Removing Unrelated Features Based on Linear Dynamical System for Motor-Imagery-Based Brain-Computer Interface

Jie Wu[1], Li-Chen Shi[1], and Bao-Liang Lu[1,2]

[1]Center for Brain-Like Computing and Machine Intelligence
Department of Computer Science and Engineering
[2]MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems
Shanghai Jiao Tong University
800 Dong Chuan Rd., Shanghai 200240, China
bllu@sjtu.edu.cn

Abstract. Common spatial pattern (CSP) is very successful in constructing spatial filters for detecting event-related synchronization and event-related desynchronization. In statistics, a CSP filter can optimally separate the motor-imagery-related components. However, for a single trail, the EEG features extracted after a CSP filter still include features not related to motor imagery. In this study, we introduce a linear dynamical system (LDS) approach to motor-imagery-based brain-computer interface (MI-BCI) to reduce the influence of these unrelated EEG features. This study is conducted on a BCI competition data set, which comprises EEG signals from several subjects performing various movements. Experimental results show that our proposed algorithm with LDS performs better than a traditional algorithm on average. The results reveal a promising direction in the application of LDS-based approach to MI-BCI.

Keywords: motor imagery, brain-computer interface, linear dynamic system, common spatial pattern.

1 Introduction

Brain-computer interfaces (BCIs) are communication systems which enable users to send commands to computers using only their brain activity, which is generally measured by electroencephalography (EEG) [1,2]. BCI technology has been a promising tool for disabled people as well as for healthy people [3,4,5]. Motor imagery is a very popular paradigm in BCI. EEG and event-related synchronization/desynchronization (ERS/ERD) [6] have been employed for research on brain functional activity for many decades and have become the scientific basis of motor imagery. Studies have shown that distinct phenomena such as ERD/ERS are detectable from EEGs for both real and imagined motor movements in healthy subjects [7,8,9]. Common spatial pattern (CSP) [10] is very successful in constructing spatial filters for detecting ERS/ERD. However, the features extracted after CSP still contain unrelated EEG features.
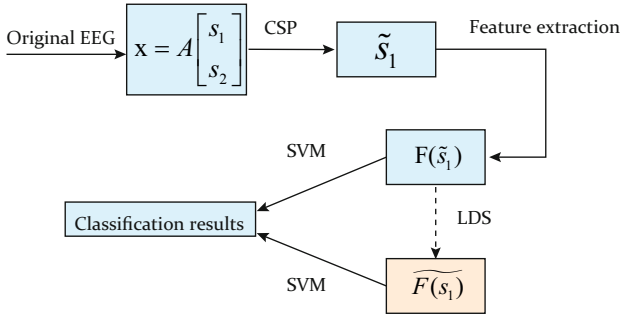
Fig. 1. The flow chart of our proposed algorithm for processing two-class motor imagery EEG recordings

Mental states have the characteristic of continuity. It is a gradual process, and so the EEG features extracted from mental states such as vigilance change continuously [15]. Recently, Shi and Lu [14] have applied linear dynamical system (LDS) approach to vigilance estimation [11,12,13] from EEGs, and their experimental results show that LDS can remove vigilance-unrelated signals effectively. LDS is a kind of state space model, which can effectively remove vigilance-unrelated features using the time dependency of vigilance changes. Motor imagery is also a kind of mental state which has the feature of continuity. In theory, by making use of the time dependency of changes of motor imagery, LDS can filter the motor-imagery-related EEG features more accurately. In this study, we introduce the LDS-based approach to motor-imagery-based brain-computer interface (MI-BCI). By using the LDS-based approach, EEG features are smoothed and the unrelated EEG influences in the EEG features are reduced. Our experimental results show that our proposed algorithm with LDS performs with a higher accuracy than the traditional algorithm on average.

The remainder of this paper is organized as follows. Section 2 describe the methodology and process of our proposed algorithm. Section 3 presents the experimental results. Finally, Section 4 discusses some conclusions.

2   Methodology

2.1   Main Idea

The flow chart of the process for our proposed algorithm with LDS in MI-BCI is shown in Fig. 1. The M recorded EEG signals $x(t) = [x_1(t), x_2(t), ..., x_M(t)]^T$ are assumed to be linear mixtures of the underlying components $s(t) = [s_1(t), s_2(t)]^T$:

$$x = A \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \tag{1}$$

Suppose s_1 are motor-imagery-related components and s_2 are unrelated components. In traditional algorithms, we use CSP to extract the motor-imagery related components $\widetilde{s_1}$ which is an estimate of s_1. Then we extract features $F(\widetilde{s_1})$
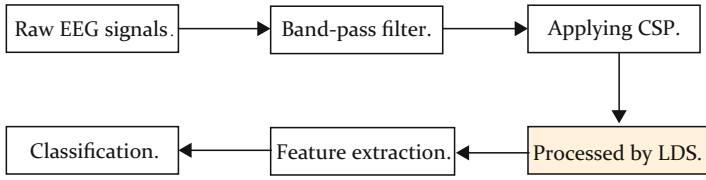
Fig. 2. Architecture of motor imagery-based brain-computer interface with linear dynamical system approach

from components $\widetilde{s_1}$ where $F(\widetilde{s_1})$ is calculated by feature extraction methods. Suppose $F(s_1)$ are the features extracted from s_1 which are pure motor-imagery related components with no noise, then in the traditional algorithms, we use $F(\widetilde{s_1})$ to estimate $F(s_1)$ and finally use a classifier such as a support vector machine (SVM) [16] to classify $F(\widetilde{s_1})$.

In traditional algorithms, a CSP filter is used to optimally separate the motor-imagery related components of left and right motor imagery in statistics [10]. However, for each single trail, the component $\widetilde{s_1}$ filtered by the CSP algorithm may contain unrelated component. So $\widetilde{s_1}$ may not estimate s_1 very precisely. As a result, the features $F(\widetilde{s_1})$ extracted from $\widetilde{s_1}$ may include features unrelated to motor imagery.

Based on this hypothesis, we try to smooth $F(\widetilde{s_1})$ into $\widetilde{F(s_1)}$ using LDS to reduce the unrelated features. Compared with $F(\widetilde{s_1})$, $\widetilde{F(s_1)}$ is expected to be a better estimate of $F(s_1)$. The architecture of the MI-BCI with LDS is shown in Fig. 2.

2.2 Common Spatial Patterns

The CSP algorithm is effective in constructing optimal spatial filters that discriminate two classes of EEG measurements in MI-BCI [10,17,18,19]. The spatial filter maximizes the variance of signals of one class and at the same time minimizes the variance of signals of the other class. Because band power is equal to the variance of band-pass filtered signals, CSP performs very well as a spatial filter for detecting ERS/ERD in EEG measurements and has been well used in in BCI systems [20,21,22].

In this study, we extract one feature every 0.2 second, so a sequence of features $Y(i)$ is obtained in a single trail (7 seconds). The feature $Y(i)$ is obtained by calculating the variance of signals in the time interval of 0.2 second. Let $Y = \frac{1}{n}\sum_{i=1}^{n} Y(i)$, so we can choose $\bar{Y}$ as the extracted feature of the trail. Let $\bar{Y}_{classA}$ and $\bar{Y}_{classB}$ be the selected features for two classes which are chosen from four classes (left hand, right hand, tongue, and foot movements). Since CSP maximizes the variance ratio of components of the two classes, we can classify $\bar{Y}_{classA}$ and $\bar{Y}_{classB}$ by SVM.

2.3   Linear Dynamical System

There are some motor-imagery-unrelated influences in the EEG features, so $y(i)$ probably contains noise. We design LDS to reduce these influences as well as to smooth the EEG features.

The motor-imagery-unrelated influences of EEG features result in a difference between the original EEG features Y calculated from traditional methods and the motor-imagery-related EEG features Y_m which can represent the feature of motor imagery more accurately. Because mental state is time dependent, the features Y_m extracted from EEG components are also time dependent. If Y is considered as the observation sequence of the latent state sequence Y_m, we can represent a state space model to filter out the above influences and recover the Y_m from Y in the form of LDS:

$$Y_m(t) = AY_m(t-1) + v(t) \tag{2}$$
$$Y(t) = CY_m(t) + w(t), \tag{3}$$

where A is the state transition matrix, C is the observation matrix, $v(t) \sim \mathcal{N}(0, \Gamma)$ and $w(t) \sim \mathcal{N}(0, \Sigma)$ are the Gaussian variables, and the initial latent state is assumed to be distributed as $Y_m(1) \sim \mathcal{N}(\mu(0), V(0))$, Eqs. (2) and (3) can also be expressed in an equivalent form in terms of Gaussian conditional distributions as follows,

$$p(Y_m(t) \mid Y_m(t-1)) = \mathcal{N}(AY_m(t-1), \Gamma) \tag{4}$$
$$p(Y(i) \mid Y_m(t)) = \mathcal{N}(CY_m(t)\Sigma). \tag{5}$$

The parameters of the LDS model are denoted by $\theta = \{A, C, \Gamma, \Sigma, \mu_0, V_0\}$. According to the LDS model, given the observations, the latent state $Y_m(t)$ can be estimated from the posterior marginal distribution corresponding to $p(Y_m(t) \mid Y)$ and this posterior marginal distribution is Gaussian,

$$p(Y_m(t) \mid Y) = \mathcal{N}(\mu(t), V(t)).$$

The mean $\mu(i)$ is just the maximum a posteriori (MAP) estimation of $Y_m(t)$. For online inference, Y are the observations from $Y(1)$ to $Y(t)$. The parameters of the marginal distribution, $\hat{\mu}(t)$ and $\hat{V}(t)$, can be determined by the following forward recursions:

$$P_{t-1} = A\hat{V_{t-1}}A^T + \Gamma$$
$$K_t = P_{t-1}C^T(CP_{t-1}C^T + \Sigma)^{-1}$$
$$\hat{\mu}_t = A\hat{\mu}_{t-1} + K_t(Y^t - CA\hat{\mu}_{t-1})$$
$$\hat{V}_t = (I - K_tC)P_{t-1},$$

where the initial conditions are:

$$K_1 = V_0C^T(CV_0C^T + \Sigma)^{-1}$$
$$\hat{\mu}_1 = \mu_0 + K_1(Y^1 - C\mu_0)$$
$$\hat{V}_1 = (I - K_1C)V_0.$$

For offline inference, Y is the whole sequence of observations from Y^1 to Y^N. The parameters of the marginal distribution, $\tilde{\mu}(t)$ and $\tilde{V}(t)$, can be determined by the online inference results and the following backward recursions:

$$J_t = \hat{V}_t A^T (P_t)^{-1}$$
$$\tilde{\mu}_t = \hat{\mu}_t + J_t(\tilde{V}_{t+1} - A\hat{\mu}_t)$$
$$\tilde{V}_t = \hat{V}_t + J_t(\tilde{V}_{t+1} - P_t)J_t^T,$$

where the initial conditions are:

$$\tilde{\mu}_N = \hat{\mu}_N$$
$$\tilde{V}_N = \hat{V}_N.$$

Though offline inference is more accurate than online inference, we use online inference in this study, because immediate feedback is required in MI-BCI. Every EEG feature with interval of 0.2 second is smoothed by LDS. The LDS model runs in constant time because the size of input is limited which is the length of a sequence of features obtained in a single trail (7 seconds). So the time complexity of the LDS model is $O(1)$. The parameters of the LDS model can be estimated by the EM algorithm. However, these estimated parameters are locally optimized. Because the form of parameters is relatively simple, we can test by hand and determine the parameters.

3   Experimental Results

This section evaluates the performance of the proposed algorithm on BCI Competition 3 data set 3a [23]. This data set comprises EEG signals from three subjects who performed left hand, right hand, tongue, and foot movements. The four classes of movements that should be discriminated were paired in six groups to yield the 2-class motor imagery data sets. For the first two subjects, there are 90 trails for each class and for third subject, 60. These data sets comprise a training set and a testing set for each subject. Half of each session is the training set and the other half is the testing set. Each trail has a duration of 7 sec. The subjects performed motor imagery from time $t = 3$ sec to $t = 7$ sec of each trail.

We compare the traditional algorithm with the proposed algorithm with LDS to see whether there is an increase in accuracy. The architecture of the CSP algorithm for two-class motor imagery in this study is shown in Fig. 3. We filter the EEG signals in 8-20 Hz and then use CSP to construct the spatial filter. We decompose the signals into 6 bands of 2 Hz, which are 8-10 Hz, 10-12 Hz, 12-14 Hz, 14-16 Hz, 16-18 Hz, and 18-20 Hz, respectively. Signals of each band are filtered by CSP, and features of 6 bands, $Y_{8-10}, Y_{10-12}, Y_{12-14}, Y_{14-16}, Y_{16-18}, Y_{18-20}$, are extracted by calculating the variance of components which are obtained after CSP. Then, $(Y_{8-10}, Y_{10-12}, Y_{12-14}, Y_{14-16}, Y_{16-18}, Y_{18-20})$ as used as the feature of a single trail. And, SVM is used as a classifier with radial basis function (RBF) kernel, and 5-fold cross validation for training. We need to estimate
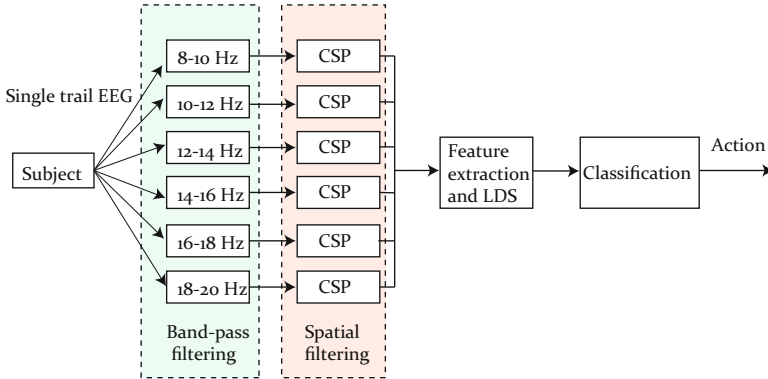
Fig. 3. Architecture of CSP algorithm for two-class motor imagery EEG data

Table 1. Comparison of classification accuracy of the traditional algorithm and that of our proposed algorithm

Data set	Task 1/2	Task 1/3	Task 1/4	Task 2/3	Task 2/4	Task 3/4	Average	Γ/C
k3b(original)	88.9%	85.5%	88.9%	80.0%	78.9%	64.4%	81.1%	
k3b(with LDS)	**91.1%**	**86.7%**	**92.2%**	**81.1%**	**83.3%**	**65.6%**	**83.3%**	0.25
k6b(original)	63.3%	58.3%	86.7%	66.7%	90.0%	90.0%	75.8%	
k6b(with LDS)	**65.0%**	**65.0%**	86.7%	**70.0%**	90.0%	90.0%	**77.8%**	0.03
l1b(original)	65.0%	66.7%	75.0%	68.3%	78.3%	63.3%	69.4%	
l1b(with LDS)	**75.0%**	66.7%	75.0%	**71.7%**	78.3%	**65.0%**	**72.0%**	0.50

the parameters $\theta = \{A, C, \Gamma, \Sigma, \mu_0, V_0\}$ of the LDS model. Considering the form of parameters is relatively simple, they can be determined by hand. We only need to enumerate Γ and C, while the other parameters A, Σ, μ_0 and V_0 have constant values. We set $A = 1, \Sigma = 1, V_0 = 0.1$ and let μ_0 be the the first value of the input sequence to the LDS model.

The experimental results of classification and values of parameter Γ/C are shown in Table 1. The first six columns show the results for the two-class classification of recorded data, while the last column holds the average classification accuracy over all tasks. Comparing the proposed algorithm with the traditional algorithm, the performance for each subject steadily is improved about 2%. However, the overall classification accuracies of l1b and k6b are still relatively low. The reason could be the poor performance of the subject on the practical motor imagery task. In summary, from the experimental results it is clear that LDS is effective for reducing motor-imagery-unrelated EEG features in MI-BCI.

4   Conclusions

In motor imagery, CSP filters can optimally separate the motor-imagery related components of two classes, but for a single trail, the components filtered by

CSP still contain unrelated component. In this paper, we introduced LDS to MI-BCI, which is used to filter the motor-imagery-related EEG features more accurately for a single trail and improve the classification accuracy of MI-BCI. The traditional algorithm and the proposed algorithm with LDS were evaluated on the four-class motor imagery data of BCI competition 3 data set 3a. The experimental results show that after adding the LDS in the traditional algorithm, the average classification accuracy rises about 2%. The experimental results suggest LDS can effectively filter out the motor-imagery-unrelated EEG features and reveal a promising direction in the application of LDS to MI-BCI.

References

1. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain-computer interfaces for communication and control. Clin. Neurophysiol. 113, 767–791 (2002)
2. Pfurtscheller, G., Neuper, C., Guger, C., Harkam, W., Ramoser, R., Schll, A., Obermaier, B., Pregenzer, M.: Current Trends in Graz Brain-computer Interface (BCI). IEEE Trans. Rehab. Eng. 8(2), 216–219 (2000)
3. Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kubler, A., Perelmouter, J., Taub, E., Flor, H.: A spelling device for the paralysed. Nature 398, 297–298 (1999)
4. Curran, E.A., Stokes, M.J.: Learning to control brain activity: A review of the production and control of EEG components for driving brain-computer interface (BCI) systems. Brain Cogn. 51, 326–336 (2003)
5. Dornhege, G., del R. Millan, J., Hinterberger, T., McFarland, D., Muller, K.-R. (eds.): Toward Brain-Computer Interfacing. MIT Press, Cambridge (2007)
6. Pfurtscheller, G., Lopes da Silva, F.H.: Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin. Neurophysiol. 110(11), 1842–1857 (1999)
7. Stavrinou, M., Moraru, L., Cimponeriu, L., Della Penna, S., Bezerianos, A.: Evaluation of Cortical Connectivity During Real and Imagined Rhythmic Finger Tapping. Brain Topogr. 19(3), 137–145 (2007)
8. McFarland, D., Miner, L., Vaughan, T., Wolpaw, J.: Mu and Beta Rhythm Topographies During Motor Imagery and Actual Movements. Brain Topogr. 12(3), 177–186 (2000)
9. Pfurtscheller, G., Brunner, C., Schlogl, A., Lopes da Silva, F.H.: Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. NeuroImage 31(1), 153–159 (2006)
10. Ramoser, H., Muller-Gerking, J., Pfurtscheller, G.: Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans. Rehab. Eng. 8, 441–446 (2000)

11. Cajochen, C., Zeitzer, J.M., Czeisler, C.A., Dijk, D.J.: Dose-response Relationship for Light Intensity and Ocular and Electroencephalographic Correlates of Human Alertness. Behavioural Brain Research 115, 75–83 (2000)
12. Molloy, R., Parasuraman, R.: Monitoring an automated system for a single failure: vigilance and task complexity effects. Human Factors 38, 311–322 (1996)
13. Weinger, M.B.: Vigilance, Boredom, and Sleepiness. Journal of Clinical Monitoring and Computing 15, 549–552 (1999)
14. Shi, L.C., Lu, B.L.: Off-line and on-line vigilance Estimation based on linear dynamical system and manifold learning. In: Proceedings of 32nd International Conference of the IEEE Engineering in Medicine and Biology Society, Buenos Aires, Argentina, pp. 6587–6590 (2010)
15. Oken, B.S., Salinsky, M.C., Elsas, S.M.: Vigilance, alertness, ro sustained attention: physiological basis and measurement. Clinical Neurophysiology 117, 1885–1901 (2006)
16. Lotte, F., et al.: A Review of Classification Algorithms for EEG-Based Brain-Computer Interface. J. Neural. Eng. 4, R1–R13 (2007)
17. Pfurtscheller, G., Neuper, C.: Motor imagery and direct braincomputer communication. Proc. IEEE 89(7), 1123–1134 (2001)
18. Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., Muller, K.-R.: Optimizing Spatial Filters for Robust EEG Single-Trial Analysis. IEEE Signal Proc. Magazine 25(1), 41–56 (2008)
19. Muller-Gerking, J., Pfurtscheller, G., Flyvbjerg, H.: Designing optimal spatial filters for single-trial EEG classification in a movement task. Clin. Neurophysiol. 110, 787–798 (1999)
20. Koles, Z.J.: The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG. Electroencephalogr. Clin. Neurophysiol. 79(6), 440–447 (1991)
21. Guger, C., Ramoser, H., Pfurtscheller, G.: Real-time EEGanalysis with subject-specific spatial patterns for a Brain Computer Interface (BCI). IEEE Trans. Neural Sys. Rehab. Eng. 8(4), 447–456 (2000)
22. Blankertz, B., Dornhege, G., Krauledat, M., Muller, K.-R., Curio, G.: The non-invasive Berlin Brain-Computer Interface: Fast Acquisition of Effective Performance in Untrained Subjects. NeuroImage 37(2), 539–550 (2007)
23. BCI competition III website: `http://www.bbci.de/competition/iii/`