

Research Article

Multilinear Discriminative Spatial Patterns for Movement-Related Cortical Potential Based on EEG Classification with Tensor Representation

Qian Cai,¹ Jianfeng Yan ², Hongfang Han,² Weiqiang Gong,³ and Haixian Wang ²

¹School of Statistics and Mathematics, Nanjing Audit University, Nanjing 211815, Jiangsu, China

²Key Laboratory of Child Development and Learning Science of Ministry of Education, School of Biological Science & Medical Engineering, Southeast University, Nanjing 210096, Jiangsu, China

³Nanjing Les Information System Technology Company Ltd., Nanjing 210007, Jiangsu, China

Correspondence should be addressed to Jianfeng Yan; 220181834@seu.edu.cn and Haixian Wang; hxwang@seu.edu.cn

Received 29 December 2020; Revised 9 April 2021; Accepted 15 May 2021; Published 27 May 2021

Academic Editor: Anastasios D. Doulamis

Copyright © 2021 Qian Cai 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 discriminative spatial patterns (DSP) algorithm is a classical and effective feature extraction technique for decoding of voluntary finger premovements from electroencephalography (EEG). As a purely data-driven subspace learning algorithm, DSP essentially is a spatial-domain filter and does not make full use of the information in frequency domain. The paper presents multilinear discriminative spatial patterns (MDSP) to derive multiple interrelated lower dimensional discriminative subspaces of low frequency movement-related cortical potential (MRCP). Experimental results on two finger movement tasks' EEG datasets demonstrate the effectiveness of the proposed MDSP method.

1. Introduction

The core task of brain-computer interface (BCI) is to extract specific useful components from complex electroencephalography (EEG) signal. Currently, the commonly extracted EEG components include event-related desynchronization/synchronization (ERD/ERS) [1], steady-state visual evoked potentials (SSVEPs) [2], event-related potentials (ERPs) [3], and movement-related cortical potentials (MRCPs) [4]. Among them, MRCP is a slow negative shift that precedes naturally voluntary movement or motor imaging. This negative shift contains two components, i.e., Bereitschaftspotential (BP) and negativity slope (NS), respectively, generating from 1-2 s and 0.4 s before movement [5]. Furthermore, MRCP is low frequency signal [6] and always submerged with the background noise. It, therefore, is difficult to reconnoiter the signal.

The detection of MRCP has been widely studied by BCI researchers. Particularly, discriminative spatial patterns (DSP) [7] are proposed to enhance the detection of MRCP

patterns. Geometrically, DSP aims to find discriminative directions onto which projected scatters between any EEGs are maximized, and meanwhile, projected scatters within any EEGs are minimized [8]. However, most of the current DSP-based methods analyze limited spatial information while ignoring the inherent structure information of original EEG signal.

In real world, the collected physiological signal usually has some specialized structures and these structures are usually in the form of tensors with second or higher order. For example, an original EEG signal is a second-order tensor, i.e., matrix, whose rows and columns represent channels and time series, respectively. When time-frequency analysis is performed on the signal, the third dimension representing frequency information will be shown. Thus, the data driven method which considers the underlying data structure is required when analyzing real physiological signal.

This paper is motivated by the tensor algebra and tensor subspace analysis [9–12]. In this work, the EEG data is firstly

encoded as a tensor with second or higher order by continuous wavelet transform (CWT). Then, tensor-based discriminant analysis theory is explored to optimize subspaces algorithm. To uncover the underlying structures in these problems for EEG analysis, this paper proposes the multilinear discriminative spatial patterns (MDSP) as a subspace learning method that includes classification.

In summary, our main contributions are as follows. Firstly, we extend the DSP algorithm restricted to 2D data to tensor objects of any order. Secondly, an iterative optimization method that alternately solves the problem of optimal projection is customized for MDSP. Finally, against the special form of features extracted by MDSP, we propose a new tensor classification method based on nearest neighbors. The advantages of our MDSP algorithm are as follows:

- (1) MDSP is a general multidimensional dimensionality reduction method. Compared with traditional time-frequency analysis methods, it can avoid the undersample problem (the dimensionality of feature is much larger than the number of training samples) [13] by working on each order of the training tensors separately.
- (2) MDSP derives multiple interrelated lower dimensional discriminative subspaces. These subspaces are not independent of each other, which is in line with the inherent structural characteristics of the actual signal.

2. Related Work

There have been theoretical analyses that provide a multilinear supplement to linear algebra. And, the work of directly using linear algebra is very common. Some such models from [14, 15] propose that linear dimensionality reduction approaches have been extended smoothly to multilinear subspace reduction algorithms, following the principle of multilinear algebra.

More neoterically, the joint solution of tensor and hotspot technologies to solve specific problems has attracted widespread attention. For instance, Makantasis et al. [16, 17] presented a new model that trains the neural network by tensor decomposition and significantly reduces the number of samples required in hyperspectral image classification. Meanwhile, common spatial patterns which are restricted to 2D data have also been extended to the hyperspectral image of arbitrary order [18]. Regarding the problem of multiview clustering, Wu et al. [19] firstly constructed a third-order tensor by stacking similarity matrices from each view and then captured the low-rank structure information by t-SVD in the tensor space. Furthermore, they also attempted to construct new multiview features from view-specific affinity matrix by low-rank tensor learning [20].

Motivated by the significant performance of tensor in the field of computer vision, an increasing number of brain scientists have attempted to apply tensor to BCI community [21, 22]. Zheng et al. [23] proposed a tensor-based multitask learning method to assess human cognitive activities from multiclass EEG signal. To overcome the limitations of noisy

and missing data of EEG, a tensor completion model is designed [24], which utilizes tensor-based algorithms to clean and complete data. Van Eynhoven et al. [25] trained multilinear subspace learning on multichannel EEG and tested on a single channel, thus creating a real breakthrough for practical application.

3. Methods

In this section, we present our tensor-based MDSP method. Firstly, we briefly review the original DSP method. Then, we introduce some concepts and definitions related with this work to help understanding. Finally, the proposed MDSP and classification extension based on tensor are presented with detailed mathematical formulas.

3.1. Discriminative Spatial Patterns. Assume that $X_i \in R^{c \times t}$ is the EEG signal of trial i with c as the number of channels and t is the number of samples in time, $y_i \in \{1, 2, \dots, p\}$ is the class label of X_i , n_j is the number of trials of class j , and $n = \sum_{j=1}^p n_j$ meaning the number of all trials. Thus, the within-class scatter matrix S_w and the between-class scatter matrix S_b are defined as follows:

$$S_w = \sum_{j=1}^p \sum_{i: y_i=j} (X_i - M_j)(X_i - M_j)^T, \quad (1)$$

$$S_b = \sum_{j=1}^p n_j (M_j - M)(M_j - M)^T,$$

where $M_j = (1/n_j) \sum_{i: y_i=j} X_i$ means the average of class j and $M = (1/n) \sum_{i=1}^n X_i$ means the average of all classes. DSP attempts to seek a set of vectors $U (U \in R^{c \times d}, d \leq c)$ to maximize the Fisher criterion given by

$$J(U) = \frac{\text{tr}(U^T S_b U)}{\text{tr}(U^T S_w U)} = \frac{\left\| \sum_{j=1}^p n_j U^T (M_j - M) \right\|^2}{\left\| \sum_{j=1}^p \sum_{i: y_i=j} U^T (X_i - M_j) \right\|^2}, \quad (2)$$

where $\text{tr}(\cdot)$ denotes the operation of matrix trace. According to the Lagrange multiplier method, $U = (u_1, u_2, \dots, u_d)$ can be obtained by computing the equation $S_b u_i = \lambda_i S_w u_i$. As a result, U is the eigenvectors corresponding to the d biggest eigenvalues of matrix $S_w^{-1} S_b$. Thus, the filtered feature $F_i \in R^{d \times t}$ corresponding to X_i are given by

$$F_i = U^T (X_i - M). \quad (3)$$

3.2. Multilinear Algebra. In this paper, we assume the bold uppercase symbols represent tensor objects, for example, $\mathbf{X} \in R^{m_1 \times m_2 \times \dots \times m_h}$ represents an h -order (also called h -mode) tensor. $\mathbf{X}_{i_1, i_2, \dots, i_h}$ is the $(i_1, i_2, \dots, i_h)^{\text{th}}$ element of \mathbf{X} , where $1 \leq i_1 \leq m_1, \dots, 1 \leq i_h \leq m_h$. Here, we first introduce two definitions related to this article.

Definition 1 (mode- k unfolding or unfolding-matricization). The mode- k unfolding is the process of rearranging the elements of a tensor (along the mode- k) to obtain a

matrix. For an n -order tensor $\mathbf{X} \in R^{m_1 \times m_2 \times \dots \times m_k \times \dots \times m_h}$, the mode- k unfolding of this is a matrix $X \in R^{m_k \times (m_1 * m_2 * \dots * m_{k-1} * m_{k+1} * \dots * m_h)}$. The operation of mode- k unfolding is denoted as $X^k \leftarrow_k \mathbf{X}$.

Definition 2 (mode- k product). The mode- k product of a tensor $\mathbf{X} \in R^{m_1 \times m_2 \times \dots \times m_k \times \dots \times m_h}$ and a matrix $U \in R^{m_k \times m'_k}$ is denoted as $\mathbf{F} = \mathbf{X} \times_k U$, $\mathbf{F} \in R^{m_1 \times m_2 \times \dots \times m'_k \times \dots \times m_h}$, where $\mathbf{F}_{i_1, i_2, \dots, i_{k-1}, j, i_{k+1}, \dots, i_h} = \sum_{i=1}^{m_k} \mathbf{X}_{i_1, i_2, \dots, i_{k-1}, i, i_{k+1}, \dots, i_h} * U_{i, j}$, $j = 1, 2, \dots, m'_k$. To simplify the notation of multimode product in this paper, we denote that $\mathbf{X} \times_1 U_1 \times_2 U_2 \dots \times_h U_h = \mathbf{X} \prod_{d=1}^h \times_d U_d$.

3.3. Multilinear Discriminative Spatial Patterns. As an improvement of DSP, MDSP utilizes multiple interrelated subspaces which can collaborate to discriminate different classes. Suppose that $\mathbf{X}_i \in R^{m_1 \times m_2 \times \dots \times m_h}$ is the EEG signal of trial i ; in this paper, $h = 2$ or 3 . The aim of MRCP is to find a set of optimal projection matrices $U_1, U_2, \dots, U_k, \dots, U_h$ ($U_k \in R^{m_k \times m'_k}$) that leads to the most accurate classification in the projected tensor subspace, where

$$\mathbf{F}_i = \mathbf{X}_i \prod_{d=1}^h \times_d U_d, \quad \mathbf{F}_i \in R^{m'_1 \times m'_2 \times \dots \times m'_k \times \dots \times m'_h}. \quad (4)$$

Similar to DSP, MDSP maximizes the between-class scatter and, at the same time, minimizes the within-class scatter measured in each optimal projection matrices. That is,

$$J(U_d |_{d=1}^h) = \arg \max_{U_d |_{d=1}^h} \frac{\left\| \sum_{j=1}^p n_j (\mathbf{M}_j - \mathbf{M}) \prod_{d=1}^h \times_d U_d \right\|^2}{\left\| \sum_{j=1}^p \sum_{i: y_i=j} (\mathbf{X}_i - \mathbf{M}_j) \prod_{d=1}^h \times_d U_d \right\|^2}, \quad (5)$$

where $\mathbf{M}_j = (1/n_j) \sum_{i: y_i=j} \mathbf{X}_i$ is the average tensor of the samples belonging to class j and $\mathbf{M} = (1/n) \sum_{i=1}^n \mathbf{X}_i$ is the total average tensor of all the samples.

Equation (5) is equivalent to a high-order nonlinear optimization problem with high-order nonlinear constraints. Therefore, there are dependencies between different projection matrices to be sought, and it is difficult to find the optimal solution. Consequently, we have to use an iterative optimization method to alternately find the discriminant subspace projection matrices MDSP need. In each iteration, to seek the optimal projection matrix U_k , we assume that other projection matrices $U_1, U_2, \dots, U_{k-1}, U_{k+1}, \dots, U_h$ are known. So, the optimization problem with high-order nonlinear constraints with equation (5) is changed to

$$U_k^* = \arg \max_{U_k} \frac{\left\| \sum_{j=1}^p n_j (\mathbf{M}_j - \mathbf{M}) \prod_{d=1}^h \times_d U_d \right\|^2}{\left\| \sum_{j=1}^p \sum_{i: y_i=j} (\mathbf{X}_i - \mathbf{M}_j) \prod_{d=1}^h \times_d U_d \right\|^2}. \quad (6)$$

In equation (6), U_k is the only unknown variable. Thus, equation (6) aims to pursue the best projection matrices U_k , which maximize the interclass scatter and, at the same time, minimize the intraclass scatter just as the general DSP in

equation (2). The entire iterative procedure of MDSP is listed in Algorithm 1.

3.4. Classification. After learning the best projection matrices, a new high-dimensional sample \mathbf{X} can be rewritten as \mathbf{F} through equation (4). Then, the class label of the new sample is predicted by the average tensor of the closest category, that is,

$$j^* = \arg \min_j \left\| \mathbf{F} - \mathbf{M}_j \times_1 U_1 \dots \times_h U_h \right\|. \quad (7)$$

4. Experiments

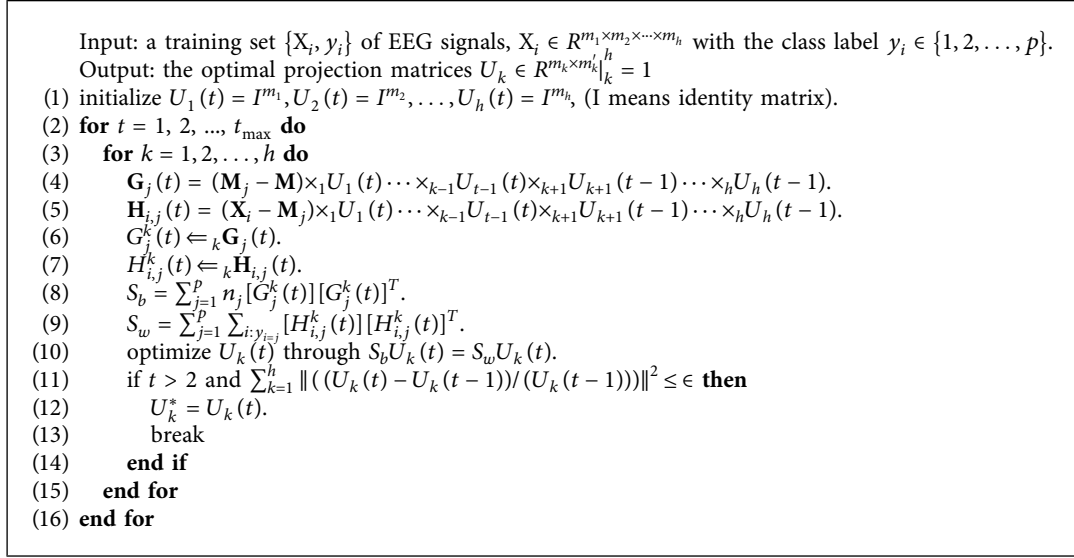
In this section, we evaluate the proposed approach for a binary classification problem. The objective of the experiments is to detect two-class finger tapping activities (i.e., left/right finger tapping) from the EEG signal. In our experiments, the tensor-based scheme for finger tapping classification is shown in Figure 1.

4.1. EEG Datasets. We handle the detection on two datasets with the problem of two-class classification (i.e., left/right finger tapping). We use EEG signal as input data. A brief description of the two datasets is as follows.

Dataset IV < self-paced 1 s > of BCI competition II (also called BCI competition 2003) [26] provided 416 trials of 500 ms length, which were ended 130 ms before pressing a key, and only 316 trials were provided with labels and were regarded as the training set. Therefore, the remaining 100 trials without labels were regarded as the testing set. The EEG signals were recorded at 1000 Hz and in a version down-sampled at 100 Hz (recommended) with a band-pass filter between 0.05 and 200 Hz.

EEG data for voluntary finger tapping movement [27] is from 14 healthy individuals. The EEG signals were recorded for three conditions of right finger tap, left finger tap, and resting state (in this paper, we only focus on right finger tap and left finger tap) with 19 channels and sampling frequency of 1024 Hz. The dataset consists of 40 trials of 6 s in length (-3 s to 3 s) for each condition and each participant. The dataset was shared after preprocessing and artefacts removal using independent component analysis. Figure 2 shows the average waveforms of EEG signals across trials of one second of a subject.

4.2. Preprocessing. For both datasets, a low-pass fifth-order Butterworth filter with cut-off frequency of 7 Hz is applied to obtain MRCPs samples. Then, we use the total segment (-630 ms to -130 ms) and part segment (-330 ms to -130 ms) before the motor movement for features extraction in dataset 1. For dataset 2, the data were subsampled to 100 Hz from the raw signals before filtering by a zero-phase low-pass filter. Each EEG trial was segmented from -2 s to



ALGORITHM 1: Training procedure of multilinear discriminative spatial patterns.

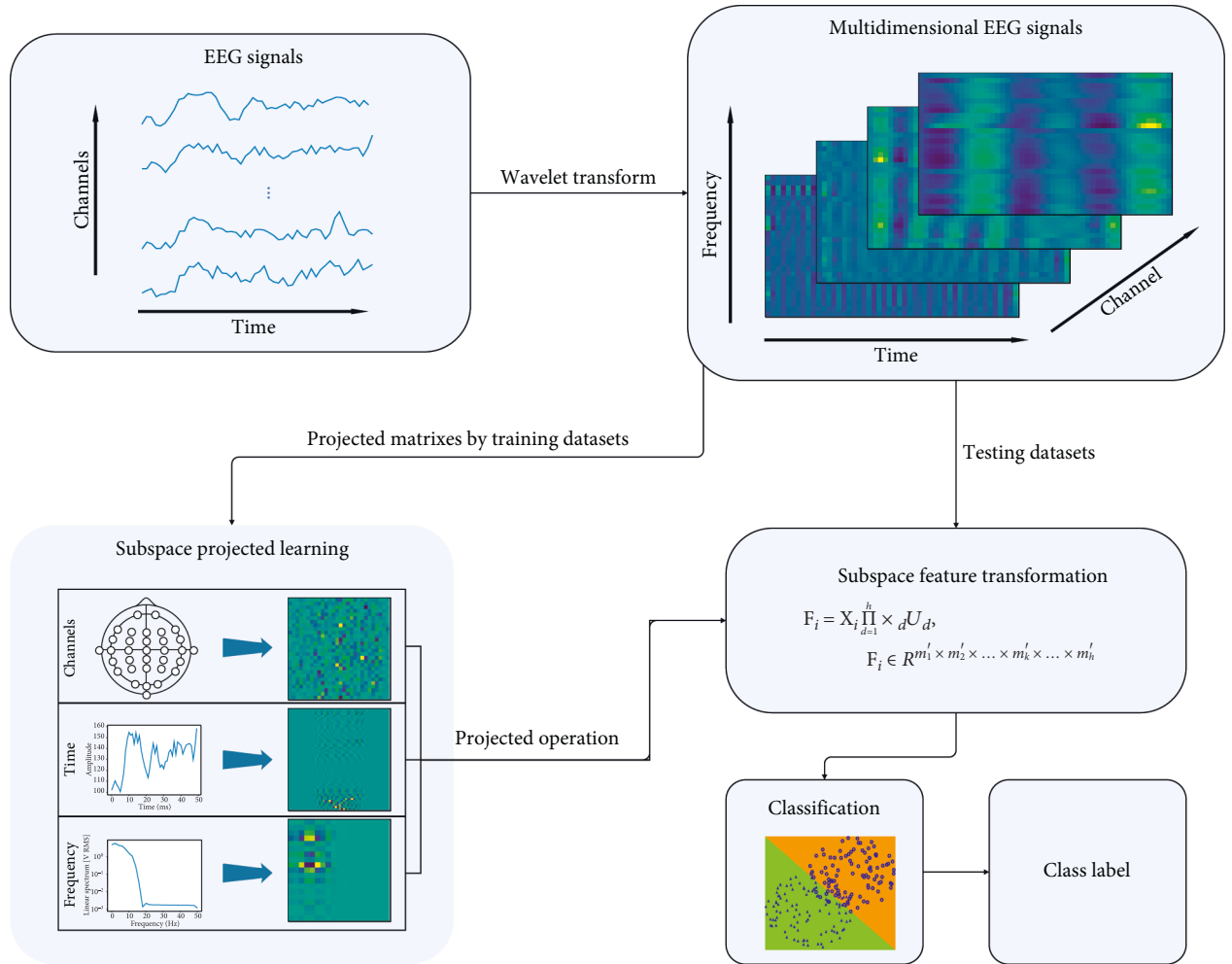


FIGURE 1: The schematic diagram for EEG classification with MDSP.

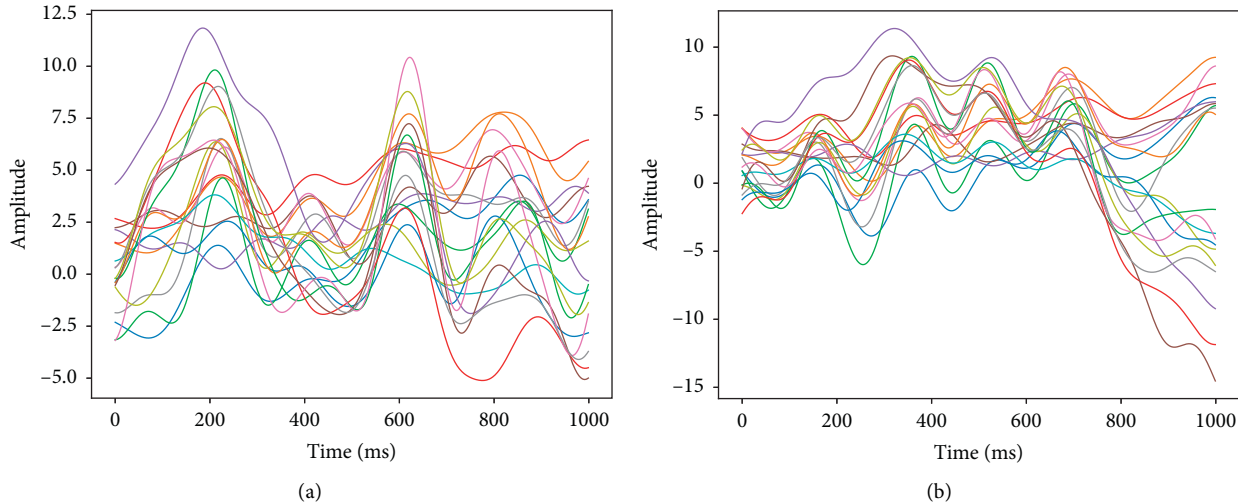


FIGURE 2: The average waveforms of EEG signals across trials of one second of a subject. (a) The waveforms evoked by the right finger tap. (b) The waveforms evoked by the left finger tap.

0 s for subsequent analysis. Then, these trials were divided into 0.5 s sliding window with 0.1 s overlap moving.

In this paper, we utilize CWT to reveal the change of EEG signal in time and frequency domains. CWT measures the similarity between the input signal and the basic signal (also called mother wavelet). In the wavelet transform, we select the complex Gaussian wavelet as mother wavelet and preset the length of scale sequence as 10. Thus, we obtain EEG tensor representation which is three-order shape representing the number of channels, time points, and steps of frequency.

4.3. Parameters' Selection. In this experiment, MDSP needs to choose various parameters, namely, each number of filters in interrelated discriminative subspace, and the start point of the time window. For dataset 1, we seek the optimal parameter combination on the training set and evaluate on the testing set like in the competition. For dataset 2, we choose the optimal parameters by a 5 × 5-fold cross-validation method on the EEG data of each subject.

Figure 3 reveals how different parameters affect the whole process. From Figures 3(a)–3(o), it can be inferred that 2–4 filters in each dimension achieve better results. Redundant filters are insignificant for the experiment and even add redundant information, thereby reducing accuracy. This is consistent with the experience of selecting the number of filters in other EEG-based subspace algorithms, such as common spatial pattern (CSP) and DSP. Furthermore, Figure 3(p) presents the best time window (among 0.1–0.8 s before finger tapping), which is also coincident with the time when the NS potential occurs.

Note that, the threshold ε is used to check convergence of the iterative optimization procedures. The mode- k convergence error at iteration t is

$$\text{err}_k(t) = \left\| \frac{U_k(t) - U_k(t-1)}{U_k(t-1)} \right\|^2. \quad (8)$$

Thus, the total convergence error is $\text{err}(t) = \sum_{k=1}^h \text{err}_k(t)$. Then, the algorithm is treated as convergence when $\text{err}(t) \leq \varepsilon$, where $\varepsilon = 0.01$. The maximum number of iteration t_{\max} is chosen to be 50 in our experiments. While the number of iteration $t \geq t_{\max}$, we think that the procedure on the combination of projection subspace does not converge, namely, filters in different subspaces cannot work well in coordination. Thus, we ignore the combination of parameters.

4.4. Results and Discussion. Table 1 gives the results of the three methods with the best parameters on different time segments for the 100 testing trials. Here, MDSP is referred to as MDSP (2D) and MDSP (3D) for problems with tensor of second and third order, respectively. From Table 1, it is observed that (1) MDSP, whether working on 2D or 3D signals, significantly outperforms DSP in each experimental trial. (2) For MDSP, the total segment 0.5 s yields better result than part segment 0.2 s, which implies that MDSP is insensitive to the choice of time periods.

Note that the recognition rates of most previous research studies require better manual selections of time segments. In our experiments, MDSP has better robustness to the time segments and automatically selects the best time information over a longer time segments, thereby reducing the influence of manual selections. Besides, our experiments focus on only low frequencies MRCP-based EEG classification (rather than joint classification of MRCP and ERD-based) because previous work has proven that MRCP can be well combined with ERD in BCI [7,28].

Figure 4 shows the effect of the number of iterations for MDSP on 2D and 3D signals while assuming that all tensor modes have the same reduced dimensionality for simplicity ($m_k|_{k=1}^h = 4$). Although the convergence of MDSP is not guaranteed from the mathematical formula, Figure 4(a) shows that the total error does not increase sharply after a certain number of iterations on the excellent channel.

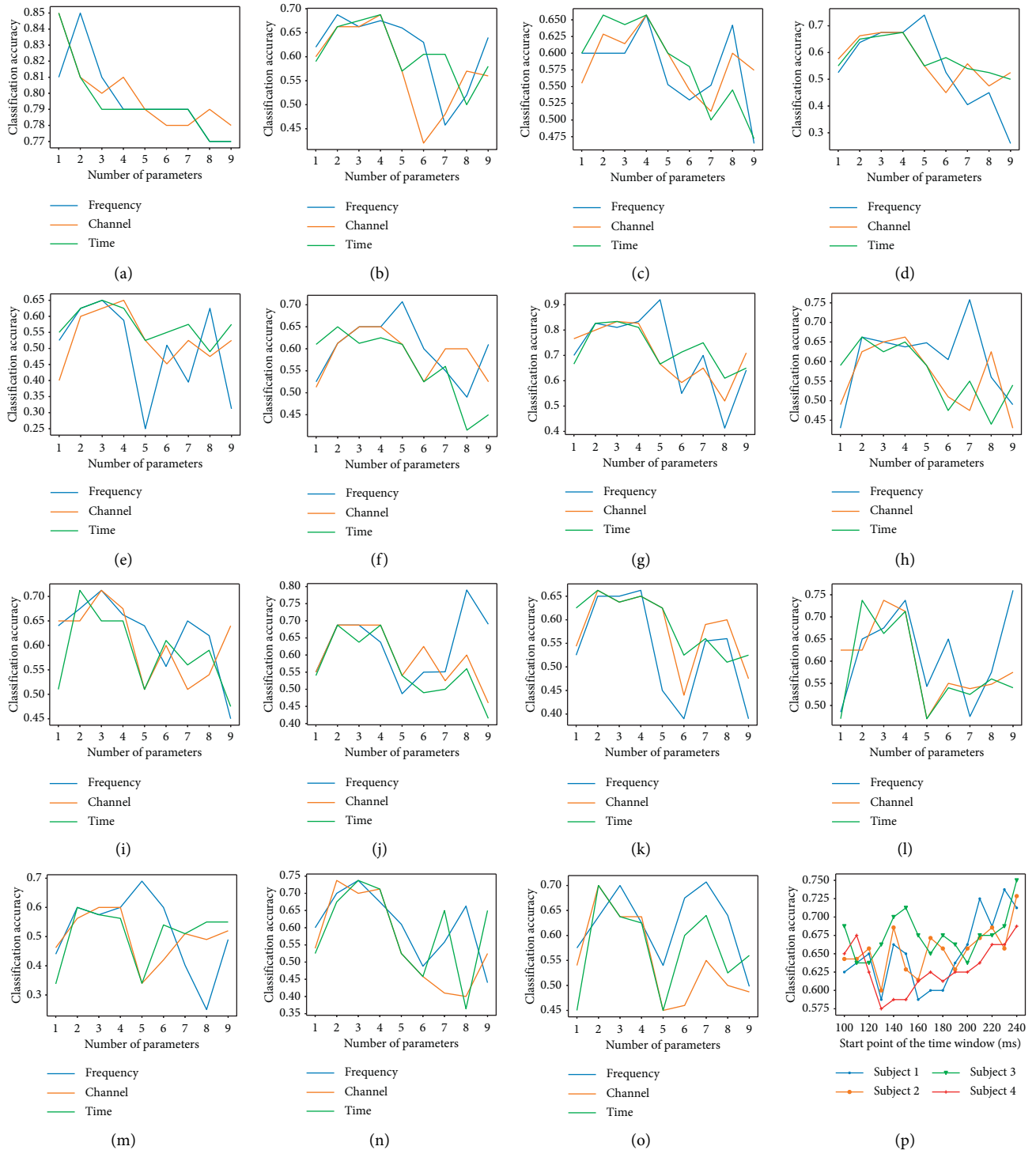


FIGURE 3: The effect of different parameters on performance for MDSP. (a–o) The effect of the number of filters, where figure (a) means subject on dataset 1 and figures (b–o) mean subjects 1–14 on dataset 2, respectively. (p) The effect of the start of time windows of subject 1–4 on dataset 2.

TABLE 1: Comparison (%) for the 100 testing trials with time segments of 0.2 s and 0.5 s.

	Accuracy		Precision		Recall		F1_score	
	0.5 s	0.2 s	0.5 s	0.2 s	0.5 s	0.2 s	0.5 s	0.2 s
DSP	67.0	72.0	66.0	69.1	67.3	77.6	66.6	73.1
MDSP (2D)	77.0	73.0	77.8	72.0	74.1	73.5	74.5	72.7
MDSP (3D)	79.0	75.0	80.4	76.1	75.5	71.4	77.9	73.7

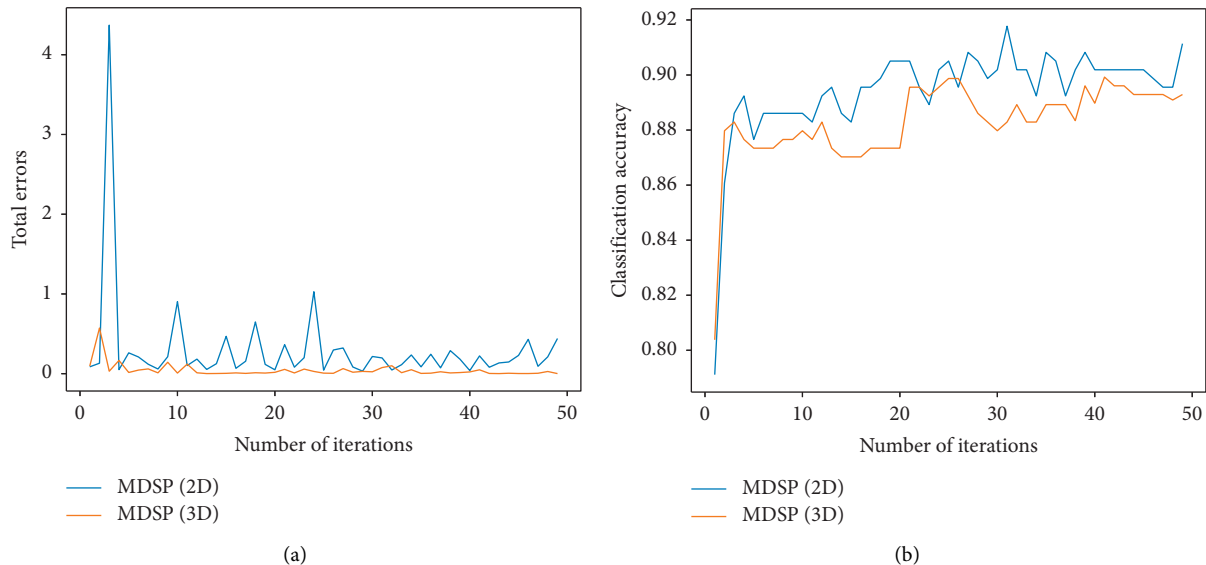


FIGURE 4: The effect of the number of iterations on performance for MDSP. (a) The total error versus the number of iterations. (b) The best accuracy versus the number of iterations.

Additionally, Figure 4(b) displays that MDSP provides a stable classification accuracy as it approaches convergence.

Table 2 lists the results of five DSP-based MRCP on dataset IV, BCI competition II. Among them, RST is named regularized spatial-temporal filter, combining DSP and regularization ideas. AST is named adaptive spatial-temporal filter, computing a spatial filter automatically by Gaussian kernel and linear ridge regression (LRR). PSTF is named pipeline of spatial-temporal filter, combined by a series of systematic approaches. Our result is equal to the best result. Furthermore, MDSP considers the tensor form and can naturally combine regularization terms to improve accuracy, just like RST.

Figure 5 is the boxplot of the classification obtained with the three methods on the 14 subjects in dataset 2. The boxplot demonstrates that MDSP (3D) performs better than MDSP (2D) and DSP, and MDSP (2D) is more stable than DSP though they perform similarly on accuracy. From these, we conclude that the collaboration of multiple subspaces can greatly enhance the classification capability.

Statistical analysis was performed with one-way analysis of variance (ANOVA) and a multiple comparisons procedure was performed as a *post hoc* analysis [31]. The statistical results demonstrated a statistically significant difference in the accuracy for the three methods ($p < 0.05$). Moreover, the *post hoc* analysis showed that MDSP (3D) had a significantly greater accuracy than MDSP (2D) and DSP (highest $p < 0.05$).

TABLE 2: Comparison (%) of classification performance of five studies of DSP-based MRCP.

Methods	Accuracy
DSP	72
RST [29]	79
AST [29]	79
PSTF [30]	75
MDSP	79

It implies that the subspaces derived from frequency-domain have additional discrimination capability compared with time domain and spatial domain.

4.5. Computational Cost. For original EEG signal $X \in R^{c \times t}$, the time complexity of the MDSP is $O(c^3 + t^3)$ for each loop. Specifically, Table 3 compares the time complexity and space complexity of MDSP and DSP. Here, r is the number of iterations that makes the MDSP optimization procedure converge and f is the dimensionality of frequency after CWT.

More generally, for any an n th-order tensor $\mathbf{X} \in R^{m_1 \times m_2 \times \dots \times m_k \times \dots \times m_n}$, the time complexity of the MDSP is $O(r(\sum_{i=1}^n m_i^3))$ and the space complexity $O(\sum_{i=1}^n m_i^2)$. Although the MDSP training procedure requires many loops to converge, it is acceptable for ordinary computer

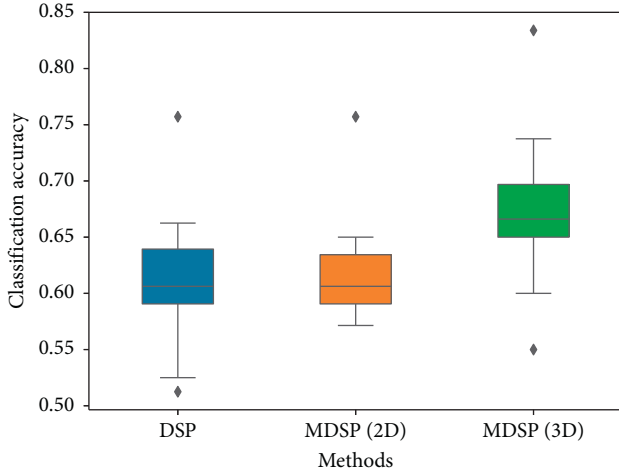


FIGURE 5: Comparison of classification accuracies of all subjects on the dataset 2 by DSP, MDSP (2D), and MDSP (3D).

TABLE 3: Computational complexity analysis.

Methods	Time complexity	Space complexity
DSP	$O(c^3)$	$O(c^2 + t^2)$
MDSP (2D)	$O(r(c^3 + t^3))$	$O(c^2 + t^2)$
MDSP (3D)	$O(r(c^3 + t^3 + f^3))$	$O(c^2 + t^2 + f^2)$

compared with traditional subspace methods, for example, LDA with the time complexity of $O(\prod_{i=1}^h m_i^3)$ and the space complexity of $O(\prod_{i=1}^h m_i^2)$.

5. Conclusions

In this paper, we proposed an effective feature extraction method, called MDSP, to seek a series of optimal interrelated projections for discrimination in multiple lower dimensional tensor subspaces. Compared with DSP, MDSP discovers more useful discriminant information by constructing multiple interrelated subspaces and thus has better discrimination ability.

Data Availability

All data included in this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Nature Science Foundation of China under Grant 61773114 and the University Synergy Innovation Program of Anhui Province under Grant GXXT-2020-015.

References

- [1] M. Wairagkar, Y. Hayashi, and S.-J. Nasuto, "Modeling the ongoing dynamics of short and long-range temporal correlations in broadband EEG during movement," *Frontiers in Systems Neuroscience*, vol. 13, 2019.
- [2] Z. Cao, W. Ding, Y.-K. Wang, F. K. Hussain, A. Al-Jumaily, and C.-T. Lin, "Effects of repetitive ssvsps on EEG complexity using multiscale inherent fuzzy entropy," *Neurocomputing*, vol. 389, pp. 198–206, 2020.
- [3] Y. Zhang, G. Zhou, J. Jin, Q. Zhao, X. Wang, and A. Cichocki, "Aggregation of sparse linear discriminant analyses for event-related potential classification in brain-computer interface," *International Journal of Neural Systems*, vol. 24, no. 1, 2014.
- [4] J.-H. Jeong, N.-S. Kwak, C. Guan, and S.-W. Lee, "Decoding movement-related cortical potentials based on subject-dependent and section-wise spectral filtering," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 3, pp. 687–698, 2020.
- [5] K. Wang, M. Xu, Y. Wang, S. Zhang, L. Chen, and D. Ming, "Enhance decoding of pre-movement EEG patterns for brain-computer interfaces," *Journal of Neural Engineering*, vol. 17, no. 1, 2020.
- [6] A. Shakeel, M.-S. Navid, M.-N. Anwar, S. Mazhar, M. Jochumsen, and I.-K. Niazi, "A review of techniques for detection of movement intention using movement-related cortical potentials," *Computational and Mathematical Methods in Medicine*, vol. 2015, Article ID 346217, 13 pages, 2015.
- [7] X. Liao, D. Yao, D. Wu, and C. Li, "Combining spatial filters for the classification of single-trial EEG in a finger movement task," *IEEE Transactions on Bio-Medical Engineering*, vol. 54, no. 5, pp. 821–831, 2007.
- [8] Q. Tang, J. Wang, and H. Wang, "L1-norm based discriminative spatial pattern for single-trial EEG classification," *Biomedical Signal Processing and Control*, vol. 10, pp. 313–321, 2014.
- [9] A.-J. McConnell, *Applications of Tensor Analysis*, Courier Corporation, North Chelmsford, MA, USA, 2014.
- [10] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.
- [11] J. Wu, S. Qiu, R. Zeng, Y. Kong, L. Senhadji, and H. Shu, "Multilinear principal component analysis network for tensor object classification," *IEEE Access*, vol. 5, pp. 3322–3331, 2017.
- [12] M. Bessaoudi, M. Belahcene, A. Ouamane, A. Chouchane, and S. Bourennane, "Multilinear enhanced Fisher discriminant analysis for robust multimodal 2d and 3d face verification," *Applied Intelligence*, vol. 49, no. 4, pp. 1339–1354, 2019.
- [13] D. Tao, X. Li, X. Wu, and S. J. Maybank, "General tensor discriminant analysis and Gabor features for gait recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 10, pp. 1700–1715, 2007.
- [14] M. Bessaoudi, A. Ouamane, M. Belahcene, A. Chouchane, E. Boutellaa, and S. Bourennane, "Multilinear side-information based discriminant analysis for face and kinship verification in the wild," *Neurocomputing*, vol. 329, pp. 267–278, 2019.
- [15] Z. Lai, Y. Xu, J. Yang, J. Tang, and D. Zhang, "Sparse tensor discriminant analysis," *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, vol. 22, no. 10, pp. 3904–3915, 2013.
- [16] K. Makantasis, A. D. Doulami, N. D. Doulami, and A. Nikitakis, "Tensor-based classification models for

- hyperspectral data analysis,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 12, pp. 6884–6898, 2018.
- [17] K. Makantasis, A. Doulamis, N. Doulamis, A. Nikitakis, and A. Voulodimos, “Tensor-based nonlinear classifier for high-order data analysis,” in *Proceedings of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2221–2225, IEEE, Calgary, Alberta, Canada, April 2018.
- [18] K. Makantasis, A. Doulamis, N. Doulamis, and A. Voulodimos, “Common mode patterns for supervised tensor subspace learning,” in *Proceedings of 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2927–2931, IEEE, Brighton, UK, April 2019.
- [19] J. Wu, Z. Lin, and H. Zha, “Essential tensor learning for multi-view spectral clustering,” *IEEE Transactions on Image Processing*, vol. 28, no. 12, pp. 5910–5922, 2019.
- [20] J. Wu, X. Xie, L. Nie, Z. Lin, and H. Zha, “Unified graph and low-rank tensor learning for multi-view clustering,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 04, pp. 6388–6395, 2020.
- [21] Y. Tang, D. Chen, L. Wang, A. Y. Zomaya, J. Chen, and H. Liu, “Bayesian tensor factorization for multi-way analysis of multi-dimensional EEG,” *Neurocomputing*, vol. 318, pp. 162–174, 2018.
- [22] A. Nikitakis, K. Makantasis, N. Tampouratzis, and I. Papaefstathiou, “A unified novel neural network approach and a prototype hardware implementation for ultra-low power EEG classification,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 4, pp. 670–681, 2019.
- [23] Q. Zheng, Y. Wang, and P. A. Heng, “Multitask feature learning meets robust tensor decomposition for eeg classification,” *IEEE Transactions on Cybernetics*, vol. 51, no. 4, pp. 2242–2252, 2021.
- [24] J. Solé-Casals, C. F. Caiafa, Q. Zhao, and A. Cichocki, “Brain-computer interface with corrupted EEG data: a tensor completion approach,” *Cognitive Computation*, vol. 10, no. 6, pp. 1062–1074, 2018.
- [25] S. Van Eyndhoven, M. Bousse, B. Hunyadi, L. De Lathauwer, and S. Van Huffel, “Single-channel EEG classification by multi-channel tensor subspace learning and regression,” in *Proceedings of the 2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1–6, IEEE, Aalborg, Denmark, UK, September 2018.
- [26] B. Blankertz, K.-R. Müller, G. Curio et al., “The BCI competition 2003: progress and perspectives in detection and discrimination of EEG single trials,” *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1044–1051, 2004.
- [27] M. Wairagkar, *EEG Data for Voluntary Finger Tapping Movement*, University of Reading Dataset, Whiteknights Lake, UK, 2017.
- [28] H. Wang and J. Xu, “Local discriminative spatial patterns for movement-related potentials-based EEG classification,” *Biomedical Signal Processing and Control*, vol. 6, no. 4, pp. 427–431, 2011.
- [29] J. Lu, K. Xie, and D.-J. McFarland, “Adaptive spatio-temporal filtering for movement related potentials in EEG-based brain-computer interfaces,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 847–857, 2014.
- [30] E. A. Zeid, A. R. Sereshkeh, and T. Chau, “A pipeline of spatio-temporal filtering for predicting the laterality of self-initiated fine movements from single trial readiness potentials,” *Journal of Neural Engineering*, vol. 13, no. 6, 2016.
- [31] X. Tang, W. Li, X. Li, W. Ma, and X. Dang, “Motor imagery EEG recognition based on conditional optimization empirical mode decomposition and multi-scale convolutional neural network,” *Expert Systems with Applications*, vol. 149, 2020.