

Research Article

Low-Rank Linear Dynamical Systems for Motor Imagery EEG

Wenchang Zhang,^{1,2} Fuchun Sun,¹ Chuanqi Tan,¹ and Shaobo Liu¹

¹The State Key Laboratory of Intelligent Technology and Systems, Computer Science and Technology School, Tsinghua University, FIT Building, Beijing 100084, China

²Institute of Medical Equipment, Wandong Road, Hedong District, Tianjin, China

Correspondence should be addressed to Fuchun Sun; fcsun@mail.tsinghua.edu.cn

Received 18 September 2016; Revised 14 November 2016; Accepted 16 November 2016

Academic Editor: Feng Duan

Copyright © 2016 Wenchang Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The common spatial pattern (CSP) and other spatio-spectral feature extraction methods have become the most effective and successful approaches to solve the problem of motor imagery electroencephalography (MI-EEG) pattern recognition from multichannel neural activity in recent years. However, these methods need a lot of preprocessing and postprocessing such as filtering, demean, and spatio-spectral feature fusion, which influence the classification accuracy easily. In this paper, we utilize linear dynamical systems (LDSs) for EEG signals feature extraction and classification. LDSs model has lots of advantages such as simultaneous spatial and temporal feature matrix generation, free of preprocessing or postprocessing, and low cost. Furthermore, a low-rank matrix decomposition approach is introduced to get rid of noise and resting state component in order to improve the robustness of the system. Then, we propose a low-rank LDSs algorithm to decompose feature subspace of LDSs on finite Grassmannian and obtain a better performance. Extensive experiments are carried out on public dataset from “BCI Competition III Dataset IVa” and “BCI Competition IV Database 2a.” The results show that our proposed three methods yield higher accuracies compared with prevailing approaches such as CSP and CSSP.

1. Introduction

With the development of the simpler brain rhythm sampling technique and powerful low-cost computer equipment over the past two decades, a noninvasive brain-computer interface (BCI) called electroencephalography (EEG) has attracted more and more attention than other BCIs such as magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and near infrared spectroscopy (NIRS). Among various EEG signals, certain neurophysiological patterns can be recognized to determine the user's intentions such as visual evoked potentials (VEPs), P300 evoked potentials, slow cortical potentials (SCPs), and sensorimotor rhythms. EEG brings hope to patients with amyotrophic lateral sclerosis, brainstem stroke, and spinal cord injury [1]. Motor imagery (MI), which is known as the mental rehearsal of a motor act without real body movement execution, represents a new approach to access the motor system for rehabilitation at all stages of the stroke recovery. People with severe motor disabilities can use EEG-BCI to realize the communication

and control and even to restore their motor disabilities [2, 3]. Therefore, an increasing number of researchers are working on MI-BCI for stroke patient rehabilitation [4, 5].

MI-BCI concentrates on sensorimotor μ - or β -rhythms that has the phenomenon known as event-related synchronization (ERS) or event-related desynchronization (ERD). However, the MI pattern recognition is still a challenge due to the low signal-to-noise ratio, highly subject-specific data, and low processing speed. For these reasons, more and more digital signal processing (DSP) methods and machine learning algorithms are applied to the MI-BCI analysis. Unlike static signals such as images and semantics, the EEG signals are dynamic that lie in a spatiotemporal feature space. Thus a large variety of feature extraction algorithms are proposed, including power spectral density (PSD) values [6, 7], autoregressive (AR) parameters [8, 9], and time-frequency features [10]. For MI-BCI pattern recognition, there are mainly three types of methods: autoregressive components (AR) [11], wavelet transform (WT) [12, 13], and CSP [14, 15]. Because of effectiveness and simplicity in extracting

spatial features, CSP becomes one of the most popular and successful solutions for MI-BCI analysis according to the winners' methods analysis of "BCI Competition III Dataset IVa" [16, 17] and "BCI Competition IV Database 2a" [18, 19]. Therefore, many researchers devote themselves to improving the original CSP method for better performances, such as common spatio-spectral pattern (CSSP) [20], common sparse spectral spatial pattern (CSSSP) [21], subband common spatial pattern (SBCSP) [22], filter bank common spatial pattern (FBCSP) [23], wavelet common spatial pattern (WCSP) [24], and separable common spatio-spectral patterns (SCSSP) [25]. Most of these improved CSP methods fuse spectral and spatial characteristics in the spatio-spectral feature space and finally achieve success by comparison experiments.

Despite its effectiveness in extracting features of MI-BCI, CSP needs a lot of preprocessing and postprocessing such as filtering, demean, and spatio-spectral feature fusion, which influence the classification accuracy easily. In this paper, we utilize linear dynamical systems (LDSs) for processing EEG signals in MI-BCI. Although LDSs succeed in the field of control, to the best of our knowledge, this model has barely been tried in the feature extraction of EEG analysis so far. Compared with CSP method, LDSs have the following advantages: first, LDS can simultaneously generate spatio-spectral dual-feature matrix; second, there is no need to preprocess or postprocess signals, and the raw data can be directly fed into the model; third, it is easy to use and of low cost; last, the extracted features from the LDS are much more effective for classification.

Furthermore, we apply low-rank matrix decomposition approaches [26–28] that have the ability to learn representational matrix even in presence of corrupted data. The noise of the data can be get rid of, hence to improve the robustness. However, there are two ways for the EEG low-rank decomposition. One aims at the EEG raw data; the other aims at features extracted from LDSs, which is proposed by us and called low-rank LDSs (LR-LDSs).

This paper mainly has the following contributions. (1) We utilize LDSs for MI-EEG feature extraction to solve the MI pattern recognition problem. (2) Low-rank matrix decomposition method is applied to improve the robustness for the raw data analysis. (3) We propose LR-LDSs on finite Grassmannian feature space. (4) Plenty of comparison experiments demonstrate the effectiveness of these approaches.

The rest of this paper is organized as follows. Section 2 provides LDSs model to realize the feature extraction of EEG signals. Section 3 presents low-rank matrix decomposition method for the EEG raw data analysis. Section 4 introduces LR-LDSs method. Then, the proper classification algorithm is explained in Section 5. Section 6 compares the three proposed methods (LDSs, LR+CSP, and LR-LDSs) with other state-of-the-art algorithms in different databases. Finally, the summary and conclusion are presented in Section 7.

2. LDSs Modeling

LDSs, also known as linear Gaussian state-space models, have been used successfully in modeling and controlling dynamical systems. In recent few years, more and more

problems extending to computer vision [29, 30], speech recognition [31], and tactile perception [32] have been solved by LDSs model. EEG signals are sequences of brain electron sampling that have typical dynamic textures. We present the features of EEG dynamic textures by LDSs modeling and apply machine learning (ML) algorithms to capture the essence of dynamic textures for feature extraction and classification.

Let $\{Y(t)\}_{t=1,\dots,\tau}$, $Y(t) \in R^m$ be a sequence of τ EEG signal sample at each instant of time t . If there is a set of n spatial filters $\varphi_\alpha : R \rightarrow R^m$, $\alpha = 1, \dots, n$, we have $x(t) = \sum_{i=1}^k A_i x(t-i) + Bv(t)$ with $A_i \in R^{n \times n}$, $B \in R^{n \times n_v}$, independent and identically distributed realization item $v(t) \in R^{n_v}$ and suppose that sequence of observed variables $Y(t)$ can be represented approximately by function of dimensional hidden state $x(t)$, $y(t) = \varphi(x(t)) + \omega(t)$, where $\omega(t) \in R^m$ is an independent and identically distributed sequence drawn from a known distribution resulting in a positive measured sequence. We redefine the hidden state of $x(t)$ to be $[x(t)^T \ x(t-1)^T \ \dots \ x(t-k)^T]^T$ and consider a linear dynamic system as an autoregressive moving average process without firm input distribution:

$$\begin{aligned} x(t+1) &= Ax(t) + Bv(t), \\ y(t) &= \varphi(x(t)) + \omega(t), \\ x(0) &= x_0 \end{aligned} \quad (1)$$

with $v(t)$, $\varphi(x(t))$ distribution unknown, however.

In order to solve the above problem, we can regard it as a white and zero-mean Gaussian noise linear dynamical system and propose a simplified and closed-form solution:

$$\begin{aligned} x(t+1) &= Ax(t) + Bv(t) \quad v(t) \sim N(0, Q), \\ y(t) &= Cx(t) + \omega(t) + \bar{y} \quad \omega(t) \sim N(0, R), \\ x(0) &= x_0, \end{aligned} \quad (2)$$

where $A \in R^{n \times n}$ is the transition matrix that describes the dynamics property, $C \in R^{m \times n}$ is the measurement matrix that describes the spatial appearance, $\bar{y} \in R^m$ is the mean of $y(t)$, and $v(t)$ and $\omega(t)$ are noise components. We should estimate the model parameters A, C, Q, R from the measurements $y(1), \dots, y(\tau)$ and transform them into the maximum-likelihood solution:

$$\begin{aligned} \widehat{A}(\tau), \widehat{C}(\tau), \widehat{Q}(\tau), \widehat{R}(\tau) \\ = \operatorname{argmax}_{A, C, Q, R} p(y(1), \dots, y(\tau)), \end{aligned} \quad (3)$$

and, however, optimal solutions of this problem bring computational complexity.

We apply matrix decomposition to simplify the computation by the closed-form solution. The singular value

decomposition (SVD) solution is the best estimate of C in Frobenius function:

$$\begin{aligned} \widehat{C}(\tau), \widehat{X}(\tau) &= \underset{C, X}{\operatorname{argmin}} \|W(\tau)\|_F \\ \text{subject to } Y(\tau) &= CX(\tau) + W(\tau); \\ C^T C &= I. \end{aligned} \quad (4)$$

Let $Y = U\Sigma V^T$, and we get the parameter estimation of \widehat{C} , \widehat{X} :

$$\begin{aligned} \widehat{C} &= U, \\ \widehat{X} &= \Sigma V^T, \end{aligned} \quad (5)$$

where $\widehat{X} = [X(1), X(2), \dots, X(\tau)]$. \widehat{A} can be determined by Frobenius:

$$\widehat{A}(\tau) = \underset{A}{\operatorname{argmin}} \|X_2(\tau) - AX_1(\tau - 1)\|_F, \quad (6)$$

where $X_2(\tau) = [X(2), X(3), \dots, X(\tau)]$. So the solution is in closed-form using the state estimated

$$\begin{aligned} \widehat{A}(\tau) &= [X(:, 2) \ X(:, 3) \ \cdots \ X(:, \tau)] \\ & * [X(:, 1) \ X(:, 2) \ \cdots \ X(:, \tau - 1)]^\dagger, \end{aligned} \quad (7)$$

where \dagger denotes matrix pseudoinverse.

We can obtain the result $[A, C]$, a couple of spatiotemporal feature matrix. The MATLAB program of LDSs can be found in Supplementary Material algorithm 1 available online at <http://dx.doi.org/10.1155/2016/2637603>.

3. Low-Rank Matrix Decomposition

EEG signals have poor quality because they are usually recorded by electrodes placed on the scalp in a noninvasive manner that has to cross the scalp, skull, and many other layers. Therefore, they are moreover severely affected by background noise generated either inside the brain or externally over the scalp. Low-rank (LR) matrix decomposition can often capture the global information by reconstructing the top few singular values and the corresponding singular vectors. This method is widely applied in the field of image denoising and face recognition (FR). Concretely, low-rank (LR) matrix recovery seeks to decompose a data matrix X into $A + E$, where A is a low-rank matrix and E is the associated sparse error. Candès et al. [33] propose to relax the original problem into the following tractable formulation:

$$\begin{aligned} \min_{A, E} \|A\|_* + \alpha \|E\|_1 \\ \text{s.t. } X &= A + E, \end{aligned} \quad (8)$$

where the nuclear norm $\|A\|_*$ (the sum of the singular values) approximates the rank of A and the l_1 -norm $\|E\|_1$ is sparse constraint.

Then, Zhang and Li [34] decompose each image into common component, condition component, and a sparse

residual. Siyahjani et al. [35] introduce the invariant components to the sparse representation and low-rank matrix decomposition approaches and successfully apply to solve computer vision problems. They add orthogonal constraint to assume that invariant and variant components are linear independent. Therefore, we decompose EEG signals as a combination of three components: resting state component, motor imagery component represented by low-rank matrix, and a sparse residual. However, in practice, it needs some digital signal processing (DSP), that is, wavelet transform or discrete Fourier transform before decomposition. Particularly, raw time-domain signals without any preprocessing are not suitable for low-rank matrix decomposition directly. The training dataset X can be decomposed by $X := A + B + E$, where $A \in R^{m \times n}$ is a low-rank matrix and collects event-related EEG signal components, $B \in R^{m \times n}$ approximates invariant and denotes resting state signal components that are sampled by subjects without any motor imagery, and $E \in R^{m \times n}$ is the matrix of sparse noise. Therefore, training dataset can be decomposed as the following formulation:

$$X := A + B + E. \quad (9)$$

On ideal condition, each sampling channel of subject's brain EEG signals in resting state is similar. In other words, sum of each different A row is minimum. B should add common constraint as the following formulation:

$$\sum_{i \neq j} \|B_i - B_j\|_F^2. \quad (10)$$

We propose optimization problem formulation as

$$\min_{A, B, E} \|A\|_* + \alpha \|E\|_1 + \beta \sum_{i \neq j} \|B_i - B_j\|_F^2 \quad (11)$$

$$\text{s.t. } X = A + B + E.$$

Then, augmented Lagrange multiplier (ALM) [36] method is utilized to solve the above problem. The augmented Lagrangian function $L(A, B, E, \lambda)$ is given by

$$\begin{aligned} L(A, B, E, \lambda) &= \|A\|_* + \alpha \|E\|_1 + \beta \sum_{i \neq j} \|B_i - B_j\|_F^2 \\ &+ \langle \lambda, X - A - B - E \rangle \\ &+ \frac{\mu}{2} \|X - A - B - E\|_F^2, \end{aligned} \quad (12)$$

where μ is a positive scalar and λ is a Lagrange multiplier matrix. We employ an inexact ALM (IALM) method described in Algorithm 1 to solve this problem, where $J(X) = \max(\operatorname{lansvd}(X), \alpha^{-1} \|X\|_F)$ in the initialization $\lambda_0 = X/J(X)$ and $\operatorname{lansvd}(\cdot)$ computes the largest singular value.

When low-rank matrix A denoting event-related EEG signal components are generated, we can utilize some feature extraction methods such as CSP and CSSP to classify MI-BCI. In other words, low-rank matrix decomposition method in this section can be considered as a preprocessing part before feature extraction and classification.

```

Input: Observation matrix  $X$ ,  $\lambda$ , penalty weights  $\alpha, \beta$ 
(1)  $\lambda_0 = X/J(X)$ ;  $\mu_0 > 0$ ;  $\rho > 1$ ;  $k = 0$ .
(2) while not converged do
(3) // Lines (4)–(12) solve  $(A_k, B_k, E_k) = \operatorname{argmin}_{A,B,E} L(A, B, E, \lambda_k, \mu_k)$ 
(4)  $j = 0$ ;  $A_k^0 = A_k$ ;  $B_k^0 = B_k$ ;  $E_k^0 = E_k$ 
(5) while not converged do
(6) // Line (7)–(8) solves  $A_{k+1} = \operatorname{argmin}_A L(A, B_k, E_k, \lambda_k, \mu_k)$ 
(7)  $(U, S, V) = \operatorname{svd}(X - B_{k+1} - E_{k+1} + \lambda_k/\mu_k)$ 
(8)  $A_k^{j+1} = US_{\mu_k^{-1}}[S]V^T$ 
(9) // Line (10) solves  $B_{k+1} = \operatorname{argmin}_B L(A_{k+1}, B, E_k, \lambda_k, \mu_k)$ 
(10)  $E_k^{j+1} = S_{\alpha\mu_k^{-1}}[X - A_k^{j+1} - B_k^j + \lambda_k/\mu_k]$ 
(11) Update  $B_k^{j+1}$  by solving  $E_{k+1} = \operatorname{argmin}_E L(A_{k+1}, B_{k+1}, E, \lambda_k, \mu_k)$ 
(12)  $j \leftarrow j + 1$ 
(13) end while
(14)  $A_{k+1} = A_k^{j+1}$ ,  $B_{k+1} = B_k^{j+1}$ ,  $E_{k+1} = E_k^{j+1}$ 
(15)  $\mu_{k+1} = \rho\mu_k$ ,  $\lambda_{k+1} = \lambda_k + \mu_k(X - A_{k+1} - B_{k+1} - E_{k+1})$ 
(16)  $k \leftarrow k + 1$ 
(17) end while
Output:  $(A_k, B_k, E_k)$ .

```

ALGORITHM 1: Low-rank decomposition via the inexact ALM method.

4. LR-LDSs on Finite Grassmannian

Beginning at an initial state x_1 , the expected observation sequence generated by a time-invariant model $M = (A, C)$ is obtained as $E[y_1, y_2, y_1, \dots] = [C^T, (CA)^T, (CA^2)^T, \dots]^T x_1$ that lies in the column space of the extended observability matrix given by $O_\infty^T = [C^T, (CA)^T, (CA^2)^T, \dots]^T \in R^{\infty \times n}$. LDSs can apply the extended observability subspace \mathcal{O} as descriptor, but it is hard to calculate. Turaga et al. [37, 38] approximate the extended observability by taking the L -order observability matrix; that is, $O(n, L) = [C^T, (CA)^T, \dots, (CA^{L-1})^T]^T$. In this way, an LDS model can be alternately identified as an n -dimensional subspace of R^{Lm} .

Given a database of EEG, we can estimate LDSs model and calculate the finite observability matrix that span subspace as a point on the Riemannian manifold. Then, based on low-rank and sparse matrix decomposition, observability matrix O can be decomposed into $D + E$ as the following formulation:

$$\begin{aligned} \min_{D,E} \quad & \|D\|_* + \alpha \|E\|_1 \\ \text{s.t.} \quad & O = D + E, \end{aligned} \quad (13)$$

where D is a low-rank matrix and E is the associated sparse error.

The inexact ALM method can be also used to solve the optimization problem like Algorithm 1. The output D represents low-rank descriptor for LDSs and can be employed for the classification of EEG trails.

5. Classification Algorithm

We extract features by the above LDSs model and get two feature matrices A and C . Unfortunately, A and C have different modal properties and dimensionalities. So they cannot

be represented directly by a feature vector. Riemannian geometry metric for the space of LDSs is hard to determine and needs to satisfy several constraints. Common classifiers such as Nearest Neighbors (NNs), Linear Discriminant Analysis (LDA), and Support Vector Machines (SVM) cannot classify features in matrix form. The feature matrix must be mapped to vector space. We use Martin Distance [39, 40], which is based on the principal angles between two subspaces of the extended observability matrices, as kernel to present distance of different LDS feature matrix. It can be defined as

$$D^2(\Theta_a, \Theta_b) = -2 \sum_{i=1}^n \log \lambda_i, \quad (14)$$

where $\Theta_a = \{C_a, A_a\}$, $\Theta_b = \{C_b, A_b\}$. λ_i is the eigenvalue solving as the following equation:

$$\begin{aligned} \begin{bmatrix} 0 & \mathcal{O}_{ab} \\ (\mathcal{O}_{ab})^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \lambda \begin{bmatrix} \mathcal{O}_{aa} & 0 \\ 0 & \mathcal{O}_{bb} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ \text{s.t. } x^T \mathcal{O}_{aa} x &= 1, \quad y^T \mathcal{O}_{bb} y = 1, \end{aligned} \quad (15)$$

where the extended observability matrices $\mathcal{O}_a = [C_a^T, A_a^T C_a^T, \dots, (A_a^T)^n C_a^T]$, $\mathcal{O}_b = [C_b^T, A_b^T C_b^T, \dots, (A_b^T)^n C_b^T]$, $\mathcal{O}_{ab} = (\mathcal{O}_a)^T \mathcal{O}_b$. Algorithm 2 in Supplementary Material presents Martin Distance function programmed by MATLAB.

We can classify EEG signals by comparing Martin Distance between training data and testing data. Nearest two samples mean that they may be of the same class. So the forecast label and predict accuracy can be calculated. Algorithm 3 in Supplementary Material is the classification method of KNN.

Considering LR-LDSs methods generating A on Finite Grassmannian, unlike two feature matrices (A, C) by LDSs, Euclidean Distance and Mahalanobis Distance can describe

TABLE 1: Experimental accuracy results (%) obtained from each subject in BCI Competition III Dataset IVa for CSP, CSSP, and our proposed algorithm (LDS).

Subject	aa	al	av	aw	ay	Mean
CSP	71.43	94.64	61.22	89.28	73.02	77.918
CSSP	77.68	96.43	63.27	90.63	79.37	81.476
LDSs	78.57	96.43	64.29	90.18	79.76	81.846
LR+CSP	77.68	96.43	63.78	90.18	79.76	81.566
LR-LDSs	79.46	98.21	63.78	90.18	80.56	82.438

the distance between two feature spaces of EEG trails after LR-LDS. They are simple, efficient, and common for measuring distance between two points. In order to improve the accuracy of classification, we can also employ metric learning methods using the label information to learn a new metric or pseudometric such as neighborhood components analysis and large margin nearest neighbor.

6. Experimental Evaluation

From the above sections, we propose three methods for EEG pattern recognition: LDSs, LR+CSP, and LR-LDSs. Two datasets of motor imagery EEG including BCI Competition III Dataset IVa and BCI Competition IV Database 2a are used to evaluate our three methods compared with other state-of-the-art algorithms such as CSP and CSSP. All experiments are carried out with MATLAB on Intel Core i7, 2.90-GHz CPU with 8 GB RAM.

6.1. BCI Competition III Dataset IVa. Dataset IVa is recorded from five healthy subjects, labeled as “aa,” “al,” “av,” “aw,” and “ay,” with visual cues indicated for 3.5 s performing right hand and foot motor imagery. The EEG signal has 118 channels and markers that indicate the time points of 280 cues for each subject, band-pass filtered between 0.05 and 200 Hz, and downsampled to 100 Hz.

Before feature extracting for comparison experiment, the raw data needs some preprocessing. Firstly, we extract a time segment located from 0.5 to 3 s and employ FastICA to remove artifacts arising from eye and muscle movements. Secondly, we chose 21 channels over the motor cortex (CP6, CP4, CP2, C6, C4, C2, FC6, FC4, FC2, CPZ, CZ, FCZ, CP1, CP3, CP5, C1, C3, C5, FC1, FC3, and FC5) that related to motor imagery.

In order to improve the performance of CSP and CSSP, we apply Butterworth filter for EEG signals filtering within a specific frequency band between 8 and 30 Hz, which encompasses both the alpha rhythm (8–13 Hz) and the beta rhythm (14–30 Hz) that relate to motor imagery. Then, we program MATLAB code to get spatial filter parameters and feature vectors by variance. Finally, a LDA classifier is used to find a separating hyperplane of the feature vectors.

In LDSs model, the value of a hidden parameter describing dimension of Riemannian feature space is closely related to final accuracy. We chose the highest accuracy performance subject “al” and the lowest accuracy performance subject “av” to show the relationship between hidden parameter and

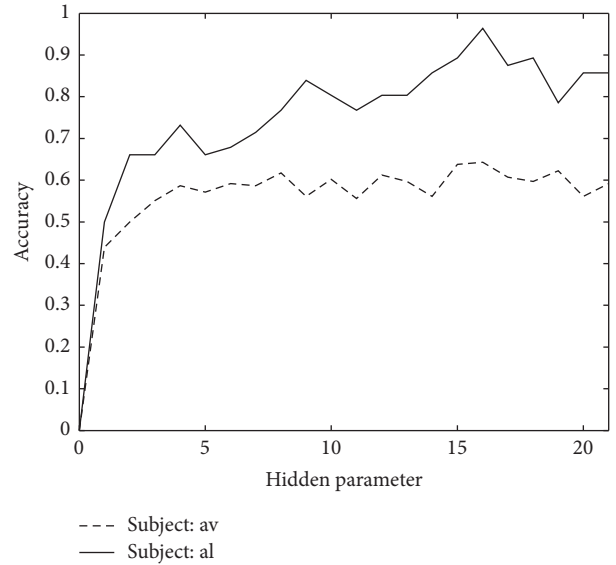


FIGURE 1: The relationship between hidden parameter and accuracy for LDSs. We choose “al” and “av,” which are the highest and lowest accuracy performance, respectively, to show the relationship between hidden parameter and accuracy.

classification accuracy. The result of experiment is presented in Figure 1, which indicates that the accuracy tends to increase when the value of hidden parameter augments approximately and the highest accuracy happens near hidden parameter value of 16.

Then five methods including CSP, CSSP, LDSs, LR+CSP, and LR-LDSs are compared with each other. The results are listed in Table 1.

From Table 1, the bold figures present the best performance results. LR-LDSs are in the majority. The last row shows that the mean of LR-LDS classification accuracy is much better than CSP and a little higher than the others. Comparing with CSP and LR+CSP, LR method is very efficient and useful to improve accuracy. LDSs related methods outperform CSP and CSSP due to their both spatial and temporal features extraction.

6.2. BCI Competition IV Database 2a. Database 2a consists of EEG data from 9 subjects. There are four different motor imagery tasks including movement of the left hand, right hand, both feet, and tongue. At the beginning of each trial, a fixation cross and a short acoustic warning tone appear.

TABLE 2: Experimental accuracy results (%) obtained from each subject in BCI Competition IV Database 2a for CSP, CSSP LDSs, LR+CSP, and LR-LDSs methods.

Subject	A01E	A02E	A03E	A04E	A05E	A06E	A07E	A08E	A09E	Mean
CSP	90.27	53.13	91.67	71.18	61.11	64.24	79.86	91.32	92.36	77.24
CSSP	90.97	56.94	92.01	72.92	61.81	65.28	79.86	93.06	92.71	78.40
LDSs	91.67	55.56	93.06	74.31	62.50	70.83	80.56	93.75	93.06	79.48
LR+CSP	92.01	58.68	95.14	74.65	61.81	65.28	81.25	94.44	93.40	79.63
LR-LDSs	92.01	59.02	94.44	75.35	63.19	69.44	81.25	95.14	93.06	80.32

After two seconds the subject is cued by an arrow pointing to either the left, right, down, or up that denote the movement of left hand, right hand, foot, or tongue for 1.25 s. Then the subjects carry out the motor imagery task for about 3 s. The BCI signals are sampled by 25 channels including 22 EEG channels and 3 EOG channels with 250 Hz and bandpass-filtered between 0.5 Hz and 100 Hz.

Different from dataset IVa, database 2a is a multiclassification problem. However, LDA is a two-class classifier. Therefore, we choose K -NN algorithms for CSP, CSSP, LDSs, LR+CSP, and LR-LDSs methods uniformly. Table 2 describes the classification accuracies results of five above concerned methods. Similar to the results of BCI Competition III Dataset IVa, the mean accuracies of LDSs, LR+CSP, and LR-LDSs are higher than CSP and CSSP methods. Furthermore, LR-LDSs method abstains the best performance.

7. Conclusion

CSP has gained much success in the past MI-BCI research. However, it is reported that CSP is only a spatial filter and sensitive to frequency band. It needs prior knowledge to choose channels and frequency bands. Without preprocessing, the result of classification accuracy may be poor. LDSs can overcome these problems by extracting both spatial and temporal features simultaneously to improve the classification performance. Furthermore, we utilize a low-rank matrix decomposition approach to get rid of noise and resting state component in order to improve the robustness of the system. Then LR+CSP and LR-LDSs methods are proposed. Comparison experiments are demonstrated on two datasets. The major contribution of our work is realization of LDSs model and LR algorithm for MI-BCI pattern recognition. The proposed LR-LDSs methods achieve a better performance than CSP and CSSP.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant nos. 91420302 and 91520201).

References

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [2] K. K. Ang and C. Guan, "Brain-computer interface in stroke rehabilitation," *Journal of Computing Science and Engineering*, vol. 7, no. 2, pp. 139–146, 2013.
- [3] N. Sharma, V. M. Pomeroy, and J.-C. Baron, "Motor imagery: a backdoor to the motor system after stroke?" *Stroke*, vol. 37, no. 7, pp. 1941–1952, 2006.
- [4] S. Silvoni, A. Ramos-Murguialday, M. Cavinato et al., "Brain-computer interface in stroke: a review of progress," *Clinical EEG and Neuroscience*, vol. 42, no. 4, pp. 245–252, 2011.
- [5] O. Bai, P. Lin, S. Vorbach, M. K. Floeter, N. Hattori, and M. Hallett, "A high performance sensorimotor beta rhythm-based brain-computer interface associated with human natural motor behavior," *Journal of Neural Engineering*, vol. 5, no. 1, pp. 24–35, 2008.
- [6] S. Chiappa and S. Bengio, "HMM and IOHMM modeling of EEG rhythms for asynchronous bci systems," in *Proceedings of the European Symposium on Artificial Neural Networks ESANN*, Bruges, Belgium, April 2004.
- [7] J. D. R. Millán and J. Mouriño, "Asynchronous BCI and local neural classifiers: an overview of the adaptive brain interface project," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 159–161, 2003.
- [8] W. D. Penny, S. J. Roberts, E. A. Curran, and M. J. Stokes, "EEG-based communication: a pattern recognition approach," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 214–215, 2000.
- [9] G. Pfurtscheller, C. Neuper, A. Schlogl, and K. Luger, "Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters," *IEEE Transactions on Rehabilitation Engineering*, vol. 6, no. 3, pp. 316–325, 1998.
- [10] T. Wang, J. Deng, and B. He, "Classifying EEG-based motor imagery tasks by means of time-frequency synthesized spatial patterns," *Clinical Neurophysiology*, vol. 115, no. 12, pp. 2744–2753, 2004.
- [11] D. J. Krusienski, D. J. McFarland, and J. R. Wolpaw, "An evaluation of autoregressive spectral estimation model order for brain-computer interface applications," in *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '06)*, pp. 1323–1326, IEEE, New York, NY, USA, September 2006.
- [12] T. Demiralp, J. Yordanova, V. Kolev, A. Ademoglu, M. Devrim, and V. J. Samar, "Time-frequency analysis of single-sweep

- event-related potentials by means of fast wavelet transform,” *Brain and Language*, vol. 66, no. 1, pp. 129–145, 1999.
- [13] D. Farina, O. F. do Nascimento, M.-F. Lucas, and C. Doncarli, “Optimization of wavelets for classification of movement-related cortical potentials generated by variation of force-related parameters,” *Journal of Neuroscience Methods*, vol. 162, no. 1-2, pp. 357–363, 2007.
- [14] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, “Optimal spatial filtering of single trial EEG during imagined hand movement,” *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, 2000.
- [15] M. Grosse-Wentrup and M. Buss, “Multiclass common spatial patterns and information theoretic feature extraction,” *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 8, article no. 11, pp. 1991–2000, 2008.
- [16] B. Blankertz, K.-R. Müller, D. J. Krusienski et al., “The BCI competition III: validating alternative approaches to actual BCI problems,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153–159, 2006.
- [17] <http://www.bbci.de/competition/iii/results/>.
- [18] M. Tangermann, K. Müller, A. Aertsen et al., “Review of the BCI competition IV,” *Frontiers in Neuroscience*, vol. 6, article no. 55, 2012.
- [19] <http://www.bbci.de/competition/iv/results/>.
- [20] S. Lemm, B. Blankertz, G. Curio, and K.-R. Müller, “Spatio-spectral filters for improving the classification of single trial EEG,” *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 9, pp. 1541–1548, 2005.
- [21] G. Dornhege, B. Blankertz, M. Krauledat, F. Losch, G. Curio, and K.-R. Müller, “Combined optimization of spatial and temporal filters for improving brain-computer interfacing,” *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 11, pp. 2274–2281, 2006.
- [22] Q. Novi, C. Guan, T. H. Dat, and P. Xue, “Sub-band common spatial pattern (SBCSP) for brain-computer interface,” in *Proceedings of the 3rd International IEEE/EMBS Conference on Neural Engineering (CNE '07)*, pp. 204–207, IEEE, Kohala Coast, Hawaii, USA, May 2007.
- [23] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, “Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface,” in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '08)*, pp. 2390–2397, Hong Kong, China, June 2008.
- [24] E. A. Mousavi, J. J. Maller, P. B. Fitzgerald, and B. J. Lithgow, “Wavelet common spatial pattern in asynchronous offline brain computer interfaces,” *Biomedical Signal Processing and Control*, vol. 6, no. 2, pp. 121–128, 2011.
- [25] A. S. Aghaei, M. S. Mahanta, and K. N. Plataniotis, “Separable common spatio-spectral patterns for motor imagery BCI systems,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 1, pp. 15–29, 2016.
- [26] E. J. Candès, X. Li, Y. Ma, and J. Wright, “Robust principal component analysis?” *Journal of the ACM*, vol. 58, no. 3, article 11, 2011.
- [27] C.-F. Chen, C.-P. Wei, and Y.-C. F. Wang, “Low-rank matrix recovery with structural incoherence for robust face recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 2618–2625, June 2012.
- [28] R. Vidal and P. Favaro, “Low rank subspace clustering (LRSC),” *Pattern Recognition Letters*, vol. 43, no. 1, pp. 47–61, 2014.
- [29] P. Saisan, G. Doretto, Y. N. Wu, and S. Soatto, “Dynamic texture recognition,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '01)*, pp. II-58–II-63, Kauai, Hawaii, USA, December 2001.
- [30] G. Doretto, A. Chiuso, Y. N. Wu, and S. Soatto, “Dynamic textures,” *International Journal of Computer Vision*, vol. 51, no. 2, pp. 91–109, 2003.
- [31] B. Mesot and D. Barber, “Switching linear dynamical systems for noise robust speech recognition,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 15, no. 6, pp. 1850–1858, 2007.
- [32] R. Ma, H. Liu, F. Sun, Q. Yang, and M. Gao, “Linear dynamic system method for tactile object classification,” *Science China: Information Sciences*, vol. 57, no. 12, pp. 1–11, 2014.
- [33] E. J. Candès, X. Li, Y. Ma, and J. Wright, “Robust principal component analysis?” *Journal of the ACM*, vol. 58, no. 3, 2011.
- [34] Q. Zhang and B. Li, “Mining discriminative components with low-rank and sparsity constraints for face recognition,” in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '12)*, pp. 1469–1477, ACM, August 2012.
- [35] F. Siahjani, R. Almohsen, S. Sabri, and G. Doretto, “A supervised low-rank method for learning invariant subspaces,” in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '15)*, pp. 4220–4228, Santiago, Chile, December 2015.
- [36] Z. Lin, M. Chen, and Y. Ma, “The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices,” 2010.
- [37] P. Turaga, A. Veeraraghavan, A. Srivastava, and R. Chellappa, “Statistical computations on grassmann and stiefel manifolds for image and video-based recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 11, pp. 2273–2286, 2011.
- [38] M. Harandi, R. Hartley, C. Shen, B. Lovell, and C. Sanderson, “Extrinsic methods for coding and dictionary learning on Grassmann manifolds,” *International Journal of Computer Vision*, vol. 114, no. 2-3, pp. 113–136, 2015.
- [39] R. J. Martin, “A metric for ARMA processes,” *IEEE Transactions on Signal Processing*, vol. 48, no. 4, pp. 1164–1170, 2000.
- [40] A. B. Chan and N. Vasconcelos, “Classifying video with kernel dynamic textures,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, 6, 1 pages, June 2007.