

Improving Decoding of the Mental Activities in BCI Systems using Overlapping FBCSP

H Rastgoo and S Azadi*

Department of Biomedical Engineering, Semnan University, Semnan, Iran.

Received 5 October 2014; revised 9 September 2015; accepted 20 November 2015

Brain Computer Interface (BCI) is an external channel from brain to the human organs. Therefore, BCI facilitate the disabled people to move their limbs. Electroencephalogram (EEG) is one of BCI recording signals which has high temporal resolution with low spatial resolution. Some researchers proposed common spatial filters to overcome the spatial resolution deficiency. However, these filters are highly sensitive to the CSP frequency band and, may decline the BCI performance. Many researchers increased the BCI performance by using filter bank CSP. They utilized a bank of band-pass filters to separate CSP in each band. However selecting of the number and the range of each frequency band becomes troubles ome. In this research work we utilized FBCSP with flexible overlapping frequency bands to enhance the decoding mental activities performance. The simulation results indicate that our proposed technique provides better classification performance ($\kappa=0.62$ with respect to the best $\kappa=0.57$ for the winner).

Keywords: BCI, motor imagery, FBCSP, overlapped FBCSP.

Introduction

BCI is an external channel between brain and human organs or other external equipment. This communication channel is especially vital for physically disabled people who has some defects in connection between brain and their motor limbs. Diseases, such as diabetes, damage nerve tissue¹ and therefore disconnect brain from limbs. There are many invasive and non-invasive approaches to record brain signals. Among the non-invasive methods, the electroencephalogram (EEG²) becomes the best recording technique due to its excellent temporal resolution with inexpensive recording devices. EEG signals may become blurred due to the volume conduction through the scalp, skull and other layers of brain. Therefore, event-related potential (ERP) is very weak³ and enhancing EEG spatial resolution is required. Koles et al.⁴ proposed common spatial pattern (CSP) to address these issues. CSP^{5,6} is effective in constructing optimal filters that discriminate two classes of EEG measurements in motor imagery based BCI⁷. However, the performance of this spatial filters are dependent on its operational frequency band⁸. Ang et al.⁷ proposed filter band common spatial pattern (FBCSP) to

explode more frequency information. FBCSP uses a bank of band-pass filters and a separate CSP in each band. Determination of frequency bands and number of filters in FBCSP are critical parameters in a BCI system. Their $\kappa=0.57$ on dataset number IIa of BCI competition IV. Javier Asensio-Cubero et al.⁸ proposed temporal segmentation for extracting CSPs and multiple Linear Discriminant Analysis (LDA) to improve the performance of BCI. They used a time range of [2, 7] while cue was appeared at $t=2s$ and therefore the cue imaginations naturally appear sometime after $t=2$ sec. Their result is $\kappa=0.59$. Alexandre Barachant et al.⁹ used Riemannian geometry to propose TSLDA and MDRM methods and obtained $\kappa=0.567$, and 0.52 , respectively. Yang et al.¹⁰ proposed automatic selection of number of spatial filters culminated in $\kappa=0.53$. Chang Liu et al.¹¹ used CSP and three classifiers KNN, SVM_D, and SVM_V with the mean $\kappa=0.59$, 0.51 , and 0.35 , respectively. Llera et al.¹² used MPMLDA classifier and obtained $\kappa=0.56$. In this paper we proposed an FBCSP with overlapped frequency bands in a wide range of frequencies to improve the performance of BCI system and remove these two parameters (band pass filters and number of them). Another advantage of proposed method is that it is not dependent on the subjects. Results indicate that this method outperforms the winner of competition IV. We gain

*Author for correspondence
E-mail: sazadi@semnan.ac.ir

$\kappa = 0.62$ compared to 0.57 for the winner in motor imagery with four classes over all nine subjects.

Problem Statement

In this section, we explain the problem statement. First, we introduce the dataset as follows.

Dataset

The dataset Ila from BCI competition IV have been provided by the Department of Medical Informatics, Institute for Biomedical Engineering, University of Technology Graz. This dataset includes EEG signal recording of 22 electrode and EOG signals of three electrodes from nine healthy subjects. All Electrodes are mono-polar and were sampled 250 Hz. This dataset is cue based and without any feedback. Data was recorded of nine healthy subjects in two different days. Each session includes six runs which separated by a short break. There are 12 trials for each of four motor imagery left hand, right hand, both feet and tongue in each run.

Therefore 288 trials are recorded in each session. So each motor imagery in each session has 72 trials. Trials of one session is used for training and the other session is used for test. A filter [0.5, 100] Hz is applied to each channel. A notch filter 50 Hz is activated to suppress line noise. All subjects were sitting in a comfortable armchair in front of a computer screen. At the beginning of a trial ($t = 0.0$ sec), a fixation cross appeared on the black screen. In addition, a short acoustic warning tone was presented. After two seconds ($t = 2$ sec), a cue in the form of an arrow pointing either to the left, right, down or up (corresponding to one of the four classes left hand, right hand, foot, or tongue) appeared and stayed on the screen for 1.25 sec. This prompted the subjects to perform the desired motor imagery task. No feedback was provided. The subjects were instructed to carry out the motor imagery task until the fixation cross disappeared from the screen at $t = 6$ sec. A short break with a black screen followed. This paradigm is illustrate in figure 1a.

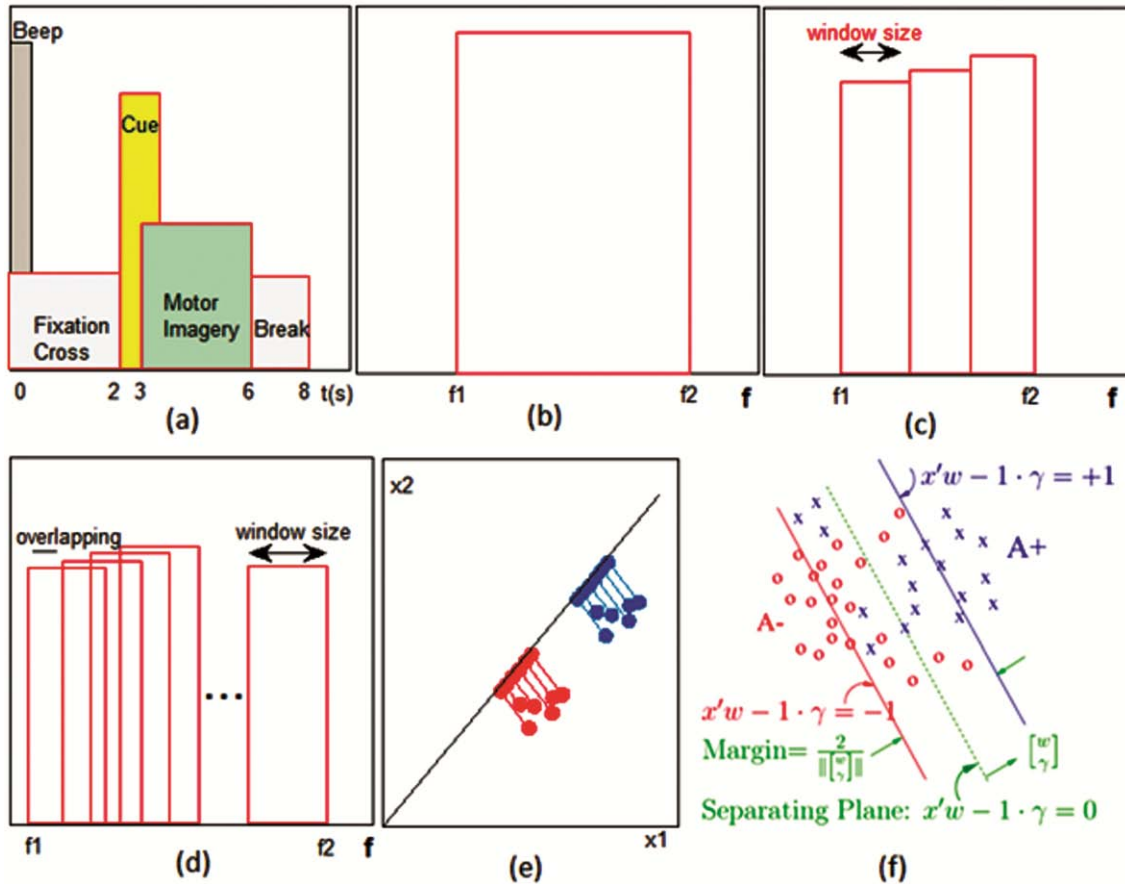


Fig. 1-(a)—Timing scheme of paradigm (Graz BCI group) Timing scheme of paradigm (Graz BCI group), b) Frequency segmentation without segmentation, c) Frequency segmentation without overlapping, d) Frequency segmentation with overlapping, e) projection of feature vectors on a line in LDA, f) the proximal support vector machine classifier: The proximal planes $x'w - \gamma = \pm 1$ around which points of the sets A^+ and A^- are clustered.

Proposed System

First a Butterworth band pass filter [8, 30] Hertz is applied to remove artifacts. Then FBCSP is used with overlapping frequency bands. Feature extraction is calculated of FBCSP output. Finally PCA reduces dimension of feature vectors to increase speed and performance.

FBCSP

The proposed Filter Bank Common Spatial Pattern (FBCSP) machine learning approach is illustrated in figure 2. It comprises two progressive stages of EEG measurements processing: multiple band pass filters using zero-phase Chebyshev Type II filters and spatial filtering using the CSP algorithm. The first stage employs a filter bank that band pass filters the EEG measurements into multiple bands. The second stage performs spatial filtering on each of these bands using the CSP algorithm. Thus, each pair of band pass and spatial filter yields CSP features that are specific to the frequency range of the band pass filter. There are various frequency bands in EEG such as the delta band (0.5-4 Hz), the theta band (4-8 Hz), the alpha

band (8-13 Hz), and the beta band (13-25 Hz)¹³. The FBCSP employs a filter bank that covers 4-40 Hz, which comprises nine band pass filters that covers 4 Hz each; and the CSP algorithm with two classes. Selecting band pass filters and number of them is critical in BCI performance.

Overlapped FBCSP

CSP uses frequency range without any segmentation while FBCSP uses segmentation of frequency range without overlapping in frequency bands. We proposed to use frequency range segmentation with overlapping in frequency bands. This causes smooth changing of features and optimal use of information in spectral domain. Figures 1b-1c shows these three schemes clearly.

Feature Extraction

Feature extraction is a crucial step in decoding mental activity. In this paper we use some common feature extraction algorithm like Hjorth¹⁴ parameters, Barlow parameters, band power estimation and time domain parameters. Hjorth et al. introduced three features Activity, Mobility, Complexity. Activity is the signal power (which is wide band filtered), Mobility is the mean frequency and Complexity the change in frequency. Band Power Estimator (BPE) calculates Activity at some specific frequency bands. Note that features in linear discriminant classifier (LDC) have to be Gaussian. After analyzing the "Gaussianity" of the features, it was decided to apply the logarithm to the Activity and leave Mobility and Complexity untransformed. Barlow parameters are quite similar Hjorth parameters and both are not Gaussian distributed. Amplitude (A), Frequency (F), and Spectral Purity Index (SPI, used to measure the irregularity of the signal. Average values (expected values $E(\cdot)$) are estimated over an exponential window and then logarithms are applied to A and SPI. Log (BP) is commonly used in many BCI systems. Signal in each channel is band-pass filtered at certain frequency bands (for example using Butterworth filters of order 5). The filtered signal was squared and the mean estimated over an exponential window. Finally, the logarithm was performed to statistically normalize the features to be classified with LDA. Suppose there is K channels in a single-trial EEG signals and B frequency bands, so you have K x B band powers in output of BPE. TDP is an extension of Hjorth and Barlow parameters so you can write the two latter in terms of TDP. TDP calculates variances of first K derivations of signal.

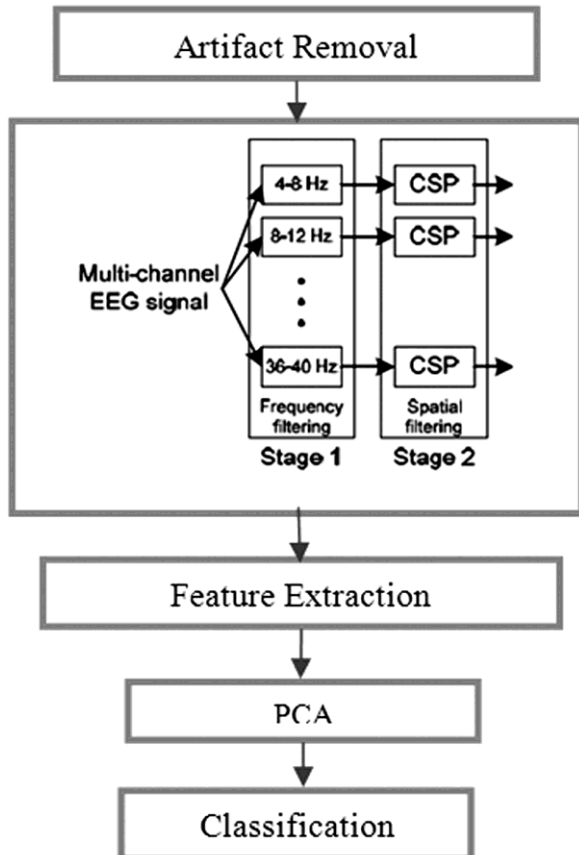


Fig. 2—A typical BCI system. FBCSP has two stages.

Classification

LDA and SVM are two common classifiers in BCI because they are linear and simple with satisfactory results. We use these classifiers and describe them briefly. Figure 1e shows the projection of feature vectors on a line in LDA. LDA seeks to reduce dimensionality while preserving as much of the class discriminatory information as possible. Assume we have a set of D -dimensional samples $\{x_1, x_2, \dots, x_N\}$, N_1 of which belong to class c_1 , and N_2 to class c_2 . We seek to obtain a scalar y by projecting the samples onto a line $y = w^T x$. Of all the possible lines we would like to select the one that maximizes the separation of the scalars y . In order to find a good projection vector, we need to define a measure of separation. Fisher suggested maximizing the difference between the means, normalized by a measure of the within-class scatter. Therefore, we are looking for a projection where examples from the same class are projected very close to each other and, at the same time, the projected means are as farther apart as possible. Standard support vector machines (SVMs) which are powerful tools for data classification, classify 2-category points by assigning them to one of two disjoint half spaces in either the original input space of the problem for linear classifiers, or in a higher dimensional feature space for nonlinear classifiers. Recently much simpler classifiers, the least squares and the proximal support vector machine (PSVM),

were implemented wherein each class of points is assigned to the closest of two parallel planes (in input or feature space) that are pushed apart as far as possible (figure 1f). This formulation, which can also be interpreted as regularized least squares or in the more general context of regularized networks, leads to an extremely fast and simple algorithm for generating a linear or nonlinear classifier that is obtained by solving a single system of linear equations.

Results and Discussions

We use 8 Hz length frequency bands starting from 6 to 36 Hz with 75% overlapping in time range [2.5, 5.5] sec. Table 1 shows the result of three case 1) without frequency range segmentation 2) with segmentation without overlapping, and 3) with segmentation with overlapping on dataset IIa of BCI competition IV. As shown in table 1, (indicated in figure 3a) the common features Hjorth, Barlow, BPE, and TDP all indicates approximately the same results. In this research work Hjorth features are used.

Figure 3a compares the results against the winner of competition. PSVM and LDA showed better results than the winner, and SVM provided very poor results. We vary the overlapping percent from 0.0 to 95 with an increment 5% to examine the BCI sensitivity. Figure 3b depicts the results obtained for this variation. As shown in figure 3b, κ increases with respect to the overlapping percentage. In addition, κ of PSVM is higher than that

Table 1—The mean kappa for four common features Hjorth, Barlow, BPE and TDP (order=7) and three classifier LDA, SVM and PSVM for all of nine subjects

		S ₁	S ₂	S ₃	S ₄	S ₅	S ₆	S ₇	S ₈	S ₉	mean
Hjorth	LDA	0.79	0.46	0.77	0.50	0.47	0.39	0.81	0.70	0.72	0.62
	SVM	0.58	0.31	0.52	0.32	0.29	0.29	0.59	0.49	0.44	0.42
	PSVM	0.79	0.47	0.78	0.51	0.45	0.37	0.83	0.73	0.72	0.63
Barlow	LDA	0.82	0.44	0.78	0.51	0.43	0.38	0.80	0.69	0.70	0.62

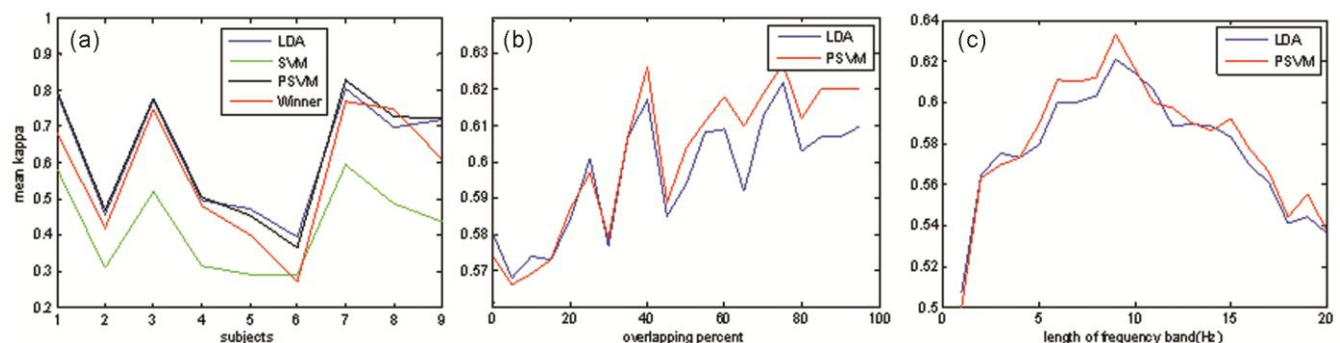


Fig. 3—The results a) the κ mean for all nine subjects of feature Hjorth and three classifiers and winner b) effect of various overlapping on BCI performance c) effect of the length of frequency band on mean kappa. Performance is decreased for long and short bands.

of LDA. There are some fluctuations on κ in low overlapping percentages.

The amplitude of fluctuations decreases in high overlapping percentages. The value of κ for PSVM is higher than 0.61 for overlapping percentage above 60%. We vary the frequency bands length from 1 to 20 Hz with an increment 1 Hz and 80% overlapping. Figure 3c indicates the result of this experiment. The BCI performance is decreased for long bands because of short features vectors. The BCI performance is decreased for short bands because of incorrect feature estimation in EEG rhythms. Therefore a frequency band length of 4-14 Hz is appropriate.

Conclusion

Using overlapped frequency bands in FBCSP shows better results than FBCSP without overlapping. We gained $\kappa=0.62$ compared to 0.57 for winner on dataset IIa of BCI competition IV. Long and short frequency bands make a decreasing in motor imagery performance. Both low and high overlapping percentage declines the value of κ . The overlapping above 60% and frequency bands of 4 to 14 Hz provides better results. We compared three classifiers PSVM, LDA and SVM. PSVM and LDA shows better results than SVM, and PSVM slightly better than LDA.

References

- 1 Datta M V, Gowda D.V, Shanmukha K & Reddy G P, Complications and risk management of diabetic foot ulcer: A review, *J Sci Ind Res*, **3(3)** (2014)363–371.
- 2 Berger H, Uber das Elektrenkephalogramm des Menschen, *Archiv für Psychiatrie und Nervenkrankheiten*, **87(1)** (1929)527–570.
- 3 potentials within several trials using Infomax ICA algorithm, *J Sci Ind Res*, **71** (2012)468–473.
- 4 Koles Z J, Lazar M S & Zhou S Z, Spatial Patterns Underlying Population Differences in the Background EEG, *Brain Topography*, **2(4)** (1990)275-284.
- 5 Ramoser H, Gerking J M & Pfurtscheller G, Optimal spatial filtering of single trial EEG during imagined hand movement, *IEEE Trans on Rehabil Engine*, **8(4)** (2000)441–446.
- 6 Lemm S, Blankertz B, Curio G & Muller K R, Spatospectral filters for improving the classification of single trial EEG, *IEEE Transactions on Biomedical Engineering*, **52(9)** (2005)1541–1548.
- 7 Ang K K, Chin Z Y, Wang C, Guan C & Zhang H, Filter bank common spatial pattern algorithm on BCI competition IV Dataset 2a and 2b, *Frontiers in NEUROSCIENCE*, **6**(2012).
- 8 Asensio-Cubero J, Gan J Q & Palaniappan R, A study on temporal segmentation strategies for extracting common spatial patterns for brain computer interfacing, *Proc of UKCI*, 2011.
- 9 Barachant A, Bonnet S, Congedo M & Jutten C, Multiclass Brain–Computer Interface Classification by Riemannian Geometry, *IEEE Trans Biomed Engine*, **59(4)** (2012)920-928.
- 10 Yang Y, Chevallier S, Wiart J & Bloch I, Automatic selection of the number of spatial filters for motor-imagery BCI, *Eur Symp on Artificial Neural Networks, Comput Intell and Machine Learning(ESANN)*, Bruges, Belgium, (April 25-27 2012).
- 11 Liu C, Wang H & Lu Z, EEG Classification for Multiclass Motor Imagery BCI, *IEEE Control and Decision Conf*, 4450-4453, Guiyang, (May 25-27 2013).
- 12 Llera A, Gomez V & Kappen H J, Adaptive Multi-class classification for Brain Computer Interfaces, *NEURAL COMPUTATION*, **26(14)** (2014)1108-1127.
- 13 Arun S, Sundaraj K & Murugappan M, Hypovigilance detection using energy of electrocardiogram signals, *JSIR*, **71** (2012)794–799.
- 14 Hjorth B, EEG analysis based on time-domain properties, *Electroencephalog and Clin Neurophysiol*, **29** (1970) 306–310.