Computer Methods and Programs in Biomedicine 207 (2021) 106150

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Sub-band target alignment common spatial pattern in brain-computer interface

Xianxiong Zhang[a], Qingshan She[a,*], Yun Chen[a], Wanzeng Kong[b], Congli Mei[c]

[a] School of Automation, Hangzhou DianZi University, Hangzhou 310018, China
[b] Key Laboratory of Brain Machine Collaborative Intelligence of Zhejiang Province, Hangzhou 310018, China
[c] College of Electrical Engineering, Zhejiang University of Water Resources and Electric Power, Hangzhou 310018, China

ARTICLE INFO

Article history:
Received 24 December 2020
Accepted 29 April 2021

Keywords:
Brain-computer interface
Transfer learning
Sub-band filtering
Target alignment
Cross-subject classification

ABSTRACT

Background and objective: In the brain computer interface (BCI) field, using sub-band common spatial pattern (SBCSP) and filter bank common spatial pattern (FBCSP) can improve the accuracy of classification by selection a specific frequency band. However, in the cross-subject classification, due to the individual differences between different subjects, the performance is limited.

Methods: This paper introduces the idea of transfer learning and presents the sub-band target alignment common spatial pattern (SBTACSP) method and applies it to the cross-subject classification of motor imagery (MI) EEG signals. First, the EEG signals are bandpass-filtered into multiple frequency bands (sub-band filtering). Subsequently, the source domain trails are aligned into the target domain space in each frequency band. The CSP algorithm is then employed to extract features among which more representative features are selected by the minimum redundancy maximum relevance (mRMR) approach from each sub-band. Then the features of all sub-bands are fused. Finally, conventional linear discriminant analysis (LDA) algorithm is used for MI classification.

Results: Our method is evaluated on Datasets IIa and IIb of the BCI Competition IV. Compared with six state-of-the-art algorithms, the proposed SBTACSP method performed relatively the best and achieved a mean classification accuracy of 75.15% and 66.85% in cross-subject classification of Datasets IIa and IIb respectively.

Conclusion: Therefore, the combination of sub-band filtering and transfer learning achieves superior classification performance compared to either one. The proposed algorithms will greatly promote the practical application of MI based BCIs.

1. Introduction

Brain computer interface (BCI) provides a new way for realizing information exchange from the brain to the external world by decoding the electrical signals of brain activity, which has been widely used in various fields of exoskeleton rehabilitation robot, fatigue detection, intelligent furniture, entertainment games and so on [1–4]. In most BCI systems, electroencephalogram (EEG) becomes the most popular input signal because it is non-invasive, inexpensive and relatively convenient to acquire compared with other signals such as magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI) and positron emission tomography (PET). Among them, motor imagery (MI) is one popular experimental paradigm. However, EEG signals are time-varying electrical brain activities measured from scalp with poor spatial resolution and low signal-to-noise ratio. Moreover, there is significant variability between different subjects and sessions. The characteristics of high non-stationarity and variability make it very difficult to analyze and identify EEG signals accurately. Therefore, it is of great significance to study effective methods to extract and recognize EEG features of different tasks.

Complex signal processing, feature extraction and machine learning algorithms are needed in EEG-based systems to decode different EEG signals. Common spatial pattern (CSP) [5–8] is a very popular and practical feature extraction method. Its performance depends on the subject-specific frequency band of EEG. When EEG signals are filtered using an inappropriate frequency range, the accuracy of classification using CSP features is poor. To address this

* Corresponding author.
E-mail address: qsshe@hdu.edu.cn (Q. She).

https://doi.org/10.1016/j.cmpb.2021.106150
0169-2607/

problem, Novi and Guan [9] proposed a sub-band CSP (SBCSP) algorithm. EEG signals were separated into multiple sub-bands using a bank of filters, and then the extracted features were fed into LDA to obtain scores which represent the classification ability of each sub-band, and finally the scores were fused to make decision. Ang and Guan [10] presented a filter bank CSP (FBCSP) algorithm, and designed a variety of feature selection approaches to select more representative features. Both of these two methods performed sub-band filtering methods on the original EEG data, obtaining new feature representations through fusion or feature selection, and finally achieved better experiment results. Obviously, sub-band filtering becomes an effective EEG signal processing method.

However, the performance of methods such as SBCSP and FBCSP in cross-session classification is more satisfactory than that in cross-subject classification. The reason is that the individual differences between subjects are bigger than that between different sessions of the same subject. In recent years, more and more researchers have introduced the idea of transfer learning to EEG classification. Transfer learning (TL) [11] is to apply the knowledge learned in a certain field to different but related fields or problems, aiming to solve the problem of data distribution and long calibration time, and has been successfully applied to BCI systems [12–16]. Zanini et al. [17] proposed a Riemannian alignment (RA) method to align the EEG covariance matrices of different subjects. The aligned covariance matrices can be used as a feature directly classified by a minimum distance to Riemannian mean (MDRM) [18] classifier. Yair et al. [19] proposed to view the data through the lens of covariance matrices and presented a method for domain adaptation using parallel transport on the cone manifold of symmetric positive-definite matrices. He and Wu [20] extended the RA method to Euclidean alignment (EA) method, to align EEG trials from different subjects in the Euclidean space to make them more consistent, so that any Euclidean space classifier can be used after EA. Zhang and Wu [21] presented a new manifold embedded knowledge transfer (MEKT) approach, which first aligned the covariance matrices by RA to extract tangent space features, and then performed domain adaptation by minimizing the joint probability distribution shift between the source and the target domains, to preserve their geometric information.

Since FBCSP has limited performance in cross-subject classification, and cross-subject classification is difficult, this paper proposes a sub-band target alignment common spatial pattern (SBTACSP) method, by combining the effectiveness of sub-band filtering with the advantages of cross-subject classification of transfer learning. This method includes a new alignment approach, namely target alignment (TA), which preserves the excellent properties of EA and is a variant of EA. SBTACSP decomposes the EEG signal into multiple sub-bands by band-pass filtering, and then aligns the source domain trails of each sub-band into the target domain space by TA. Subsequently, it extracts the CSP features and selects more representative features by using the minimum redundancy maximum relevance (mRMR) algorithm [22] for each sub-band. Finally, the fused features are classified by LDA [23]. The main contributions of this paper can be summarized as follows:

(1) By changing the way of domain alignment, a new target domain alignment method is proposed to make the marginal distribution of samples in source and target domains more similar, and the covariance matrix of the source domain is equal to that of the target domain. Without changing the distribution of the target domain, TA becomes a subject-directional alignment.
(2) A simple preprocessing method, sub-band filtering is added to TA, and the SBTACSP method is further proposed to obtain richer frequency information from each sub-band and improve the classification accuracy.
(3) Compared with EA-CSP-LDA and SBEACSP, the SBTACSP method obtains better classification performance. TA is more suitable for combining with sub-band filtering than EA.

The remainder of this paper is organized as follows: Section 2 presents the related work on TL, FBCSP and RA & EA. Section 3 introduces the details of the proposed SBTACSP. Section 4 describes the experimental results and compares the performance of SBTACSP with several advanced data alignment and sub-band methods. Section 5 has a discussion of other papers applying transfer learning in a BCI field. Finally, Section 6 draws a conclusion.

2. Related work

This section introduces the background of transfer learning, FBCSP and RA & EA, which will be used in the following sections.

2.1. Transfer learning

TL refers to the use of information rich source domain samples to improve the performance of the target domain model. There are two important concepts [24,25] in TL: source domain D_s represents the domain different from the test sample, but with rich supervision information; target domain D_t represents the domain of the test sample, with no label or only a few labels. The source domain and the target domain often belong to the same kind of tasks, but the distribution is different. In most BCI studies, a domain is referred to a subject or a session. For example, in EEG-based MI classification studied in this paper, a source domain is composed of EEG epochs from an existing subject, and the target domain from a new subject. When there are L source domains $\{D_s^l\}_{l=1,\ldots,L}$, we can perform TL for each of them separately, and treat the combination of L source domains as a single source domain.

2.2. Filter bank common spatial pattern

Common Spatial Pattern (CSP) [26] is an effective method for extracting features and is widely used in EEG classification. The basic idea to learn spatial filters that maximize the variance of bandpass-filtered EEG signals from one class while minimizing their variance from the other class, so that the features with the biggest difference are obtained. Hence, the CSP aims to find a spatial filtering matrix W_{csp}, which is obtained by maximizing or minimizing the following cost function

$$J(W_{csp}) = \frac{W_{csp}^T C_1 W_{csp}}{W_{csp}^T C_2 W_{csp}} \tag{1}$$

where C_l, $l = 1$, 2is the mean covariance matrix from the EEG signals belonging to class l.

The spatial filtering matrix W_{csp} of CSP is also

$$W_{csp} = U^T P, \tag{2}$$

where U is the orthogonal matrix and P is the whitening feature matrix.

FBCSP [10] is an algorithm for optimizing spatial and spectral filters. Therefore, FBCSP first uses filter banks to filter EEG signals into multiple frequency bands. The algorithm uses 4Hz band-pass filters (4-8Hz, 8-12Hz, ..., 36-40Hz), and then two pairs of spatial filters are optimized by CSP for each band-pass signal. Finally, from the 36 features (9 bands × 4CSP filters), the Mutual Information based Best Individual Feature (MIBIF) algorithm is used to select 4 related features, which are used as the input of LDA. FBCSP has been proved to be effective by winning the Fifth International BCI Competition.

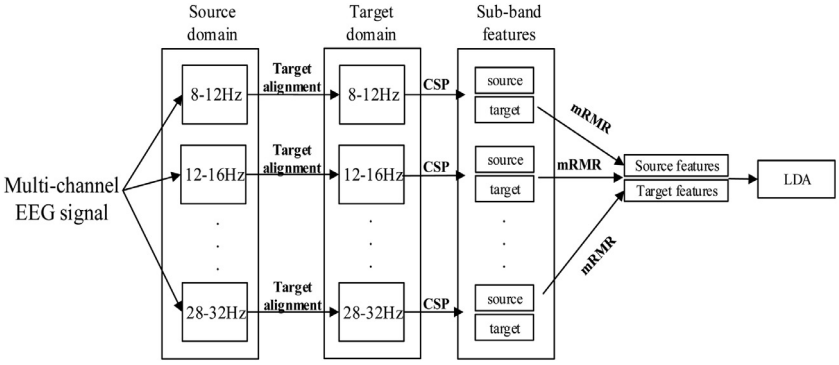

Fig. 1. Architecture of the proposed SBTACSP method.

2.3. Riemannian alignment & Euclidean alignment

RA algorithm [18] first calculates Riemannian mean of the covariance matrices from all EEG signals in resting periods as the reference matrix, in which the subject does not perform any task, to align the covariance matrices in the Riemannian manifold [27]. It reduce the variability between subjects or sessions by the following transformation:

$$\tilde{C}_i = \bar{R}^{-1/2} C_i \bar{R}^{-1/2}, \tag{3}$$

where C_i is the covariance matrix of the i-th trial, and $\tilde{C}_i$ is the corresponding aligned covariance matrix. Finally, all $\tilde{C}_i$ can be classified by MDRM classifier.

Different from RA, EA [20] aligns EEG trials instead of covariance matrices, but it also needs to calculate the reference matrix. Suppose a subject has n trails,

$$\bar{E} = \frac{1}{n}\sum_{i=1}^{n} X_i X_i^T, \tag{4}$$

where $\bar{E}$ is the Euclidean mean of all EEG trails. Then, we perform alignment by

$$\tilde{X}_i = \bar{E}^{-1/2} X_i. \tag{5}$$

After EA, CSP algorithm can be directly used to extract features for classification.

3. Method

In order to effectively select the working frequency band and features of EEG signal and reduce the data difference between subjects, this paper proposes the sub-band target alignment common spatial pattern (SBTACSP) method. The overall framework is shown in Fig. 1. It includes five stages, including sub-band filtering, target alignment, feature extraction, feature selection and classification.

3.1. Sub-band filtering

Different from EA using single frequency band of EEG signal, the SBTACSP method applies transfer learning to multiple sub-bands respectively, and combines the extracted CSP features to obtain more frequency information. Hence, SBTACSP first uses the 50-order FIR filter to filter the EEG signal in multiple sub-bands, and removes muscle artifacts and DC drift. According to the experimental results, the classification accuracy of 4-40 Hz frequency band is lower than that of 8-30 Hz in reference [21]. Therefore, different from FBCSP, the frequency range of 8-32 Hz band is selected and filtered with 4 Hz as the bandwidth into six sub-bands, namely 8-12 Hz, 12-16 Hz, ..., 28-32 Hz.

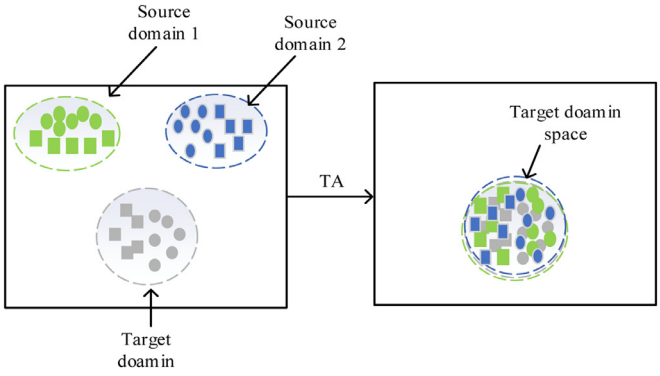

Fig. 2. Schematic illustration of the proposed TA.

3.2. Target alignment

TA is a variant of EA. Firstly, the reference matrix is calculated in the same way as Eq. (2):

$$\bar{E}_t = \frac{1}{n_t}\sum_{i=1}^{n_t} X_i^t X_i^{tT}, \tag{6}$$

where $\bar{E}_t$ is the reference matrix of the target domain, n_t is the number of samples in the target domain, X_i^t denotes the i-th EEG trail in the target domain. Then, n_s trails of source domain are aligned,

$$\tilde{X}_i^s = \bar{E}_t^{1/2} \bar{E}_s^{-1/2} X_i^s, i \in [1, n_s], \tag{7}$$

where X_i^s denotes the i-th EEG trail in the source domain. Therefore, Eqs. (6) and (7) show that TA method aims to align the source domain trails into the target domain space, so that the distribution of the source domain trails is similar to that of the target domain after alignment, as shown in the Fig. 2.

After the alignment, it is proved that the mean covariance matrix of the source domain is equal to that of the target domain:

$$\begin{aligned}
\frac{1}{n_t}\sum_{i=1}^{n_s} \tilde{X}_i^s \tilde{X}_i^{sT} &= \frac{1}{n_s}\sum_{i=1}^{n_s} \bar{E}_t^{1/2} \bar{E}_s^{-1/2} X_i^s X_i^{sT} \bar{E}_s^{-1/2} \bar{E}_t^{1/2} \\
&= \bar{E}_t^{1/2} \bar{E}_s^{-1/2} \left(\frac{1}{n_s}\sum_{i=1}^{n_s} X_i^s X_i^{sT} \right) \bar{E}_s^{-1/2} \bar{E}_t^{1/2} \\
&= \bar{E}_t^{1/2} \bar{E}_s^{-1/2} \bar{E}_s \bar{E}_s^{-1/2} \bar{E}_t^{1/2} \\
&= \bar{E}_t = \frac{1}{n_t}\sum_{i=1}^{n_t} X_i^t X_i^{tT}
\end{aligned} \tag{8}$$

TA and EA are homologous algorithms. They calculate the reference matrix in the same way, and both transform and adjust

EEG trails in Euclidean space, and do not need any label information from new subjects. Moreover, the mean covariance matrices of source and target domains are equal after EA and TA. However, there's a big differences between them. EA aligns all trails to Euclidean space, which means that the trails in the source and target domains are transformed and the mean covariance matrices are equal to the identity matrix. Therefore, all of domains are aligned to the same center and independent. But TA utilizes source domain trails and target domain trails at the same time to align source domain trails to target domain space. In this process of TA, only source domain trails are transformed, which makes the target domain trails avoid the risk of information loss by TA. Therefore, TA is a subject-directional alignment that makes the distribution of source data more consistent with that of target data.

In our method, each frequency sub-band will be domain-aligned after sub-band filtering, and the information of each sub-band is preserved as much as possible. Therefore, TA is more suitable for combining with sub-band filtering than EA.

3.3. Sub-band CSP

In this paper, we use the CSP algorithm to extract features for each sub-band after alignment using Eq. (6), and the form of the transformed signal at the k-th sub-band is

$$Z^{(k)} = \tilde{W}_{csp}^{(k)T} \tilde{X}^{(k)}, \tag{9}$$

where $W_{csp}^{(k)}$ is the CSP projection matrix of the k-th sub-band, and $\tilde{X}^{(k)} \in \mathbb{R}^{N\times T}$ denotes the single-trial EEG from the k-th sub-band after alignment. In Eq. (9), the spatial filtered signal $Z^{(k)}$ uses $W_{csp}^{(k)}$ to maximize the differences in the variance of the 2 classes of band-pass filtered EEG. The m pairs of CSP features of the i-th trail for the EEG from the k-th sub-band are then given by

$$v_i^{(k)} = \log\left(\frac{diag(\tilde{W}_{csp}^{(k)^T}\tilde{X}_i^{(k)}, \tilde{X}_i^{(k)^T}\tilde{W}_{csp}^{(k)})}{trace\left[\tilde{W}_{csp}^{(k)^T}\tilde{X}_i^{(k)}, \tilde{X}_i^{(k)^T}\tilde{W}_{csp}^{(k)}\right]}\right), \tag{10}$$

where $v_i^{(k)} \in \mathbb{R}^{1\times 2m}$; $\tilde{W}_{csp}^{(k)}$ represents the first m and the last m columns of $W_{csp}^{(k)}$; $diag(\bullet)$ denotes the diagonal elements of the square matrix.

The feature vectors of the 6 sub-bands are$v^{(1)}, v^{(2)}, v^{(3)}, v^{(4)}, v^{(5)}, v^{(6)} \in \mathbb{R}^{n\times 2m}$.

3.4. Feature selection and classification

According to the feature extraction in the previous section, we can get the six features vectors $v^{(k)} = [f_1, f_2, ..., f_{2m}], k \in [1, 6]$and true labels of the training data.

If the features of the six sub-bands are fused together, the dimension of the final feature will be particularly high, which will affect the performance of the classifier and increase the redundancy between features. In order to decrease the feature dimension, reduce the difficulty of learning task and improve the efficiency of the model, this paper uses the minimum Redundancy Maximum Relevance (mRMR) algorithm to select more representative CSP features from each sub-band.

In the mRMR method, the mutual information is used to measure the relevance of the features to generate a ranking, in which the features are sorted according to their mutual information with classes and other features. The most relevant feature is the one with the maximum mutual information with the class and with the minimum mutual information with the other features [22]. This is achieved by maximizing the following expression:

$$r_i = \frac{\frac{1}{n_f}\sum I(Y_s, f_i^T)}{\frac{1}{n_f^2}\sum I(f_s, f_i)}, \tag{11}$$

where $n_f = 2m$, $I(\bullet)$ is the mutual information. By the mRMR, we can filter out p features that represent the information of corresponding sub-band, and finally obtain a new feature matrix $\tilde{F} \in \mathbb{R}^{n\times 6p}$, where n is the number of trails and $6p$ is the dimension of a feature vector.

The mRMR feature selection algorithm requires a user-defined parameter p to select a certain number of features. After fusing the features of all sub-bands, this paper uses the LDA algorithm to design classifiers and compares the classification performance with the other submitted methods.

4. Experiments

In this section, in order to verify the effectiveness of the proposed method, we evaluate it in two scenarios, including single-source to single-target (STS) transfer and multi-sources to single-target (MTS) transfer.

4.1. BCI datasets

Two publicly available BCI Competition datasets are used to evaluate the performance of SBTACSP compared to the state-of-the-art methods MDRM, RA-MDRM, CSP-LDA, EA-CSP-LDA, FBCSP and SBEACSP. These datasets are widely used for evaluating MI EEG classification algorithms and described shortly.

(1) BCI Competition IV Dataset IIa: The EEG signals of 9 healthy subject (A01-A09) were collected in this dataset. Each subject was asked to imagine four kinds of MI tasks of left hand, right hand, foot and tongue after the cue appearance. Each experiment lasted for 4 seconds. According to the international 10/20 system, 22-channel EEGs were recorded at 250Hz. There were 288 trails for every subject. In the study, only two classes (left-hand and right-hand) of EEG data are employed, and each class has 72 trails.

(2) BCI Competition IV Dataset IIb: The EEG signals of 9 healthy subject (B01-B09) were collected in this dataset. Each subject was asked to imagine only two kinds of MI tasks of left hand and right hand after the cue appearance. Each experiment lasted for 4 seconds. There were 120 trails for every subject. According to the international 10/20 system, 3-channel EEGs were recorded at 250Hz. In the cross-subject transfer, we only used the EEG of the training session, including complete label information. Next, we extracted EEG signals between [0.5, 3.5] seconds after the cue appearance as our trails for both datasets [28].

4.2. Parameter settings and cross-validation results

For Dataset IIa, the choice of m pairs of CSP features is set to 3 because the greater choice of m does not significantly improve the classification accuracy [29,30]. For Dataset IIb, the choice of m pairs of CSP features is set to 1. These are only 3 channels of EEG signals available, thus $W_{csp}^{(k)} \in \mathbb{R}^{3\times 3}$in Eq. (2) limited the maximum selection of $m = 1$ for $\tilde{W}_{csp}^{(k)}$ [31].

Considering that each frequency band has several features, if they are directly spliced together to form the final feature, it will lead to a series of performance problems due to too high feature dimension. Therefore, the parameter p is determined as 2, which means that each sub-band will have two features selected by mRMR, and fused to form the final feature of 12 dimensions.

Next, to evaluate our method we randomly divide the dataset into 10 subsets, where each subset has similar proportion of samples from each class. One subset is selected orderly as testing data, and the rest subsets are used as training data. This process is repeated 10 times, and we can obtain the accuracies of 10 times. We

Table 1
Classification accuracy (%) of 10-fold cross-validation with RA-MDRM, EA-CSP-LDA, FBCSP and SBTACSP of BCI Competition IV Datasets IIa and IIb.

Subject	Dataset IIa				Dataset IIb			
	RA-MDRM	EA-CSP-LDA	FBCSP	SBTACSP	RA-MDRM	EA-CSP-LDA	FBCSP	SBTACSP
1	87.67	89.67	**91.71**	87.62	73.33	72.50	**84.17**	78.33
2	**63.19**	60.48	59.00	60.86	59.17	60.83	**63.33**	58.83
3	96.43	96.47	**95.76**	94.43	47.50	53.33	51.67	**61.67**
4	63.95	70.95	**77.24**	76.05	85.00	86.67	86.67	**86.67**
5	61.24	52.57	87.47	**87.52**	60.00	56.67	76.67	**80.00**
6	68.10	67.90	63.14	**68.62**	57.50	57.50	70.83	**71.67**
7	75.71	77.67	88.10	**86.81**	54.17	51.67	**62.50**	59.17
8	95.76	96.43	97.19	**97.42**	59.17	57.50	59.38	**61.67**
9	81.33	84.14	85.67	**87.57**	65.83	65.83	**67.50**	65.00
Ave	77.04	77.37	82.81	**82.99**	62.41	62.50	69.11	**69.17**

Table 2
Classification accuracy (%) of BCI Competition IV Dataset IIa in STS and MTS transfers.

Subject		MDRM	RA-MDRM	CSP-LDA	EA-CSP-LDA	FBCSP	SBEACSP	SBTACSP
S	A01	60.23	68.66	61.81	71.01	51.65	73.35	**81.68**
T	A02	50.95	54.43	52.43	55.21	**63.11**	56.25	55.47
S	A03	62.24	76.48	62.24	77.26	60.52	**84.38**	82.38
	A04	56.68	65.71	54.77	63.19	53.30	64.06	**66.15**
	A05	50.35	58.07	50.17	53.04	54.25	**60.50**	56.42
	A06	51.39	62.50	54.25	60.59	54.43	64.00	**66.93**
	A07	55.21	62.50	53.99	59.11	65.28	65.97	**66.84**
	A08	52.60	**80.47**	73.44	77.86	68.77	71.09	75.26
	A09	57.11	**72.83**	60.85	67.19	59.81	65.28	67.71
	Avg	55.20	67.16	58.22	64.94	59.01	67.25	**68.76**
M	A01	63.19	73.61	76.39	86.81	72.53	85.42	**90.28**
T	A02	49.31	56.94	54.86	56.25	**60.56**	59.03	58.33
S	A03	62.50	84.72	84.72	**98.61**	76.39	95.14	97.92
	A04	67.36	68.75	73.61	72.92	65.83	72.92	**75.69**
	A05	50.00	63.19	47.92	51.39	56.25	**64.58**	58.33
	A06	50.00	68.06	54.17	66.67	53.31	68.75	**73.61**
	A07	54.17	70.83	65.28	69.44	**76.39**	73.61	72.22
	A08	59.72	86.11	72.92	**88.89**	68.06	82.64	83.33
	A09	56.94	**79.86**	78.47	72.92	80.03	61.11	66.67
	Avg	57.02	72.45	67.59	73.77	67.71	73.69	**75.15**

employ the accuracy average to measure the performance. The results of 10-fold cross-validation between SBTACSP and RA-MDRM, EA-CSP-LDA and FBCSP on the two datasets are compared, and the results are shown in Table 1. It can be seen that in Table 1, when we use the fusion feature, we obtain the average accuracy of 82.99 in Data IIa and 69.17 in Data IIb. Although the accuracy of SBTACSP is only a little higher than FBCSP, it is much higher than EA-CSP-LDA. It shows that TA does not reduce the effect of sub-band filtering, and the combination of TA and sub-band filtering is better than the single TL method.

4.3. Cross-subject transfer

In order to verify the effectiveness of the proposed method, we compared the SBTACSP method with six kinds of BCI classification algorithms. The six algorithms can be divided into three categories:

(1) Covariance-based classification algorithms: MDRM [18] and RA-MDRM [17].
(2) Vector-based classification algorithms: CSP-LDA [32] and EA-CSP-LDA [20].
(3) Sub-bands-based classification algorithms: FBCSP [10] and SBEACSP.

Here, RA-MDRM and EA-CSP-LDA are TL algorithms, and the SBEACSP is obtained by the EA approach while the TA algorithm is used in the SBTACSP. In the following, we evaluated six algorithms in STS and MTS transfers respectively.

Tables 2 and 3 showed the classification accuracy of BCI Competition IV Datasets IIa and IIb in STS and MTS transfers, respectively. Generally speaking, SBTACSP was superior to the other six algorithms in average classification accuracy. Specifically, for Dataset IIa in STS transfer, our method achieved the best average accuracy on subjects A01, A04, A06 and A07, while SBEACSP gained the best results on A03 and A05, and FBCSP performed best on A02, and RA-MDRM achieved the best results on A08 and A09. In MTS transfer, our method achieved the best average accuracy on subjects A01, A04 and A06, while SBEACSP gained the best results on A05, and FBCSP performed best on A02 and A07, and EA-CSP-LDA gained the best results on A03 and A08, and RA-MDRM achieved the best results only on A09. For Dataset IIb in STS transfer, our method performed best on subjects B03, B04, B06, B07 and B09, while SBEACSP obtained the best average accuracy only on B02, and EA-CSP-LDA just performed best on B05, and RA-MDRM achieved the best results on B01 and B08. In MTS transfer, our method performed best on subjects B02, B03, B04, B06, B07 and B09, while FBCSP achieved the best results only on B08, and EA-CSP-LDA performed best on B01 and B05. Taking MTS transfer for example, the classification accuracy by our method was 90.28% on subject A01, better than EA-CSP-LDA (86.81%) and RA-MDRM (73.61%). For all 9 subjects, our method yielded the highest accuracy (75.15%), a 7.44 improvement over FBCSP, a 2.7% improvement over RA-MDRM, a

Table 3
Classification accuracy (%) of BCI Competition IV Dataset IIb in STS and MTS transfers.

Subject		MDRM	RA-MDRM	CSP-LDA	EA-CSP-LDA	FBCSP	SBEACSP	SBTACSP
S	B01	55.52	**67.39**	62.39	63.31	55.21	63.44	66.15
T	B02	50.83	51.77	52.19	50.31	51.56	**58.23**	57.08
S	B03	51.98	53.02	50.52	50.83	55.42	51.35	**58.44**
	B04	57.60	74.27	60.10	73.75	53.38	72.29	**74.79**
	B05	52.60	56.77	54.38	**58.33**	56.98	50.42	55.00
	B06	51.15	54.27	53.96	54.58	53.75	64.48	**65.00**
	B07	51.67	54.78	53.44	54.69	55.94	57.08	**57.08**
	B08	51.77	**57.19**	53.75	55.21	56.25	54.06	54.48
	B09	56.35	57.40	60.42	61.98	61.67	61.77	**65.21**
	Avg	53.28	58.54	55.68	58.33	55.57	59.24	**61.47**
M	B01	75.00	74.17	73.33	**75.83**	69.17	75	72.50
T	B02	57.50	51.67	54.17	55.83	57.50	62.50	**63.33**
S	B03	55.83	58.33	50.00	50.00	58.33	58.17	**60.00**
	B04	60.83	83.33	60.00	77.50	69.17	82.50	**85.83**
	B05	54.17	61.67	52.50	**62.50**	49.17	56.33	58.33
	B06	56.67	57.50	51.67	62.50	60.00	68.50	**72.50**
	B07	54.17	55.00	63.33	57.50	54.17	59.50	**63.33**
	B08	55.00	59.17	55.00	55.83	**65.83**	59.17	55.83
	B09	59.17	60.83	63.33	67.50	68.02	70.00	**70.00**
	Avg	58.70	62.41	58.15	62.78	61.26	65.74	**66.85**

1.46% improvement over SBEACSP, and a 1.38% improvement over EA-CSP-LDA. Comparing the STS and MTS transfers, the accuracy of the MTS transfer was much higher than that of the STS transfer, which showed that the MTS transfer can appropriately improve the accuracy.

In our cross-subject scenario, FBCSP did not achieve good results. The effect of FBCSP in cross-session transfer [33] was better than that in cross-subject transfer, the reason was that the data of the training and testing set of the cross-session transfer belonged to the same subject, and the difference between the training and testing set was far less than that of cross-subject transfer. In addition to the STS transfer of Dataset IIb, the accuracy of FBCSP was higher than that of CSP-LDA, which indicated that the sub-band filtering had achieved a positive effect. Meanwhile, comparing EA-CSP-LDA and SBTACSP, the accuracy of SBTACSP was at least 1.38% higher than that of EA-CSP-LDA, and the highest could be 4.07% higher. This also showed that it was necessary to perform sub-band filtering before alignment. Compared to FBCSP, SBTACSP achieved relatively better performance, which indicates that TA could effectively solve the problem of data difference between subjects. Additionally, EA-CSP-LDA had superiority over FBCSP, which showed that TL had great advantages in cross-subject transfer. Based on the above three points, incorporating sub-band filtering with TA is feasible. Finally, SBTACSP performed better than SBEACSP, which showed that our proposed variant TA was more suitable for combining it with sub-band filtering. In addition, compared with the existing TL algorithms (RA-MDRM and EA-CSP-LDA), the features extracted by the proposed SBTACSP are more representative and more conducive to classification, which makes the classification accuracy higher on two BCI datasets.

Additionally, principal component analysis (PCA) is employed to dimensionality reduction. The classification accuracies are obtained and showed in Table 4. As we can see from Table 4, although the fusion feature can improve the expression of EEG features to some extent, there are still some redundant information among these twelve features. Therefore, when we do the PCA transform, the classification accuracy is expected to be improved. In the STS scenario, when the dimension of feature space is 4, we can get the highest average classification accuracy with 69.81% for Data IIa; when the dimension of feature space is 3, the highest average classification accuracy is 64.13% for Data IIb. In the MTS scenario, when the dimension of feature space is 9, we can get the highest average classification accuracy with 75.46% for Data IIa; when the dimension of feature space is 8, the highest average classification accuracy is 67.59% for Data IIb. We can find that when the highest accuracy is achieved, the dimension of feature space is less than 12. There is redundant information between sub-bands.

Table 4
Classification accuracy of different dimensions of feature vector.

Dimension	Dataset IIa		Dataset IIb	
	STS 67.24	MTS	STS	MTS 62.78
1	67.24	70.29	61.82	62.78
2	67. 86	71.22	62.29	65.56
3	68.43	69.75	**64.13**	66.67
4	**69.81**	71.76	63.92	66.57
5	69.65	72.76	63.54	65.19
6	69.40	73.14	63.20	67.41
7	68.80	73.84	62.62	67.41
8	68.30	75.30	62.53	**67.59**
9	68.25	**75.46**	62.45	67.41
10	68.64	74.92	61.98	67.22
11	68.57	75.15	61.53	66.57
12	68.76	75.15	61.47	66.85

4.4. Feature visualization

In order to verify the distribution effect of features extracted by SBTACSP, this paper uses a non-linear dimensionality reduction technique, t-stochastic neighbor embedding (t-SNE) [34] to display the difference in feature distribution among the CSP features and the features by SBTACSP.

Taking the MTS transfer as an example, Figs. 3 and 4 show the feature distributions from two representative subjects for Dataset IIa and IIb respectively, each row corresponding to a different test subject. The figures on the left column show the features extracted directly by CSP, and the figures on the right column show that the features extracted by SBTACSP. In each subplot, red * and blue * represent features of class 1 from source domain and target domain respectively, and green circles and black circles represent features of class 2 from source domain and target domain respectively. After SBTACSP feature extraction, it can be seen that the distribution of each class of features in the source domain is closer to that of the target domain. In other words, the red and blue * are clustered together, and the green and black circles are clustered together to achieve the purpose of feature transfer. Therefore, the visualization effect is consistent with the classification result.

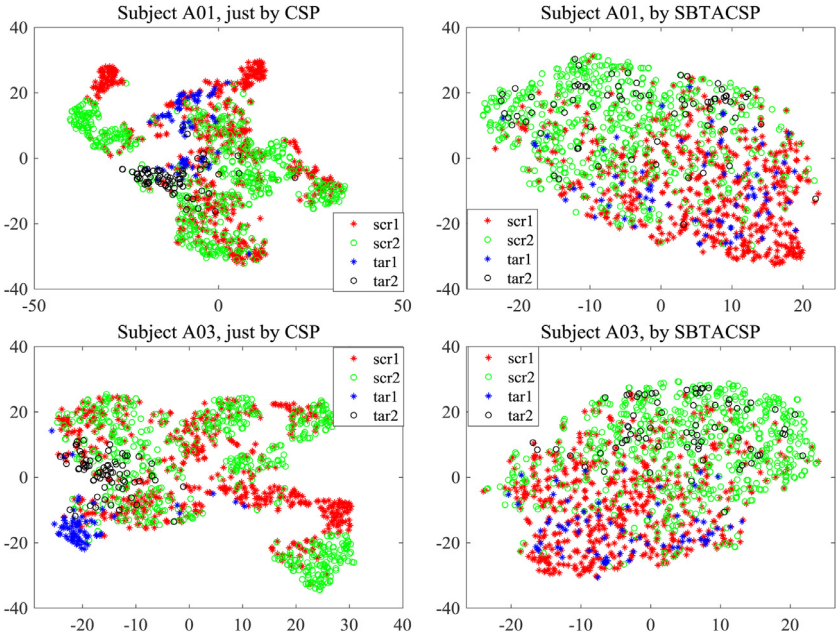

Fig. 3. Feature visualization of CSP and SBTACSP of two subjects from Dataset IIa in MTS transfer.

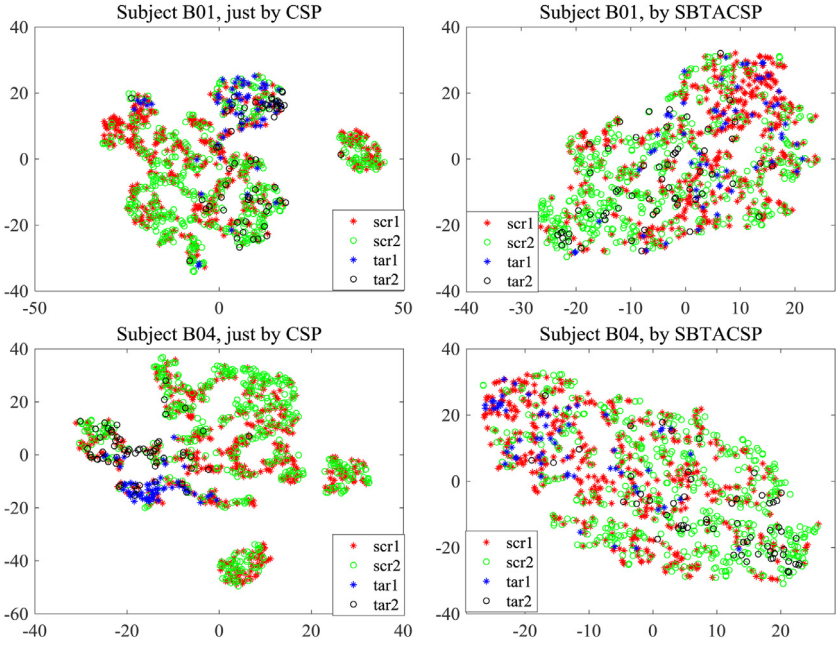

Fig. 4. Feature visualization of CSP and SBTACSP of two subjects from Dataset IIb in MTS transfer.

5. Discussion

In these experiments, the proposed SBTACSP method exhibited an excellent performance in classification, as demonstrated in two MI EEG datasets. As shown by Tables 2 and 3 in our method, the mean accuracy of MTS transfer is at least 6.39% higher than that of STS transfer in Dataset IIa, and the mean accuracy of MTS transfer is at least 5.38% higher than that of STS transfer in Data IIb. It is shown that more samples of subjects are added to the training, which can effectively improve the accuracy of classification. However, there is a case in which STS classification is better than MTS classification, such as subject A09.

Finally, we compare our algorithm to other existing methods applying TL in a BCI field: EA-CSP-wAR [24] and TS-TL [25].The results are presented in Table 5. The average accuracies on both datasets obtained by SBTACSP after PCA are higher than these of EA-CSP-wAR, while the result for all nine subjects is comparable with the result reported by TS-TL.

Table 5
Comparison between our algorithm and other methods on BCI Competition IV Dataset IIa and IIb in MTS transfer.

	EA-CSP-wAR	TS-TL	SBTACSP
Dataset IIa	73.91	75.52	75.46
Dataset IIb	66.10	-	67.59

6. Conclusion

In this paper, a new SBTACSP transfer learning method is proposed by combining the advantages of sub-band filtering and domain alignment. We demonstrated the effectiveness of our method for identifying 2-class MI tasks on BCI Competition IV Dataset IIa and IIb. Different from EA-CSP-LDA, the presented method uses sub-band filtering to obtain more frequency band information and better classification features. On the other hand, compared with FBCSP, SBTACSP uses unsupervised TA to improve classification performance based on transfer learning theory. Experimental results show that compared with MDRM, RA-MDRM, CSP, EA-CSP-LDA, FBCSP and SBEACSP, SBTACSP can effectively learn better feature representations and obtain excellent classification performance in two BCI MI-EEG datasets. In general, MTS transfer is better than STS transfer, but not all subjects' classification results follow this conclusion. Therefore, it is feasible to use domain transferability estimation method to select the most beneficial subjects in multi-source transfer learning to solve this problem. Additionally, TA ignores the difference of conditional distribution which represents the distribution of each class. Hence, in the future, we can utilize the knowledge of conditional distribution from some advanced TL algorithms, such as joint distribution adaptation (JDA) [35] and joint geometrical and statistical alignment (JGSA) [36] to improve the classification and evaluation of MI-EEG data.

Declaration of Competing Interest

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

Acknowledgments

This work was supported by National Natural Science Foundation of China under grant No. 61871427,Key Research & Development Project of Zhejiang Province under grant No. 2020C04009 and Key R & D projects of Shandong Province under grant No. 2019JZZY021005.

References

[1] F. Lotte, L. Bougrain, A. Cichocki, M. Congedo, et al., A review of classification algorithms for EEG-based brain-computer interfaces: a 10-year update, J. Neural Eng. 15 (3) (2018) 1741–1769.
[2] C.G. Coogan, B. He, Brain-computer interface control in a virtual reality environment and applications for the internet of things, IEEE Access 6 (2018) 10840–10849.
[3] R.P. Balandong, et al., A review on EEG-based automatic sleepiness detection systems for driver, IEEE Access 6 (2018) 22908–22919.
[4] V. Jayaram, et al., Transfer learning in brain-computer interfaces, IEEE Comput. Intell. Mag. 11 (1) (2015) 20–31.
[5] M. Dai, D. Zheng, S. Liu, P. Zhang, Transfer kernel common spatial patterns for motor imagery brain-computer interface classification, Comput. Math. Methods Med. 2018 (2018) 1–9.
[6] Z.J. Koles, et al., Spatial patterns underlying population differences in the background EEG, Brain Topogr. 2 (4) (1990) 275–284.
[7] J. Müller-Gerking, et al., Designing optimal spatial filters for single trial EEG classification in a movement task, Clin. Neurophysiol. 110 (5) (1999) 787–798.
[8] H. Kang, et al., Composite common spatial pattern for subject-to-subject transfer, Signal Process. Lett. 16 (8) (2009) 683–686.
[9] Q. Novi, C.T. Guan, T.H. Dat, P. Xue, Sub-band common spatial pattern (SBCSP) for brain-computer interface, in: 2007 3rd International IEEE/EMBS Conference on Neural Engineering, 2007, pp. 204–207.
[10] K.K. Ang, Z.Y. Chin, H. Zhang, C. Guan, Filter bank common spatial pattern (FBCSP) in brain-computer interface, in: Proceedings of the IEEE International Joint Conference on Neural Networks, 2008, pp. 2391–2398.
[11] H. He, D. Wu, Different set domain adaptation for brain-computer interfaces: a label alignment approach, IEEE Trans. Neural Syst. Rehabil. Eng. 28 (5) (2020) 1091–1108.
[12] H. Ibrahim, K. Abbas, H. Imali, et al., Multiclass informative instance transfer learning framework for motor imagery-based brain-computer interface, Comput. Intell. Neurosci. 2 (22) (2018) 1–12.
[13] P.L.C. Rodrigues, C. Jutten, M. Congedo, Riemannian procrustes analysis: transfer learning for brain-computer interfaces, IEEE Trans. Biomed. Eng. 66 (8) (2019) 2390–2401.
[14] D. Wu, Online and offline domain adaptation for reducing BCI calibration effort, IEEE Trans. Hum.-Mach. Syst. 47 (4) (2017) 550–563.
[15] D. Wu, et al., Driver drowsiness estimation from EEG signals using online weighted adaptation regularization for regression (OwARR), IEEE Trans. Fuzzy Syst. 25 (6) (2017) 1522–1535.
[16] D. Wu, et al., Switching EEG headsets made easy: reducing offline calibration effort using active wighted adaptation regularization, IEEE Trans. Neural Syst. Rehabil. Eng. 24 (11) (2016) 1125–1137.
[17] P. Zanini, M. Congedo, et al., Transfer learning: a Riemannian geometry framework with applications to brain-computer interfaces, IEEE Trans. Biomed. Eng. 65 (5) (2018) 1107–1116.
[18] F. Yger, et al., Riemannian approaches in brain-computer interfaces: a review, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (10) (2017) 1753–1762.
[19] O.Y air, M. Ben-Chen, R. Talmon, Parallel transport on the cone manifold of SPD matrices for domain adaptation, IEEE Trans. Signal Process. 67 (7) (2019) 1797–1811.
[20] H. He, D. Wu, Transfer learning for brain-computer interfaces: a Euclidean space data alignment approach, IEEE Trans. Biomed. Eng. 67 (2) (2020) 399–410.
[21] W. Zhang, D. Wu, Manifold embedded knowledge transfer for brain-computer interfaces, IEEE Trans. Neural Syst. Rehabil. Eng. 28 (5) (2020) 1117–1127.
[22] H. Peng, F. Long, C. Ding, Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, IEEE Trans. Pattern Anal. Mach. Intell. 27 (2005) 1226–1238.
[23] I.E. Altman, G. Marco, F. Varetto, Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience, J. Bank. Finance 18 (3) (1994) 505–529.
[24] D. Wu, R. Peng, et al, "Transfer learning for brain-computer interfaces: a complete pipeline," arXiv: 2007.03746, 2020.
[25] P. Guar, K. McCreadie, et al., Tangent space features-based transfer learning classification model for two-class motor imagery brain-computer interface, Int. J. Neural Syst. 29 (10) (2019) 1950025.
[26] W. Wu, Z. Chen, X. Gao, Y. Li, E.N. Brown, S. Gao, Probabilistic common spatial patterns for multichannel EEG analysis, IEEE Trans. Pattern Anal. Mach. Intell. 37 (3) (2015) 639–653.
[27] D. Wu, Y. Xu, B.L. Lu, Transfer learning for EEG-based brain-computer interfaces: a review of progress made since 2016, IEEE Trans. Cogn. Dev. Syst. (2020) 1–15.
[28] A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, J. Neurosci. Methods 134 (1) (2004) 9–21.
[29] G. Pfurtscheller, C. Neuper, Motor imagery and direct brain-computer communication, Proc. IEEE 89 (7) (2001) 1123–1134.
[30] H. Ramoser, J. Müller-Gerking, G. Pfurtscheller, Optimal spatial filtering of single trial EEG during imagined hand movement, IEEE Trans. Rehabil. Eng. 8 (4) (2000) 441–446.
[31] K.K. Ang, Z.Y. Chin, H. Zhang, C. Guan, Filter bank common spatial pattern (FBCSP) algorithm using online adaptive and semi-supervised learning, in: The 2011 International Joint Conference on Neural Networks, San Jose, CA, 2011, pp. 392–396.
[32] C.M. Bishop, Pattern Recognition and Machine Learning, springer, New York, 2006.
[33] K.K. Ang, Z.Y. Chin, H. Zhang, et al., Mutual information-based selection of optimal spatial-temporal patterns for single-trial EEG-based BCIs, Pattern Recognit. 45 (6) (2012) 2137–2144.
[34] L. van der Maaten, G. Hinton, Visualizing data using t-SNE, J. Mach. Learn. Rese. 9 (2008) 2579–2605.
[35] M. Long, J. W, et al., Transfer feature learning with joint distribution adaptation, in: Proceedings of the 2013 IEEE International Conference on Computer Vision, 274, 2013, pp. 2200–2207.
[36] J. Zhang, W. Li, et al., Joint geometrical and statistical alignment for visual domain adaptation, in: 30th IEEE Conference on Computer Vision and Pattern Recognition, 547, 2017, pp. 5150–5158.