School of Instrument Science and Engineering, Southeast University, Nanjing 210096,China*

## HIGHLIGHTS

- Discriminative spatial filter vectors are selected adaptively.
- Significant CSP features are selected on time-frequency domains adaptively.
- The scale coefficient of sparse vector is used as weight of feature.
- A Weighted Naïve Bayesian Classifier is proposed for classification.
- The new method is a potential approach for improving the performance of BCI.

## ARTICLE INFO

*Article history:*
Received 3 November 2016
Received in revised form
14 December 2016
Accepted 20 December 2016
Available online 21 December 2016

*Keywords:*
Brain-computer interface (BCI)
Electroencephalogram (EEG)
Discriminative spatial filter
Sparse regression
Weighted Naïve Bayesian Classifier

## ABSTRACT

*Background:* Common spatial pattern (CSP) is most widely used in motor imagery based brain-computer interface (BCI) systems. In conventional CSP algorithm, pairs of the eigenvectors corresponding to both extreme eigenvalues are selected to construct the optimal spatial filter. In addition, an appropriate selection of subject-specific time segments and frequency bands plays an important role in its successful application.
*New method:* This study proposes to optimize spatial-frequency-temporal patterns for discriminative feature extraction. Spatial optimization is implemented by channel selection and finding discriminative spatial filters adaptively on each time-frequency segment. A novel Discernibility of Feature Sets (DFS) criteria is designed for spatial filter optimization. Besides, discriminative features located in multiple time-frequency segments are selected automatically by the proposed sparse time-frequency segment common spatial pattern (STFSCSP) method which exploits sparse regression for significant features selection. Finally, a weight determined by the sparse coefficient is assigned for each selected CSP feature and we propose a Weighted Naïve Bayesian Classifier (WNBC) for classification.
*Results:* Experimental results on two public EEG datasets demonstrate that optimizing spatial-frequency-temporal patterns in a data-driven manner for discriminative feature extraction greatly improves the classification performance.
*Comparison with existing methods:* The proposed method gives significantly better classification accuracies in comparison with several competing methods in the literature.
*Conclusions:* The proposed approach is a promising candidate for future BCI systems.


## 1. Introduction

A brain-computer interface (BCI) system supplies a novel control and communication channel between the brain of a human person and an external device, such as a prosthetic device (Shin et al., 2012; Jin et al., 2014; Jin et al., 2015). Due to its inexpensiveness, simplicity and high temporal resolution, the non-invasive scalp-recorded Electroencephalography (EEG) signal is most widely used in BCI applications (Novi et al., 2007). In the researches of non-invasive EEG based BCIs, motor imagery (MI) related BCI has become a hot research theme (Zhou et al., 2012). MI tasks involve the imaginations of motor actions, e.g., the imaginations of hand movements, however physical implementation is not allowed (Alvarez-Meza et al., 2015). MI leads to measurable potential

* Corresponding author.
*E-mail addresses:* 15006187659@sina.cn, wangam@seu.edu.cn (A. Wang).

http://dx.doi.org/10.1016/j.jneumeth.2016.12.010
0165-0270/

changes in the EEG signals termed as event-related synchronization/desynchronization (ERS/ERD) patterns, which are short lasting amplitude enhancement and attenuation in the rhythmic components of EEG signal (Ince et al., 2009; Thomas et al., 2009). In recent years, ERD/ERS based BCIs have received a lot of attentions due to its potential application in motor function rehabilitation and its assisting for the motor function impaired patients (Zhang et al., 2013a).

One of the major challenges in the ERS/ERD based BCI systems is the correct and effective class-discriminative features extraction from the blurred EEG signals (Zhou et al., 2012). To address this technical issue, common spatial pattern (CSP) (Wang and Zheng, 2008; Higashi and Tanaka, 2013; Lee and Kim, 2013) method is widely used for motor imagery related BCIs. CSP aims to maximize the projected scatters between two classes of EEGs by finding directions (i.e., spatial filters) and the spatial filters are computed via the simultaneous diagonalization of two covariance matrices. However, there are two major drawbacks for the CSP-based approaches. The first drawback is that the estimations of the covariance matrices are severely affected by outliers or noises (Wang and Zheng, 2008; Lee and Kim, 2013). When designing an optimal spatial filter with CSP, to the best of our knowledge, a majority of the previous methods considered pairs of the eigenvectors corresponding to the $K$ smallest and $K$ largest eigenvalues of the estimated covariance matrices (Shin et al., 2012; Novi et al., 2007; Blankertz et al., 2008; Zhang et al., 2015a; Xu et al., 2014; Wu et al., 2012; Suk and Lee, 2013). To some extent, we can't guarantee that eigenvectors corresponding to both extreme eigenvalues, calculated from a covariance matrix which has been contaminated, can form an optimal spatial filter for class-discriminative features extraction. Since the temporal, frequency, and spatial nonstationarities of ERS/ERD result in high intrasubject and intersubject variability in MI related BCIs, the second deficiency of CSP is that its successful application greatly depends on the appropriate selection of subject-specific temporal segments and frequency bands. Although an entire filter band (i.e., 8–30 Hz) and a predefined time segment are usually adopted for CSP algorithm in MI classification, an increasing number of researches have suggested that the optimization of time interval or filter band can significantly improve classification accuracy. In the literature, common sparse spectral pattern (CSSP) (Lemm et al., 2005), filter bank CSP (FBCSP) (Ang et al., 2008), discriminative CSP (DCSP) (Thomas et al., 2009), optimal spatiospectral filter networks (OSSFN) (Zhang et al., 2011) have been proposed for choosing the optimal frequency band automatically. Besides, some approaches have been proposed to select the optimal time segments automatically in MI classification (Xu et al., 2014; Ang et al., 2012; Hsu, 2011). Moreover, works that engage comprehensive considerations over the spatial-frequency-temporal features have been reported in recent years. Zhou GX et al. used tensor/matrix factorizations to explore spatial, temporal, and spectral differences and similarities in multi-channel EEG signals (Zhou et al., 2016a; Zhou et al., 2016b). Multiset canonical correlation analysis (MCCA) (Zhang et al., 2014a) and common feature analysis (CFA) (Zhang et al., 2015b) methods were proposed to learn more effective reference signals in BCI applications.

The main contributions of our work are three-fold. To overcome the first drawback of conventional CSP algorithm, we assume that a set of spatial filters, irrespective of the eigenvalues, is more proper for class-discriminative features extraction. Meanwhile, unlike Lee KY et al.'s work (Lee and Kim, 2013) which treated features in a pair-wise manner, we consider features from the spatial filters, that are eigenvectors, independently. In this paper, we propose a novel method of finding optimal spatial filters on each decomposed time-frequency segment. It selects discriminative bases of each filter based on our proposed dissimilarity metric called Discernibility of Feature Sets (DFS). By constructing spatial filters with an individual set of bases for each segmented time-frequency range in a data-driven manner, the nonstationary features of EEG signals can be managed and optimal spatial filters for classification can be obtained. Currently, sparse learning has been widely applied to feature extraction, selection and classification in BCI systems (Zhang et al., 2014b; Barthelemy et al., 2013; Zhang et al., 2013b; Zhang et al., 2016a). To deal with the second deficiency of CSP, we propose a novel sparse time-frequency segment common spatial pattern (STFSCSP) method by exploiting sparse regression (Qu et al., 2015; Zhang et al., 2012; Kolar and Liu, 2015) for automatic CSP features extraction on time-frequency domains. In the proposed method, CSP features are extracted on multiple signals that are located in a set of time-frequency segments. Sparse regression is applied to select the significant CSP features in a supervised way. Thirdly, a weight determined by the sparse coefficient which is the measure of significance for discrimination is assigned for each selected CSP feature. We notice that Bayesian analysis is frequently adopted in BCI systems and usually achieves good performance (Zhang et al., 2016c; Hoffmann et al., 2008; Zhang et al., 2016b). Hence we propose a Weighted Naïve Bayesian Classifier (WNBC) for MI based EEG classification. We provide the experimental results on two publicly available data sets, and our approach has been compared with the baseline CSP, as well as the state-of-the-art CSP-based methods, which aim at optimizing spatiospectral patterns for discriminative feature extraction in BCI. We show that our approach achieves superior classification accuracy.

The remainder of this paper is structured as follows. In Section 2, we present the data description and a detailed explanation of our method. Section 3 describes experimental results and discussion. Finally, the conclusions are given in Section 4.

## 2. Materials and methods

### *2.1. EEG dataset description and preprocessing*

#### *2.1.1. Dataset I:BCI competition III dataset IVa*

In this data set, five healthy subjects ("aa", "al", "av", "aw", "ay") participated in the collection of EEG signals and EEG signals were recorded from 118 Ag/AgCl electrodes (Blankertz et al., 2006). Subjects were instructed to perform MI according to the visual cue which was shown for 3.5 s for each trial. In this task, three kinds of motor imageries were performed, that is, right foot, left hand and right hand. However, only the motor imageries of right foot and right hand were provided for the competition. The EEG signals were further band-pass filtered between 0.05 Hz and 200 Hz and sampled at 1000 Hz. There were 280 trials in total, 140 trials per task for each subject. The numbers of trials for training and test across subjects were as follows: subject(right-foot, right-hand)-"aa"(88,80), "al"(112,112), "av"(42,42), "aw"(26,30) and "ay"(10,18) were for training and the rest were for test, the ratio of trials correctly classified to the total number of test trails was used as a criteria for evaluation. Considering the usual frequency band for ERD/ERS, a fifth order Butterworth band-pass filter was used to remove noises over 40 Hz and slow baseline signal under 6 Hz. In order to reduce the computational cost of subsequent processing, this dataset was down sampled to 250 Hz.

#### *2.1.2. Dataset II:BCI competition IV dataset IIa*

In this data set, nine subjects participated in the collection of EEG signals and EEG signals were recorded from 22 electrodes (Tangermann et al., 2012). Subjects were instructed to perform four types of MI: left hand, right hand, both feet and tongue. Two sessions on different days were recorded for each subject. Each session was comprised of 6 runs and one run consisted of 48 trials (12 for each class). Therefore, one session consisted of 288 trials totally. The