Journal of Medical Systems (2019) 43: 169
https://doi.org/10.1007/s10916-019-1270-0

IMAGE & SIGNAL PROCESSING

Improvement Motor Imagery EEG Classification Based on Regularized Linear Discriminant Analysis

Rongrong Fu[1] · Yongsheng Tian[1] · Tiantian Bao[1] · Zong Meng[1] · Peiming Shi[1]

Received: 20 November 2018 / Accepted: 3 April 2019 / Published online: 7 May 2019

Abstract

Mental tasks classification such as motor imagery, based on EEG signals is an important problem in brain computer interface systems (BCI). One of the major concerns in BCI is to have a high classification accuracy. The other concerning one is with the favorable result is guaranteed how to improve the computational efficiency. In this paper, Mu/Beta rhythm was obtained by bandpass filter from EEG signal. And the classical linear discriminant analysis (LDA) was used for deciding which rhythm can give the better classification performance. During this, the common spatial pattern (CSP) was used to project data subject to the ratio of projected energy of one class to that of the other class was maximized. The optimal projection dimension was determined corresponding to the maximum of area under the curve (AUC) for each participant. Eventually, regularized linear discriminant analysis (RLDA) is possible to decode the imagined motor sensed using electroencephalogram (EEG). Results show that higher classification accuracy can be provided by RLDA. And optimal projection dimensions determined by LDA and RLDA are of consistent solution, this improves computational efficiency of CSP-RLDA method without computation of projection dimension.

Keywords Electroencephalogram classification · Common spatial pattern · Regularized linear discriminant analysis

Introduction

Brain-computer interfaces (BCI) based on sensorimotor rhythms (SMRs) of electroencephalogram (EEG) have developed a direct motor control pathway between a human brain and an external device [1, 2]. Motor imagery is a cognitive task consisting of kinesthetically imagining a motor movement while without executing movement [3], which has been widely used as a communication approach in non-invasive BCI. The signature of motor imagery performance can be reflected in oscillations of Mu and Beta rhythms over cortex.

Pattern recognition techniques are employed for the classification of motor imagery EEG [4–7]. Common spatial pattern method has been used in BCI applications as a signal enhancement method for discrimination of motor imagery task by itself and in combination with other pattern recognition techniques [8, 9]. When EEG measurements are filtered with an inappropriate frequency range, BCI systems based on CSP-feature and pattern classification methods generally yield poor accuracies [10]. Linear discriminant analysis (LDA) relying on CSP features provides a feasible tool to classify motor imagery EEG. Classical LDA aims to find optimal discriminant features by maximizing the ratio of the between-class distance to the within-class distance of a given data set under supervised learning conditions [11]. Classical LDA simply applies an eigen-decomposition on the scatter matrices, however it is difficult to process undersampled data. Classical LDA cannot address the singularity problem, it fails when scatter matrices are singular. Regularized linear discriminant analysis (RLDA) have been proposed to overcome the singularity problem in classical LDA in the past. By increasing the magnitude of the diagonal elements of the scatter matrices (usually by adding a scaled identity matrix), the singularity

This article is part of the Topical Collection on *Image & Signal Processing*

✉ Rongrong Fu
frr1102@aliyun.com

Yongsheng Tian
tian_ys@foxmail.com

Tiantian Bao
baott@stumail.ysu.edu.cn

Zong Meng
mzysu@ysu.edu.cn

Peiming Shi
hspace@ysu.edu.cn

[1] Key Lab of Measurement Technology & Instrumentation of Hebei Province, Yanshan University, Qinhuangdao 066004, China

problem can be addressed by RLDA [10]. Therefore, combined CSP-feature and RLDA method, suitable filtered motor imagery EEG can be classified with better classification performance. Though this CSP-RLDA method can give high performance, it is really time-consuming process. If the parameter of CSP (the dimensionality of projected data) can be determined before using RLDA, this will save a lot running time of CSP-RLDA method for processing three dimensional epoched motor imagery EEG.

Mu rhythm is extracted from motor imagery EEG measurements in the sensory motor area with the frequency of 8~12 Hz. Part of the frequency of the Beta rhythm which is in 18~26 Hz is the harmonic of the Mu rhythm, and it is also related to the movement and the motor imagination. In this paper, the method of CSP-LDA is studied first for addressing two aspect of the classification task. One aspect is the classification results of the two kinds of rhythm (Mu and Beta) can be compared by using the method of CSP-LDA. The other aspect is the effect of the number of dimensionality of CSP projected data can be determined by CSP-LDA method. Experiments show that the dimensionality determined by CSP-LDA have great guiding significant for the CSP-RLDA method. This makes CSP-RLDA mothed can save running time and provide high classification performance at the same time.

CSP-based features

CSP algorithm aims to find the spatial filters that can differentiate two classes of EEG signals [12]. All the EEG trails have been epoched, filtered and saved as three-dimensional tensor, and this tensor is represented as nChannels×nSamples×nTrails. CSP algorithm is applied to a two-class paradigm (right and left motor imagery) to obtain features for EEG classification. The composite spatial covariance $\boldsymbol{C}$ can be calculated as the sum of normalized covariance matrix of two classes ($\boldsymbol{C_1}$ and $\boldsymbol{C_2}$). With the matrixes of eigenvalues ($\boldsymbol{\Lambda}$) and eigenvectors ($\boldsymbol{Uc}$), the whitening transformation can be obtained by $P = \Lambda^{-\frac{1}{2}}U^{\mathrm{T}}$, therefore all eigenvalues of PCP^{T} are equal to one. Similarly, $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$ are transformed as $S_1 = PC_1P^{\mathrm{T}}$ and $S_2 = PC_2P^{\mathrm{T}}$. If $\boldsymbol{S_1}$ is decomposed into $S_1 = B\Lambda_1B^{\mathrm{T}}$, then $S_2 = B\Lambda_2B^{\mathrm{T}}$. The optimal discriminative information of two populations can be achieved by projecting whitened EEG onto the first and last several eigenvectors of $\boldsymbol{B}$ [13]. The spatially filtered signal $\boldsymbol{Z}$ of a single trial EEG with the size of NChannels×NSamples is given by

$$Z = W^{\mathrm{T}}E \tag{1}$$

where $W = B^{\mathrm{T}}P$ is projection matrix. The CSP-based features formed from $\boldsymbol{Z}$ can be extracted as follow,

$$f_p = \log\left(\left(\mathrm{var}(Z_p)\right)/\left(\sum_{i=1}^{2m}\mathrm{var}(Z_i)\right)\right) \tag{2}$$

where $\boldsymbol{Z_p}$ is the first and last m rows of $\boldsymbol{Z}$, i=1,2,…,2 m. The logarithmic transformation makes the variance feature's distributions close to Gaussian [14, 15].

Classical linear discriminant analysis

Given an $m \times n$ data matrix $\boldsymbol{X}$, which is treated as n column vectors $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_n}$ ($\boldsymbol{x} \in \Re^m$), each column corresponds to a data point and each row corresponds to a particular feature. The optimized features $\boldsymbol{y} \in \Re^l$ can be computed by linear transformation matrix $\boldsymbol{A} \in \Re^{m \times l}$ as follows [16],

$$\boldsymbol{A}:\ \boldsymbol{x} \in \Re^m \longrightarrow \boldsymbol{y} = \boldsymbol{A}^T\boldsymbol{x} \in \Re^l \tag{3}$$

The resulting data matrix $\boldsymbol{Y} \in \Re^{l \times n}$ contains l rows which leads to the l-dimensional reduced space, there are l features for each data point. Given the within-class scatter matrix $\boldsymbol{S_w}$, the between-class scatter matrix $\boldsymbol{S_b}$, and the total scatter matrix $\boldsymbol{S}$, S_w^L and S_b^L represent the between-class scatter matrix and within-class scatter matrix in the lower-dimensional space. With the linear transformation $\boldsymbol{A}$, S_w^L and S_b^L become,

$$S_w^L = A^T S_w A \tag{4}$$

$$S_b^L = A^T S_b A \tag{5}$$

An optimal transformation $\boldsymbol{A}$ would maximize S_b^L and minimize S_w^L simultaneously, so that $J(A) = S_b^L/S_w^L = \frac{A^T S_b A}{A^T S_w A}$ scattering matrix criterion involving $\boldsymbol{S_w}$, $\boldsymbol{S_b}$ is maximized. Classical LDA computes the optimal $\boldsymbol{A}$, such that

$$A = \arg\max_A\left[\left(A^T S_w A\right)^{-1}A^T S_b A\right] \tag{6}$$

For each Gaussian class with the common covariance matrix, classical LDA is equivalent to the optimal Bayesian classifier, with a difference of a threshold value. From geometric interpretation, the optimized features $\boldsymbol{y}$ is the projection of $\boldsymbol{x}$ onto the subspace spanned by the eigenvectors of $\left(S_w^L\right)^{-1}S_b^L$.

Materials and methods

Experiments and data sets

Data sets used in this work is motor imagery task from BCI Competition IV. The data was acquired with an EEG array of 59 electrodes at a sampling frequency of 1000 Hz, data sets were band-pass filtered and downsampled at 100 Hz. As a

result, seven data sets (labeled as A, B, C, D, E, F and G) were obtained, each with 100 trails from two out of the three available cues. Special information is the data sets containing both real and artificial data set. Data sets from C, D and E are generated artificial by Guido Nolte and Carmen Vidaurre. The generating way and the true distribution of artificial data were undisclosed for public. The artificial data sets were provided to test the proposed method whether can tell which participants were real and which artificial. The detail description can be found from reference [17].

Epoch the continuous signals of 59 EEG channels into two separable classes, reshape them in two tensors as the number channels the number of samples the number of trials as shown in Fig. 1.

Regularized linear discriminant analysis

Classical LDA cannot handle singular scatter matrices, which limits its applicability to low-dimensional data. To overcome the singularity problem of the within-class scatter matrix $\boldsymbol{S_w}$, the regularization can be done by adding a regularization parameter to the diagonal elements of $\boldsymbol{S_w}$ in the RLDA technique [18, 19], thus regularized scatter matrix is become a non-singular matrix. In this section, the regularized Fisher's criterion becomes,

$$J(A,\alpha) = \frac{A^T S_b A}{A^T[(1-\alpha)S_w + \alpha I]A} \tag{7}$$

Denote cost function as $f = \boldsymbol{A}^T\boldsymbol{S_b}\boldsymbol{A}$, which subjects to $\boldsymbol{A}^T[(1-\alpha)\boldsymbol{S_w} + \alpha\boldsymbol{I}]\boldsymbol{A} = c$, where c is a constant. Rewrite the constraint function as $g = \boldsymbol{A}^T[(1-\alpha)\boldsymbol{S_w} + \alpha\boldsymbol{I}]\boldsymbol{A} - c = 0$. Using the Lagrange's multiplier λ, we have $F = f\text{-}\lambda g$. The partial derivative of F with respect to $\boldsymbol{A}$ is given as follows,

$$\frac{\partial F}{\partial A} = 2S_b A - \lambda[2(1-\alpha)S_w A + 2\alpha A] = 0 \tag{8}$$

thus $\alpha A = \frac{S_b A}{\lambda} - (1-\alpha)S_w A$.

Substituting value of $\alpha\boldsymbol{A}$ into constraint function g, then g becomes,

$$g = A^T(1-\alpha)S_w A + A^T\left[\frac{S_b A}{\lambda} - (1-\alpha)S_w A\right] - c = 0 \tag{9}$$

Further, we get $A^T S_b A = c\lambda$.

$$\hat{J}(A^*,\alpha) = \frac{A^{*T} S_b A^*}{A^{*T}[(1-\alpha)S_w + \alpha I]A^*} = \lambda_{\max} \tag{10}$$

Thereby, $\boldsymbol{A^*}$ and λ can be computed as eigen-value decomposition of $[(1-\alpha)\boldsymbol{S_w} + \alpha\boldsymbol{I}]^{-1}\boldsymbol{S_b}$.

Results

EEG rhythms comparison and dimensionality of CSP projected data by classical LDA

Preprocessing needs achieved before feature extraction, Sensory Motor Rhythms (SMR) containing Mu and Beta rhythms are generated in the sensorimotor cortex. Related information can be used by motor imagery based BCI to translate a subject's motor intention into a control signal to have efficient control over an output device.

The effect of Mu rhythm and Beta rhythm on the classification results were studied. Meanwhile, the effect dimensionality of CSP projected data can be chosen by classical LDA for each participant. Given both real and artificial data sets, 10-fold cross-validation was carried out to compute the classification performance for comparing the Mu and Beta rhythm. This 10-fold cross-validation procedure can be descripted first in here. All the available samples were partitioned into 10 equal segments randomly. Then samples in one segment are held out as validation data and the remaining 9 segments are applied as training set. Training data were used to obtain orientation matrix and classification accuracy is computed by validation data. This classification accuracy computation is carried out 10 times to evaluate average classification accuracy. The 10-fold cross-validation results of all participants for Mu rhythm (first column of Fig. 2) and Beta rhythm (second column of Fig. 2) are shown in Fig. 2.

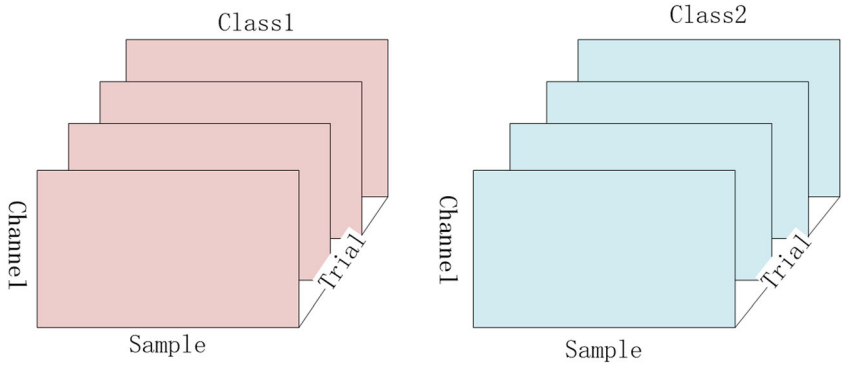

Fig. 1 EEG data saved in tensor structure

The performance of classifier can be given by Relative Operating Characteristic (ROC) curves. ROC curve is the entire set of possible true and false positive fractions attained by dichotomizing a continuous test result T with different thresholds, which is of monotone increasing trend from 0 to 1. ROC curves of different CSP-projected dimensionality (m) can be plotted in the same figure (as shown in Fig. 3), it can provide the intuitive comparing appreciation of m values, the curve with the bulge shape and near to the upper left hand corner of the figure means that this index can contribute more evaluation value. The area under ROC curve (AUC) is an important statistic, and for these areas, the larger the better.

The comparison of classification results of the two kinds of rhythm for both real and artificial participants can be given by CSP and classical LDA (as shown in Fig. 4). We found that the Mu rhythm has better classification effect than Beta rhythm for all real participants (A, B, F and G). For artificial participants, the classification accuracy evaluated by AUC values on "participants" C and E can also provide that Mu rhythm performs better than Beta rhythm. However, this cannot be found in "participant" D for any dimensionality of CSP projected data. These results show it outperforms Mu rhythm for 6 out of 7 participants by CSP and classical LDA, and "participant" D as the only one disagreement is just artificial generated. Therefore, Mu rhythm is chosen as an appropriate frequency range for EEG measurements in preprocessing part.

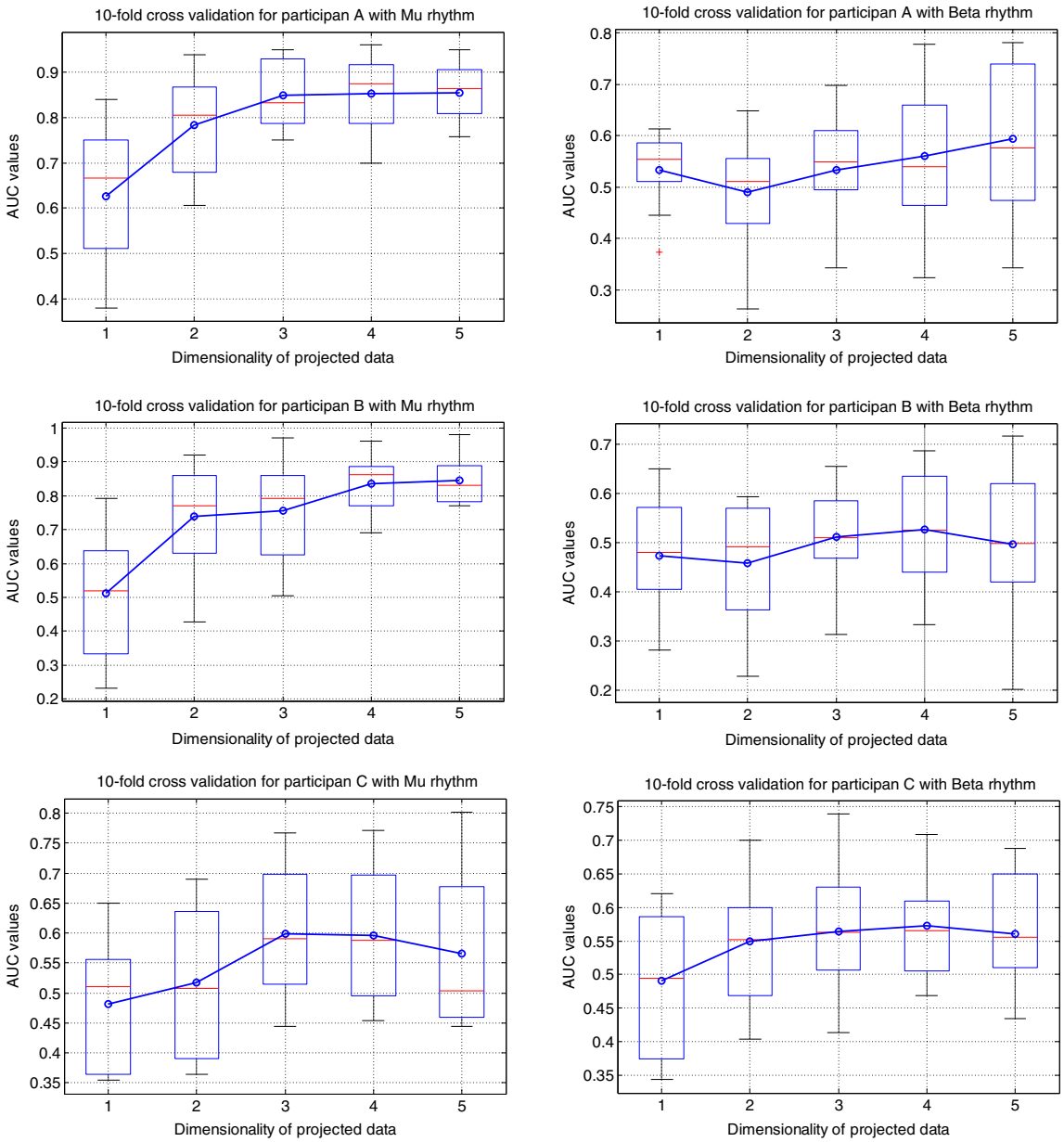

Fig. 2 10-fold cross-validation results of all participants for Mu rhythm and Beta rhythm

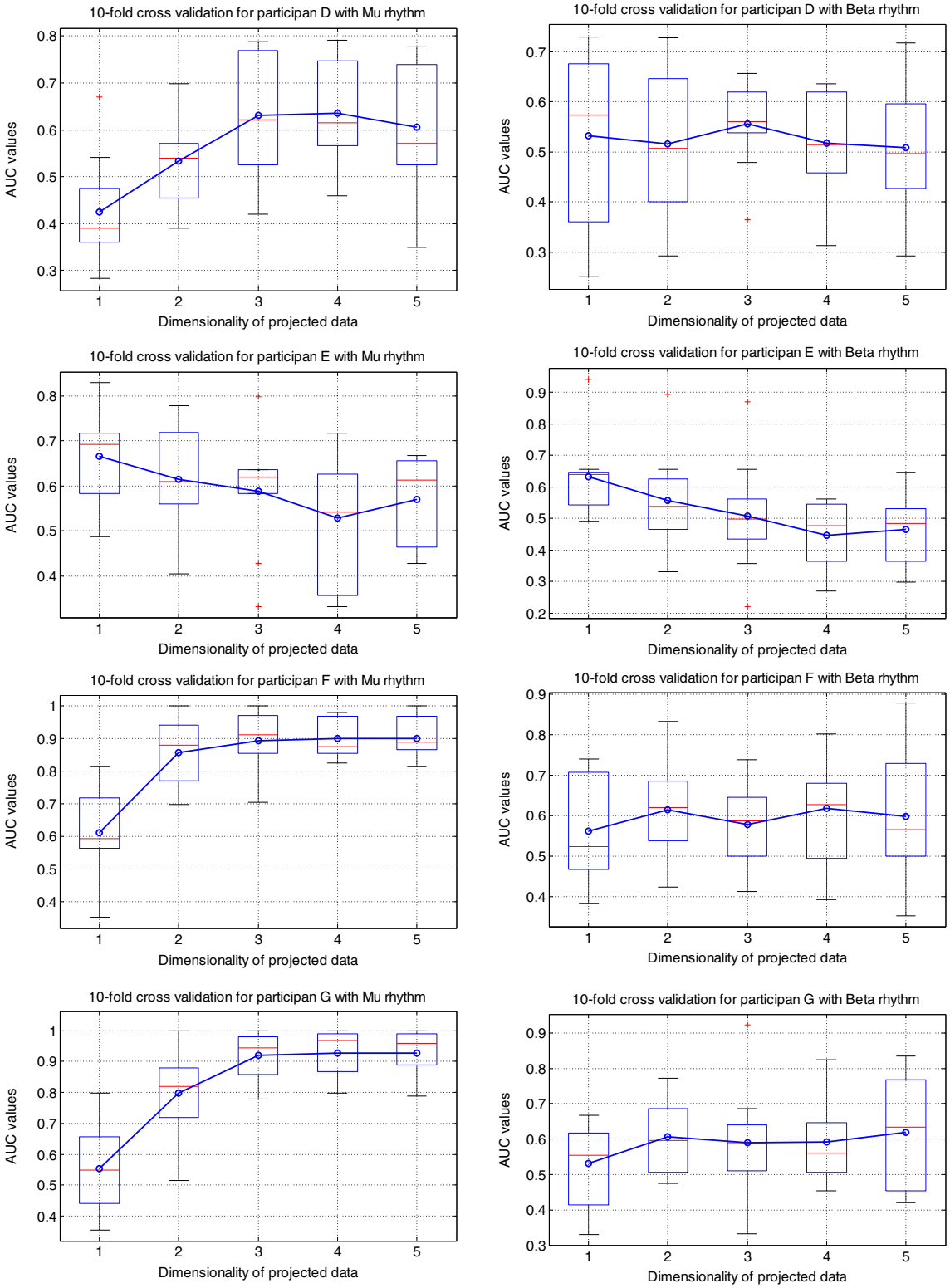

Fig. 2 (continued)

From the classification results of Mu rhythm, the best dimensionality of projected data for each participant can be obtained as shown in Table 1.

Meanwhile, the classification results of Mu and Beta rhythm by CSP and Classical LDA showed that the results of "participant" E are very different from others. Data set from

"participant" E is artificially generated data, not a real motor imagery EEG signal. As shown in Fig. 4, the classification performance of the real motor imagery EEG datasets are similar. When the projection dimension (m) is from 1 to 5, the Mu rhythm is superior to Beta rhythm in separability. And the projection dimensions are concentrated at 4, 5 to get the maximum AUC value, which gives the best separability. But for "participant" C, D, and E, since they are artificial data, the AUC values are relatively small under both Mu and Beta rhythm. And the data themselves are poorly separable. There is no common in projection dimension and the separability from dataset of "participant" C, D and E. When the projection dimension m takes the value from 1 to 5, the maximum AUC value may appear in any dimension. For dataset of "participant" C, the data separability is best when the projection dimension of Mu rhythm is 3. Meanwhile, the AUC reaches the highest value for dataset of "participant" D when the projection dimension equals 2, and the highest AUC value is given under the Beta rhythm. For dataset of "participant" E, when projection dimension equals 1, the data separability is best under μ rhythm. Therefore, the three artificially data have no commonality in both rhythm extraction and dimension selection.

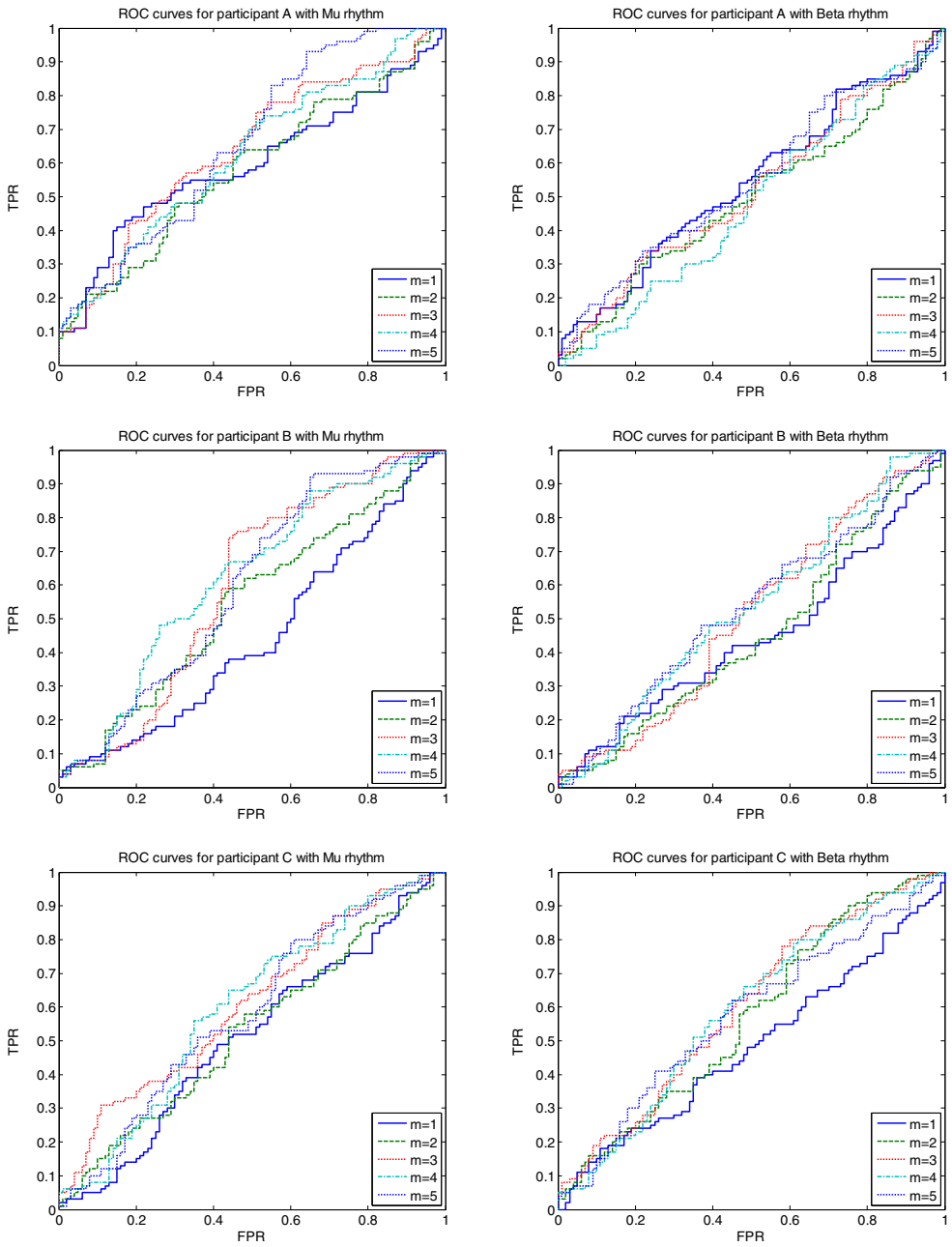

Fig. 3 ROC curves of different CSP-projected dimensionality (m)

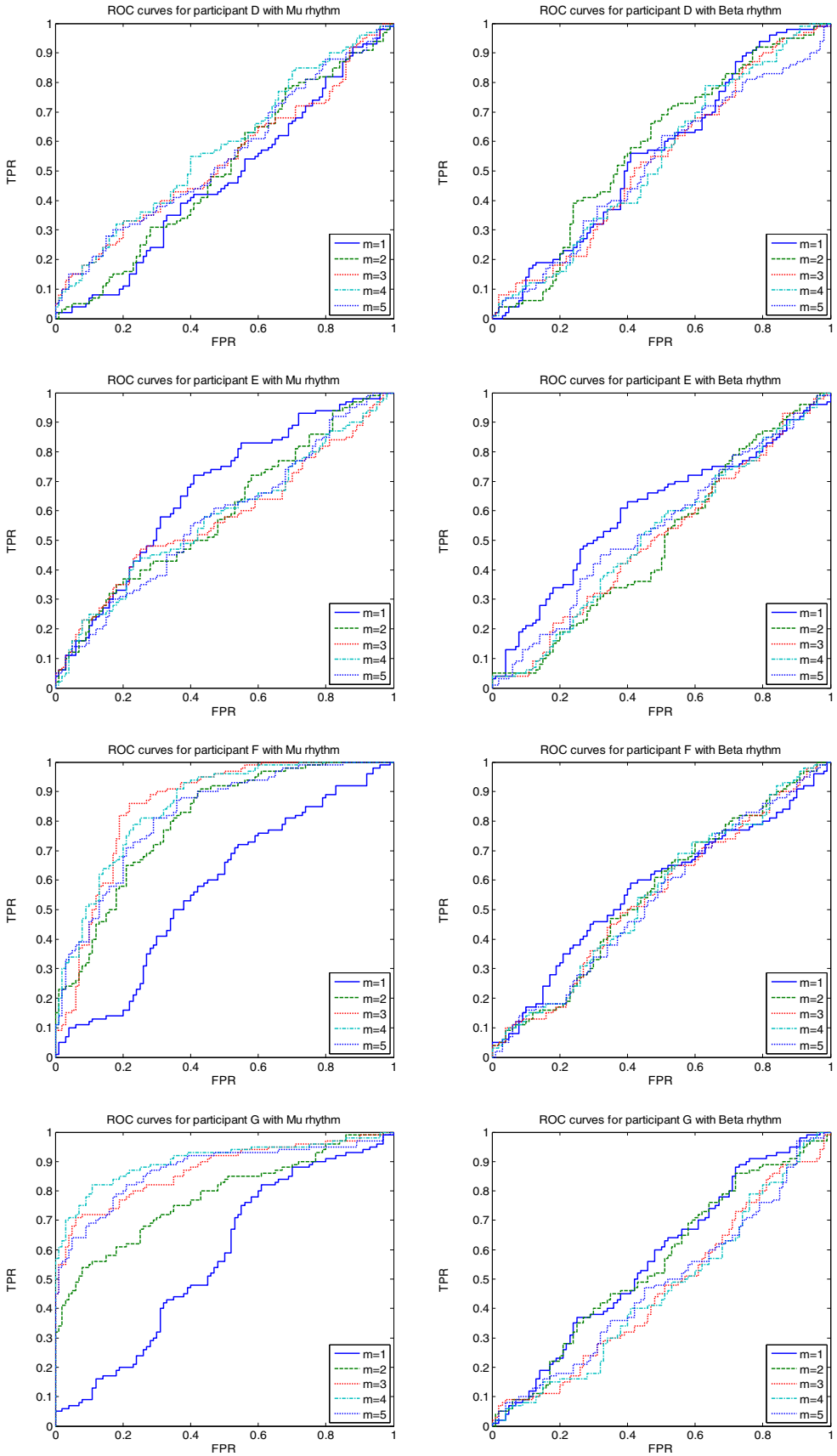

Fig. 3 (continued)

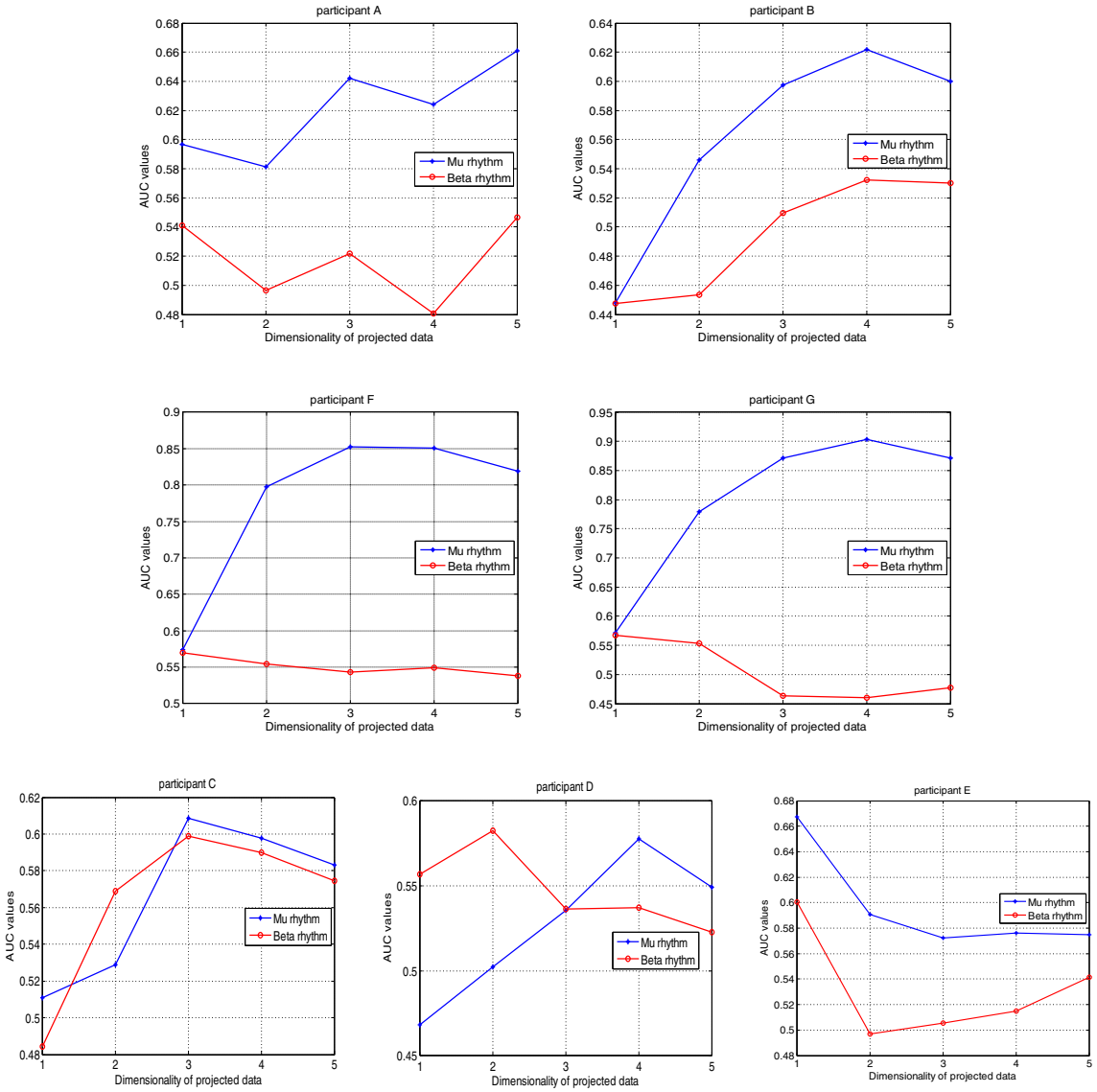

Fig. 4 Classification results of Mu and Beta rhythm for both real and artificial participants CSP and classical LDA. Real participants are A, B, F and G; Artificial participants shown in the last row in Fig. 4 are C, D and E

Classification results by CSP and regularized LDA

Given the Mu rhythms for each participant were obtain by filtering the EEG measurements, the effect dimensionality of CSP projected data was studied by regularized LDA and the consistent comparison between CSP-LDA and CSP-RLDA was concerned in classification results by CSP and regularized LDA.

In the RLDA technique, a fixed value of regularization parameter may not give the best classification performance. Therefore, cross-validation procedure was used in estimating the regularization parameter α, as mentioned in Eq. (7). Let [0,1] be the range of α value to be explored and optimal regularization parameter α be any value in this range. And the dimensionality of CSP projected data (m) also influence the classification result. The best classification performance need to find under different pairs of value α and m. Therefore, this turn out to be a two-parameter optimization problem [20], the cross validated estimate of classification accuracy can maximize jointly. The classification results of Mu rhythm for both real and artificial participants CSP and RLDA can be observed from Fig. 5.

In Fig. 5, we built a parameters plane with different pair of values of regularization parameter α and dimensionality of projected data (m). As shown in Fig. 5, α values are of the range from 0 to 1 with the interval of 0.1, and m values are with the range from 1 to 5. The classification accuracy at every point on this points' grid plane can be computed by CSP and RLDA. In this α - m grid plane, the high classification accuracy can suggest the optimal values of α and m, three-dimensional coordinate values can be showed in each sub-figure of Fig. 5.

Table 1 Classification results by CSP and Classical LDA

Participants	A	B	C	D	E	F	G
Property	real	real	artificial	artificial	artificial	real	real
Dimensionality(m)	5	4	3	4	1	3	4
Accuracy	0.6611	0.6286	0.6087	0.5778	0.6672	0.8521	0.9034

Involving the parameters, the results given by CSP and Regularized LDA can be collected in Table 2.

From the results in Tables 1 and 2, the m value determined by CSP-LDA also can give high accuracy in CSP-RLDA, which means the m value given by CSP-LDA have instruction meaning in parameter choosing for the method of CSP-RLDA. The accuracy evaluated by AUC value showed that the method of CSP-RLDA outperforms CSP-LDA method for all four real participants.

Discussion

The comparison of CSP-LDA and CSP-RLDA for all seven data sets including the four real participants and three "artificial participants" can be given in visualized way by ROC curves with AUC values as shown in Fig. 6. The results of AUC value in Fig. 6 of each participant give that much higher classification accuracies by RLDA method of participant A, B, F, and G than the results given by classical LDA method. The classification accuracy showed that the four real participants achieved performances above 0.8, especially for participant A and B, there are great increase in classification accuracy. However, for the three artificial "participants" C, D and E, the classification accuracies are very close and pretty low, this means the classification results can not get better even by the improved classification algorithm.

The same feature extraction and classification were performed for all seven participants, and the randomness of the results was avoided by cross-validation. The classification accuracy obtained by cross-validation was a reliable evaluation of the classification model. For the same feature data, if the algorithm is improved, the classification performance is greatly improved as well, it not only shows that the improved algorithm itself is more superior, but also shows that the data are separable in the feature space. Participants A, B, F, and G fall into this situation. And if the data themselves are not separable, then even a powerful classifier can not give a good classification result, which is the case for "participants" C, D and E. Therefore, the datasets from participants A, B, F, and G can be considered to be meaningful and measured by real participants when they performed different motor imagery tasks. The datasets of "participants" C, D and E are not collected from real motor imagery tasks. Datasets from "participants" C, D and E not separable and they are judged to be artificially generated data. The results obtained using this proposed method are consistent with the results published in the BCI competition. The results of the BCI competition are that A, B, F and G are real EEG data, while C, D and E were provided by Guido Nolte and Carmen Vidaurre. The details of the synthesis of the data were not announced. The great difference in classification results can be regarded as the basis of distinguishing the artificial data with the real data.

To evaluate the performance of the proposed method under the settings of BCI competition, the mean squared error (MSE) was used for performance evaluation with respect to the target vector with the values [−1,1] for two-class task recognition. The classification accuracy of all seven participants evaluated by MSE was presented in Fig. 7 for comparison purposes and completeness.

MSE values given by "participants" C, D, and E all are high by both LDA and RLDA method, as the blue shaded area in Fig. 7. However, MSE values obtained from datasets of A, B, F, and G are much smaller than that given by artificial datasets. Especially, the RLDA method gives pretty small MSE values for datasets of real participants. The MSE values of real participants are given in Table 3.

As shown in Table 3, the MSE in the classification of the data corresponding to participants A, B, F and G. The average MSE values obtained by CSP-LDA and CSP-RLDA were 0.465 and 0.195, which would have ranked the proposed method in the seventh and first of the BCI competition with the same data sets. Comparing CSP-LDA method, CSP-RLDA can give better classification performance. The potential for CSP-RLDA to improve classification performance over that of CSP-LDA depends on the situation. In situations for which the sample size are much larger than the dimension of feature space, no regularization is needed, and the regularization parameter α tends to produce a small value during the model selection procedure. However, the EEG data in this experiment is high dimensional data with small sample size. Classical LDA gets the optimal projection direction by the eigenvector of $(\boldsymbol{S}_w)^{-1}S_b$. The within-class scatter matrix $(\boldsymbol{S}_w)$ is estimated by using the pooled covariance matrix. When sample size is not considerably larger than the dimension of feature space, discriminant analysis is poorly-posed. When applying the classical LDA to such data, singularity problem may be occurred due to poor posed inverse problem in eigenvalue decomposition. Classical LDA cannot handle singular scatter matrices, which limits its applicability to low-

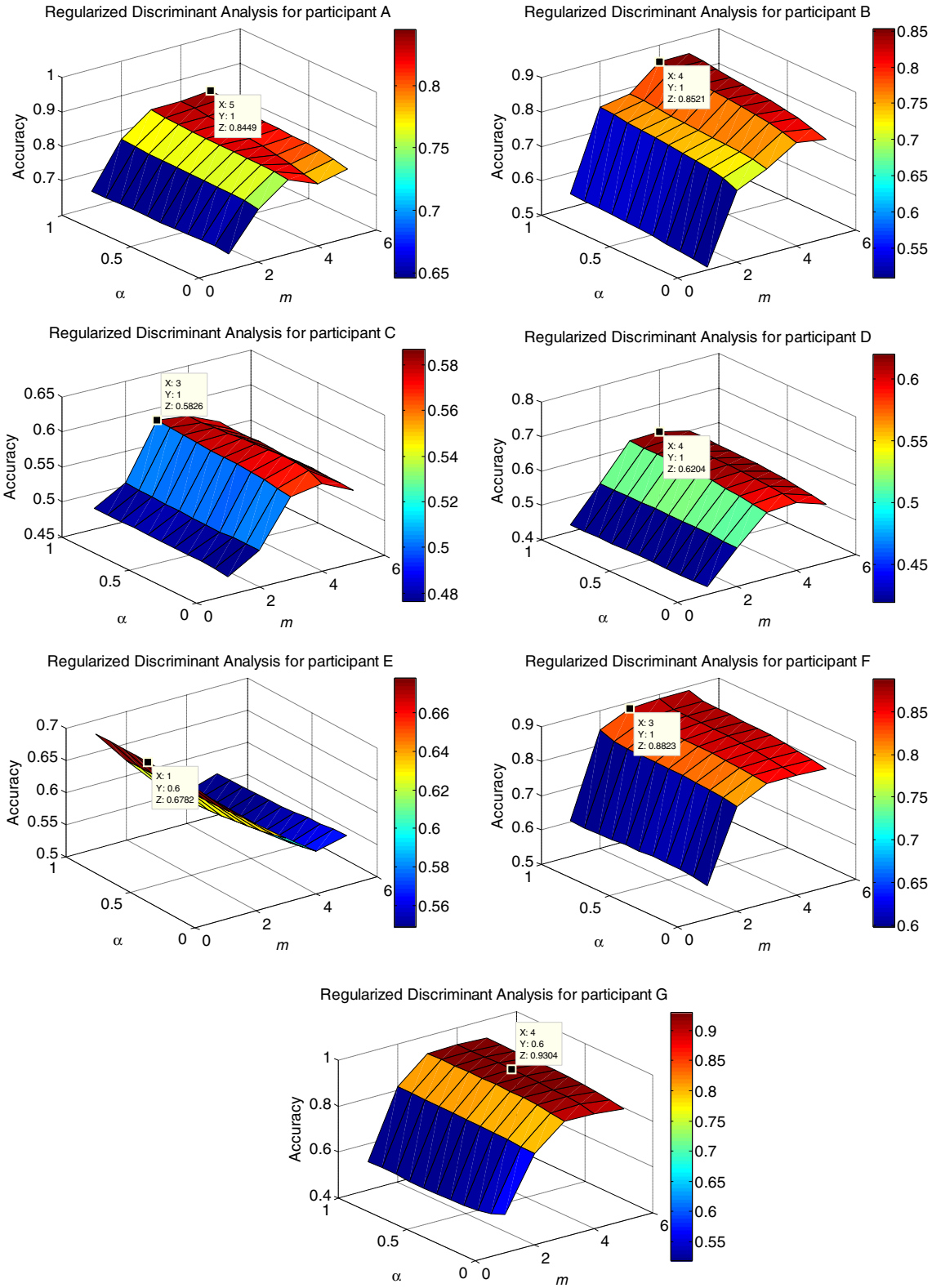

Fig. 5 Classification results of Mu rhythm for both real and artificial participants CSP and RLDA

dimensional data only. To overcome the singularity problem of the within-class scatter matrix S_w, the regularization can be done by adding a regularization parameter to the diagonal elements of S_w in the RLDA technique. Regularization technique can successfully deal with solving solution of poor posed inverse problem. RLDA combining regularization

Table 2 Classification results by CSP and Regularized LDA

Participants	A	B	C	D	E	F	G
Property	real	real	artificial	artificial	artificial	real	real
regularization parameter (α)	1	1	1	1	0.6	1	0.6
Dimensionality(m)	5	4	3	4	1	3	4
AUC	0.8449	0.8521	0.5826	0.6204	0.6782	0.8823	0.9304

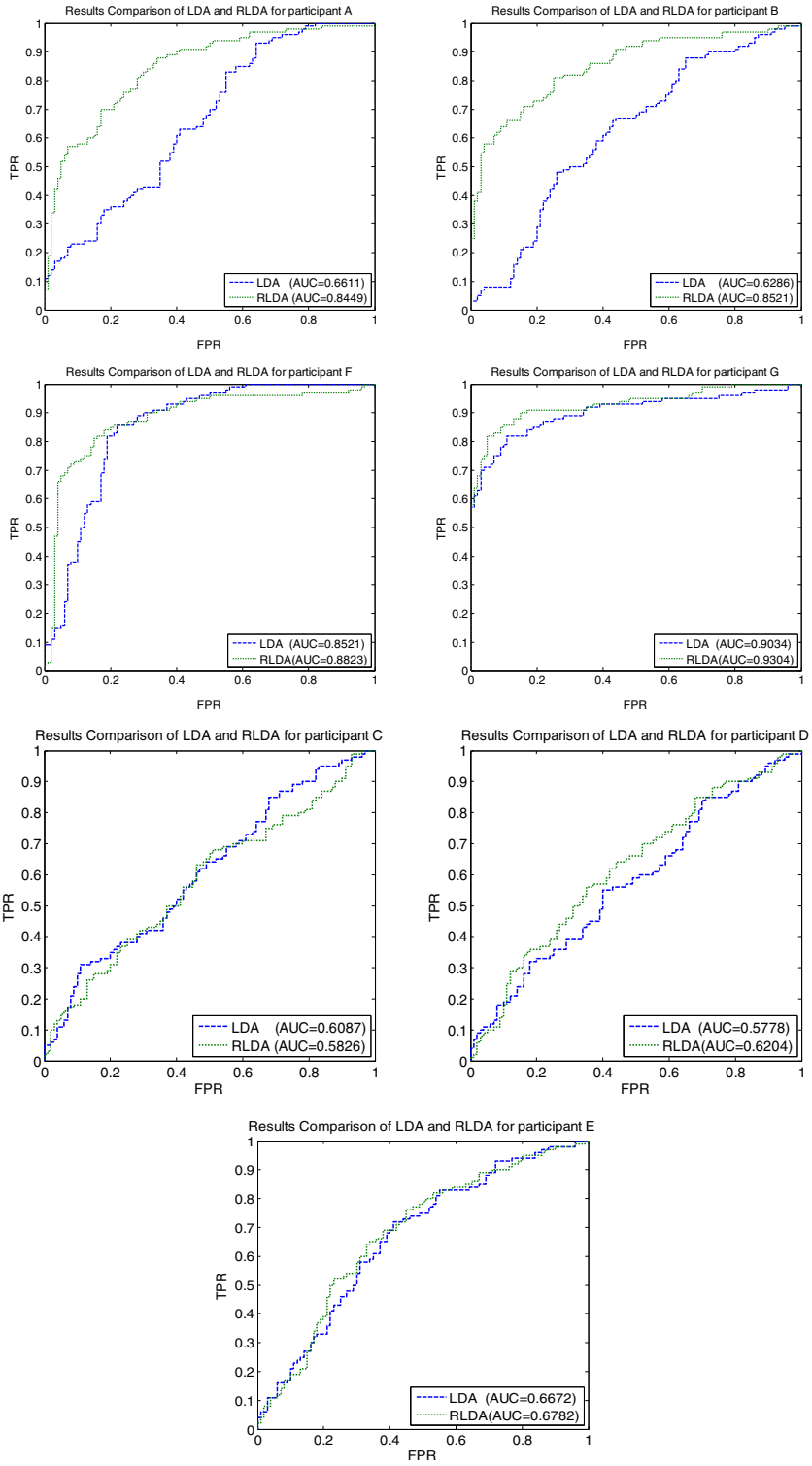

Fig. 6 Comparison of Classification Accuracy for all seven participants

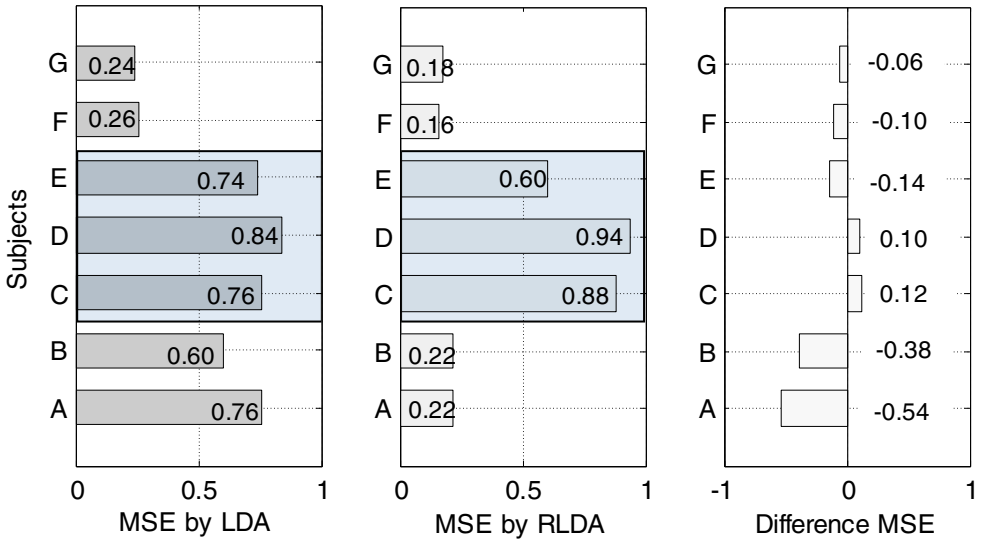

Fig. 7 Classification accuracy of all seven participants evaluated by MSE

parameter can lead to superior performance, especially in small sample setting, even small value of regularization parameter can largely improve the performance.

Conclusion

This paper presents a theoretical and computational study of common spatial pattern and regularized discriminant analysis. Given EEG measurements are filtered with a fix frequency range, Mu rhythm suggests an appropriate frequency range for BCI systems based on CSP-feature and pattern classification methods. Then, we present the dimensionality of CSP projected data by classical LDA for each participant. The criterion used in classical LDA is improved by adding regularized term to avoiding the singularity problem, based on this improved optimization criterion, regularized LDA is given in theoretically. Combined CSP method, the CSP-RLDA gives the classification accuracy for different pairs of CSP-dimensionality and regularized parameter. Experimental results on all real participants show the superiority of CSP-RLDA method over CSP-LDA method in separating motor imagery EEG measurements. The CSP-dimensionality determined by LDA method are consistent with they suggested by RLDA method. This means that though the classification accuracy obtained by LDA method are not great, in determining CSP-dimensionality, LDA method has instructing meaning in RLDA method. Based on this point, future CSP-RLDA method can be used without the computation of CSP-dimensionality.

Table 3 MSE values for real participants under BCI competition's conditions

Participants	A	B	F	G	Average
MSE by CSP-LDA	0.76	0.60	0.26	0.24	0.465
MSE by CSP-RLDA	0.22	0.22	0.16	0.18	0.195

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 51605419, 51475407), Natural Science Foundation of Hebei Province (Grant No. E2018203433), China Postdoctoral Science Foundation (Grant No. 2016 M600193), Hebei Province Funding Project for Returned Overseas Scholar (Grant No. CL201727).

Funding National Natural Science Foundation of China (Grant No. 51605419, 51475407), Natural Science Foundation of Hebei Province (Grant No. E2018203433), China Postdoctoral Science Foundation (Grant No. 2016 M600193), Hebei Province Funding Project for Returned Overseas Scholar (Grant No. CL201727).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Clarification and Statement This manuscript by Rongrong Fu, Yongsheng Tian, Tiantian Bao, Zong Meng, Peiming Shi titled "Improvement Motor Imagery EEG Classification based on Regularized Linear Discriminant Analysis" is an original unpublished work and the manuscript or any variation of it has not been submitted to any other publication previously. All of the authors have agreed with the submission.

References

1. Shi, T., Cui, W., and Ren, L., Multimedia remote interactive operations based on EEG signals constructed BCI with convolutional neural network. Multimed. Tools Appl.:1–15, 2019.
2. Burns, A., Adeli, H., and Buford, J. A., Brain-computer interface after nervous system injury. Neurosci. A Rev. J. Bringing Neurobiol. Neurol. Psych. 20(6):639–651, 2014.

3. Baxter, B. S., Edelman, B. J., Nesbitt, N., and He, B., Sensorimotor rhythm bci with simultaneous high definition-transcranial direct current stimulation alters task performance. Brain Stimul. 9(6): 834–841, 2016.
4. Michalopoulos, K., Zervakis, M., Deiber, M. P., and Bourbakis, N., Classification of eeg single trial microstates using local global graphs and discrete hidden markov models. Int. J. Neural Syst. 26(6):1650036, 2016.
5. Wang, H., Chang, W., and Zhang, C., Functional brain network and multichannel analysis for the p300-based brain computer interface system of lying detection. Expert Syst. Appl. 53:117–128, 2016.
6. Kirar, J. S., and Agrawal, R. K., Relevant feature selection from a combination of spectral-temporal and spatial features for classification of motor imagery EEG. J. Med. Syst. 42(5):78, 2018.
7. Samuel, O. W., Geng, Y., Li, X., and Li, G., Towards efficient decoding of multiple classes of motor imagery limb movements based on EEG spectral and time domain descriptors. J. Med. Syst. 41(12):194, 2017.
8. Zhang, Y., Wang, Y., Jin, J., and Wang, X., Sparse bayesian learning for obtaining sparsity of eeg frequency bands based feature vectors in motor imagery classification. Int. J. Neural Syst. 27(2):1650032, 2017.
9. Li, X., Lu, X., and Wang, H., Robust common spatial +patterns with sparsity. Biomed. Sign. Proc. Con. 26:52–57, 2016.
10. Pfurtscheller, G., Brunner, C., Schlögl, A., and Fh, L. D. S., Mu rhythm (de)synchronization and eeg single-trial classification of different motor imagery tasks. Neuroimage 31(1):153–159, 2006.
11. Ye, J., Janardan, R., Li, Q., and Park, H., Feature reduction via generalized uncorrelated linear discriminant analysis. IEEE Trans. Knowl. Data Eng. 18(10):1312–1322, 2006.
12. Hatamikia, S., and Nasrabadi, A. M., Subject transfer bci based on composite local temporal correlation common spatial pattern. Comput. Biol. Med. 64:1–11, 2015.
13. Ramoser, H., Muller-Gerking, J., and Pfurtscheller, G., Optimal spatial filtering of single trial eeg during imagined hand movement. IEEE Trans. Rehab. Eng. A Publ. IEEE Eng. Med. Biol. Soc. 8(4): 441–446, 2000.
14. Gutiérrez, D., and Salazar-Varas, R., Using eigenstructure decompositions of time-varying autoregressions in common spatial patterns-based eeg signal classification. Biomed. Sign. Proc. Con. 7(6):622–631, 2012.
15. Thomas, K. P., Guan, C., Lau, C. T., Vinod, A. P., and Kai, K. A., A new discriminative common spatial pattern method for motor imagery brain–computer interfaces. IEEE Trans. Biomed. Eng. 56(11):2730–2733, 2009.
16. Han, X., and Clemmensen, L., Regularized generalized eigendecomposition with applications to sparse supervised feature extraction and sparse discriminant analysis. Pattern Recogn. 49:43–54, 2016.
17. Blankertz, B., Dornhege, G., Krauledat, M., Müller, K. R., and Curio, G., The non-invasive berlin brain-computer interface: Fast acquisition of effective performance in untrained subjects. Neuroimage 37(2):539–550, 2007.
18. Ji, S., and Ye, J., Kernel uncorrelated and regularized discriminant analysis: A theoretical and computational study. IEEE Trans. Knowl. Data Eng. 20(10):1311–1321, 2008.
19. Friedman, J. H., Regularized discriminant analysis. J. Am. Stat. Assoc. 84(405):165–175, 1989.
20. Fu, R., Wang, H., and Zhao, W., Dynamic driver fatigue detection using hidden markov model in real driving condition. Expert Syst. Appl. 63:397–411, 2016.