A Canonical Correlation Analysis-Based Transfer Learning Framework for Enhancing the Performance of SSVEP-Based BCIs

Qingguo Wei, Yixin Zhang, Yijun Wang, *Member, IEEE*, and Xiaorong Gao, *Member, IEEE*

***Abstract*— A steady-state visual evoked potential (SSVEP)- based brain-computer interface (BCI) can either achieve high classification accuracy in the case of sufficient training data or suppress the training stage at the cost of low accuracy. Although some researches attempted to conquer the dilemma between performance and practicality, a highly effective approach has not yet been established. In this paper, we propose a canonical correlation analysis (CCA)-based transfer learning framework for improving the performance of an SSVEP BCI and reducing its calibration effort. Three spatial filters are optimized by a CCA algorithm with intra- and inter-subject EEG data (IISCCA), two template signals are estimated separately with the EEG data from the target subject and a set of source subjects and six coefficients are yielded by correlation analysis between a testing signal and each of the two templates after they are filtered by each of the three spatial filters. The feature signal used for classification is extracted by the sum of squared coefficients multiplied by their signs and the frequency of the testing signal is recognized by template matching. To reduce the individual discrepancy between subjects, an accuracy-based subject selection (ASS) algorithm is developed for screening those source subjects whose EEG data are more similar to those of the target subject. The proposed ASS-IISCCA integrates both subject-specific models and subject-independent information for the frequency recognition of SSVEP signals. The performance of ASS-IISCCA was evaluated on a benchmark data set with 35 subjects and compared with the state-of-the-art algorithm task-related component analysis (TRCA). The results show that ASS-IISCCA can significantly improve the performance of SSVEP BCIs with a small number of training trials from a new user, thus helping to facilitate their applications in real world.**

Manuscript received 20 February 2023; revised 16 May 2023; accepted 16 June 2023. Date of publication 21 June 2023; date of current version 28 June 2023. This work was supported by the National Natural Science Foundation of China under Grant 62066028 and Grant 61663025. *(Corresponding author: Qingguo Wei.)*

Qingguo Wei and Yixin Zhang are with the Department of Electronic Information Engineering, School of Information Engineering, Nanchang University, Nanchang 330031, China (e-mail: wqg07@163.com; 102800646@qq.com).

Yijun Wang is with the State Key Laboratory on Integrated Optoelectronics, Institute Semiconductors, Chinese Academy of Science, Beijing 100083, China (e-mail: wangyj@semi.ac.cn).

Xiaorong Gao is with the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China (e-mail: gxr-dea@tsinghua.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2023.3288397

***Index Terms*— Brain-computer interface, steady-state visual evoked potential, transfer learning, feature extraction, intra- and inter-subject canonical correlation analysis.**

I. Introduction

STEADY-STATE visual evoked potential (SSVEP) is the brain's response to repetitive visual stimulus with maximum amplitude over the occipital lobe and is characterized by sinusoidal-like waveforms at a stimulation frequency and its harmonics [1], [2]. An SSVEP-based brain-computer interface (BCI) usually contains multiple commands, each of which corresponds to a stimulus with a specific frequency. By gazing at different stimuli, a user can output different commends for controlling an external device. Compared to other types of BCIs, an SSVEP-based BCI has the advantages of higher signal-to-noise ratio and stronger distinguishability, and thereby attracts increasing attention in the research field [3], [4].

To improve the communication speed of SSVEP BCIs, numerous studies were focused on multiple target coding and accurate frequency detecting in the past decades. In terms of the latter, the representative methods include power spectrum density analysis (PSDA) [5], minimum energy combination (MEC) [6], canonical correlation analysis (CCA) [7], [8] and its variants [9], [10], [11], [12], [13], [14], [15], correlation component analysis [16] and task-related component analysis (TRCA) [17]. From a training standpoint, these frequency detection methods can be divided into two main categories: training-free and subject-specific. Training-free methods like PSDA [5] and CCA [7], [8], are more practical because they do not need any calibration data, but have low performance due to large inter-subject variability arising from the complex EEG activity; Subject-specific methods like IT-CCA [11] and TRCA [17], have higher performance at the cost of long training sessions, because they require sufficient calibration data from a specific user to optimize the parameters of his/her classification model. Due to the dilemma between performance and practicality, SSVEP-based BCIs have not yet supported widespread daily use. Thereby, a major challenge that SSVEP-based BCI research faces is how to reduce or suppress training time while maintaining high classification accuracy [18].

Recently, researchers attempted to conquer the dilemma of SSVEP-based BCIs by adopting subject-independent training methods, which require the labeled data from other subjects (named source subjects hereinafter) to estimate the template of a new subject (named target subject hereinafter). These methods exploit inter-subject similarity to transfer a surrogate template of source subjects, which is yielded by averaging labeled data across all source subjects, to a target subject in the hope that the surrogate template can replace his/her real template, e.g., transfer template-based CCA (tt-CCA) [19], online tt-CCA (ott-CCA) [19], transfer template-based combined CCA (combined t-CCA) [20] and a unsupervised adaptive variant of combined t-CCA) [20]. However, these methods ignore the discrepancies in data distribution between subjects arisen from inter-subject variability and non-stationarity of EEG signals, which may impede transferability of data from a subject to another. Thereby, the performance of these methods is far less than that of subject-specific methods. To address this issue, two transfer learning (TL) methods were proposed recently based on a small number of training trials from the target subject. Chiang et al. [21] proposed a least-squares transformation (LST) method for transforming labeled data from several source subjects to fit individual data and compensate the shortage of training data from the target subject. The LST can better match the EEG data of a source subject with those of a target subject and significantly improve the decoding performance of SSVEP BCIs. Wang et al. [22] proposed an inter- and intra-subject maximal correlation (IISMC) algorithm for enhancing the performance of SSVEP BCIs, and presented a TL framework based on random selection of transferred subjects. The algorithm achieved high accuracy and ITR and has great potential for developing high-speed BCIs.

So far, a highly effective TL approach for SSVEP BCIs has not yet been established. In this paper, we proposed a novel CCA-based spatial filtering algorithm, named intra- and inter-subject canonical correlation analysis (IISCCA), for creating three different types of spatial filters, each of which is estimated by a training data-driven CCA (TDCCA) algorithm [15] with EEG data from the same domain or two different domains. The IISCCA not only contains the subject-specific information, but also incorporates similarity between subjects. To reduce the variability between subjects, an accuracy-based algorithm for subject selection (ASS) is designed for screening those source subjects whose data distributions are more similar to those of the target subject. Subsequently, we developed a cross-subject TL framework to improve the performance of SSVEP BCIs and reduce their calibration time. The proposed algorithm ASS-IISCCA was evaluated on a 40-target SSVEP dataset containing 35 subjects and compared with the state-of-the-art algorithm TRCA in terms of accuracy and ITR. The results show that ASS-IISCCA outperforms TRCA especially when the number of training trials is small and can significantly reduce the training time.

II. Methods

This section details related works, the basic principle of IISCCA and the algorithmic steps of ASS.

A. Related Works

1) Canonical Correlation Analysis (CCA): CCA is a multivariable statistical analysis method for measuring underlying correlations between two multidimensional variables. It aims at finding two weight vectors to maximize the correlation between the two variables [23], [24]. Given two variables X and Y, CCA creates two weight vectors w_x and w_y such that the correlation between the two linear combinations $x = w_x^T X$ and $y = w_y^T Y$ is maximized

$$\begin{aligned}\rho &= \max_{w_x, w_y} \frac{E[xy^T]}{\sqrt{E[xx^T]E[yy^T]}} \\ &= \max_{w_x, w_y} \frac{E[w_x^T XY^T w_y]}{\sqrt{E[w_x^T XX^T w_x]E[w_y^T YY^T w_y]}}\end{aligned} \tag{1}$$

where superscript T represents the transpose operation. Maximizing Eq. (1) is equivalent to solving a generalized eigenvalue problem. The maximum of ρ with respect to w_x and w_y is called maximum canonical correlation.

a) Standard CCA: In SSVEP-based BCI applications, the weight vector w_x can be used as a spatial filter to extract useful information for frequency recognition from multi-channel SSVEP signals. The input X is a single-trial testing signal, whereas the input Y is a reference signal that can be either a universal template (UT) for all users arising from sine-cosine signals [7] or an individual template (IT) for a specific user yielded by averaging training data across trials [11]. Given a stimulus with frequency f_i, $i = 1, 2, \cdots, N_f$, the UT can be constructed as follows

$$Y_i = \begin{bmatrix} \sin(2\pi f_i t) \\ \cos(2\pi f_i t) \\ \vdots \\ \sin(2\pi N_k f_i t) \\ \cos(2\pi N_k f_i t) \end{bmatrix}; \quad t = [1, 2, \cdots, N_p]/F_s \tag{2}$$

where N_k and N_p denote the number of harmonics and the number of sampling points respectively, and F_s is the sampling frequency of SSVEP signals. Assume that the training signals of N_t trials for a stimulus target i are represented as X_i^h, $h = 1, 2, \cdots, N_t$, the IT is calculated as follows

$$Y_i = \bar{X}_i = \frac{1}{N_t}\sum_{h=1}^{N_t} X_i^h \tag{3}$$

For either case, the stimulus frequency are recognized by maximum canonical correlation. Given a single-trial testing signal X, its frequency f_t is decided as the frequency of the template with the maximum canonical correlation as

$$f_t = \underset{f}{\operatorname{argmax}}\ \rho_f, \quad f = f_1, f_2, \cdots, f_K \tag{4}$$

The UT- and IT-based CCA algorithms are named as UTCCA and ITCCA respectively hereinafter.

b) Training data-driven CCA (TDCCA): We proposed a training data-driven CCA algorithm for a specific user [15]. Based on the assumption that the template signal of a stimulus frequency is considered as the SSVEP component of a single-trial EEG signal, which is invariant for different trials of the same stimulus. Thereby, the objective function of TDCCA can also be formulated with Eq. (1), but unlike ITCCA, the two variables X and Y are the continuous EEG signal of multiple training trials $X = [X^1, X^2, \cdots, X^{N_t}]$ and continuous template signal $Y = [Y^1, Y^2, \cdots, Y^{N_t}] = [\bar{X}, \bar{X}, \cdots, \bar{X}]$ respectively, where $X^h \in R^{N_c \times N_p}, h = 1, 2, \cdots, N_t$ and $Y^h = \bar{X} = (1/N_t)\sum_{h=1}^{N_t} X^h$ are a single-trial training signal and its template respectively.

Unlike ITCCA, TDCCA extracts features using the filter w_x and recognizes the stimulus frequency by template matching. Given a single-trial testing signal $\tilde{X}$ and the template signal from ith stimulus target $\bar{X}_i$, they are spatially filtered with $w_{x,i}$, leading to two vectors $x_i = w_{x,i}^T \tilde{X}$ and $y_i = w_{x,i}^T \bar{X}_i$. Then the correlation coefficient between x_i and y_i is used as the feature signal for target recognition

$$r_i = \hat{\rho}(w_{xi}^T \tilde{X}, w_{xi}^T \bar{X}_i), \quad i = 1, 2, \cdots, N_f \tag{5}$$

where $\hat{\rho}(a, b)$ denotes Pearson correlation coefficient between two vectors a and b. Finally, the stimulus frequency corresponding to the testing trial can be decided by matching the testing vector x_i with each of the template vectors y_i as

$$f_t = \arg\max_i \; r_i, \quad i = 1, 2, \cdots, N_f \tag{6}$$

2) Task-Related Component Analysis (TRCA): TRCA is currently one of the most popular algorithms for SSVEP identification [17]. It extracts task-related components by maximizing the reproducibility of inter-trial covariances. The goal of TRCA is to optimize a spatial filter w so as to the temporal profile has the maximal similarity among different training trials. Assume $X^i \in \mathbb{R}^{N_c \times N_s}$ and $X^j \in \mathbb{R}^{N_c \times N_s}$ are the EEG signals from ith trial and jth trial respectively. The constrained optimization of TRCA boils down to the following Rayleigh-Ritz quotient problem

$$w = \underset{w}{\mathrm{argmax}} \; \frac{w^T S w}{w^T Q w} \tag{7}$$

where S and Q are the sum of inter-trial covariance matrices and the sum of auto-covariance matrices respectively, and are represented respectively as follows

$$S = \sum_{\substack{i,j=1 \\ i \neq j}}^{N_t} \mathrm{cov}(X_i, X_j), \quad Q = \sum_{\substack{i,j=1 \\ i = j}}^{N_t} \mathrm{cov}(X_i, X_j) \tag{8}$$

The filter w can be obtained as the eigenvector of matrix $Q^{-1}S$ corresponding to its largest eigenvalue. Having obtained the spatial filter w, the feature extraction and frequency recognition are performed by Eq. (5) and Eq. (6) respectively.

3) Filter Bank Analys Is and Ensemble Spatial Filter: In the previous studies, two effective methods were proposed to increase the SNR of SSVEP signals. Filter bank analysis [25], [26] aims to decompose an SSVEP signal into multiple sub-band signals and combine harmonic components embedded in these sub-bands. Specifically, the method decomposes the wide band of 8-88 Hz into 10 different sub-bands, each of which ranges between m×8 Hz and 88 Hz. Suppose that a single-trial testing signal and the template signals for the bth sub-band are denoted as $X^b \in R^{N_c \times N_l}$ and $\chi^b \in R^{N_b \times N_c \times N_l}$ respectively. In the sub-band, the feature value from the ith stimulus can be calculated as $\hat{r}_i^b = f(\bar{\chi}_i^b, X^b)$, where f denotes a kind of spatial filtering algorithm such TRCA and IISCCA. The feature signal used for target recognition is calculated as

$$r_i = \sum_{b=1}^{N_b} a(b) \cdot (\hat{r}_i^b)^2 \tag{9}$$

where N_b is the number of sub-bands and $a(b) = b^{-1.25} + 0.25$ is the weight coefficient defined in [41]. Finally, the frequency of the testing signal is recognized using Eq. (6). In the study, the number of sub-bands N_b was set to 5 [17].

Ensemble spatial filter [17] aims at incorporating the spatial filters from all stimulation targets. This method is based on the assumption that the filters within the same frequency band from different targets are similar to each other and integrating them can improve the performance of spatial filtering. Thereby, an ensemble spatial filter is created by concatenating these filters $W = [w_1, w_2, \cdots, w_{N_f}]$. Then Eq. (5) for computing correlation coefficient for the ith stimulus target is modified as

$$\hat{r}_i = \hat{\rho}(\tilde{X}^T W_i, \bar{X}_i W_i), \quad i = 1, 2, \cdots, N_f \tag{10}$$

B. Intra- and Inter-Subject CCA (IISCCA)

Existing training data-based algorithms such as ITCCA, TDCCA and TRCA extract task-related knowledge with EEG data from a single target subject, and thus can be used for estimating one spatial filter. As an extension of TDCCA [15], IISCCA aims to extract task-related knowledge with EEG data from a target subject, a set of source subjects and both a target subject and a set of source subjects, and thus can be used for estimating three spatial filters. Thereby, IISCCA includes two types of CCA algorithms, i.e., intra-subject CCA and inter-subject CCA. Assume that the target subject and each of the N_s source subjects have the EEG data from N_t labeled trials, $X^h \in R^{N_c \times N_l}$, $h = 1, 2, \cdots, N_t$. The two types of CCA algorithms are described as follows

1) Intra-Subject CCA: The goal of the intra-subject CCA is to find two weight vectors to maximize the correlation between the two variables arisen respectively from either a target subject or a set of source subjects. For a stimulus target, this CCA includes two optimization problems, one of which is the same as TDCCA, which is based solely on the training data from the target subject. The optimization problem formulated in Eq. (1) is rewritten as follows

$$\rho_1 = \arg\max_{w_x, w_y} \frac{E[w_t^T X_t Y_t^T w_y]}{\sqrt{E[w_x^T X_t X_t^T w_x] E[w_y^T Y_t Y_t^T w_y]}} \tag{11}$$

where X_t and Y_t are the same as X and Y for TDCCA respectively. The weight vector w_x is adopted as a spatial filter for feature extraction and renamed as w_1 hereinafter. The other

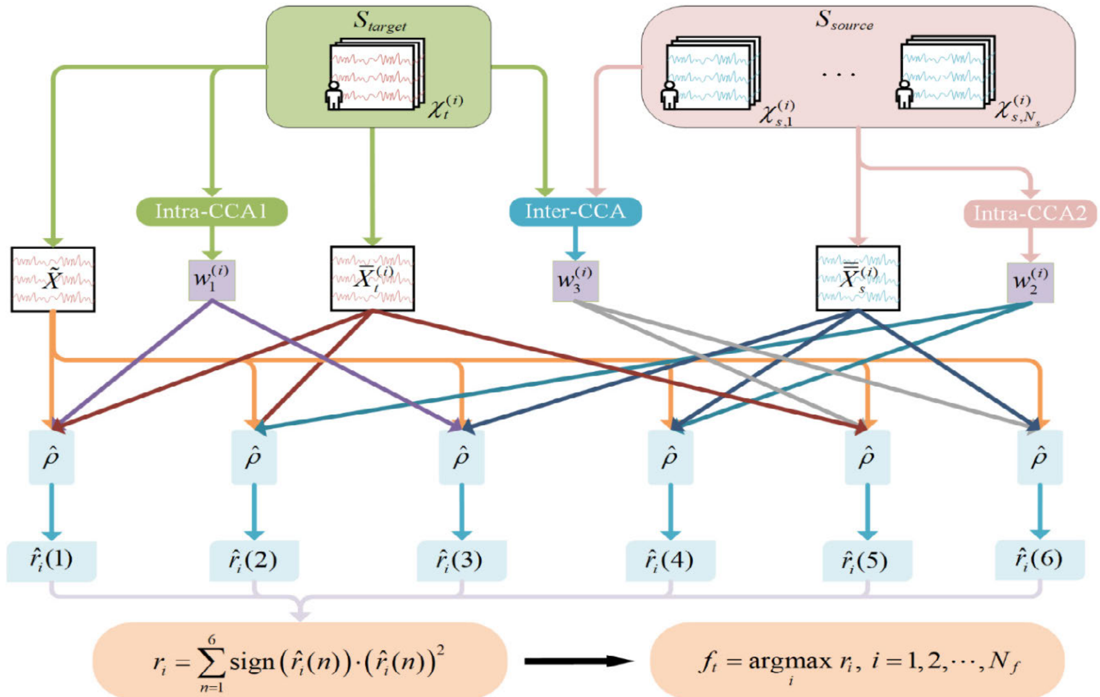

Fig. 1. The flow chart of the proposed IISCCA algorithm for target recognition. In all variables, the index i denotes ith stimulus target/frequency. $\chi_t^{(i)}$ and $\chi_{s,m}^{(i)}$, $m = 1, 2, \cdots, N_s$ are the EEG data from the target subject and mth source subject respectively. $\tilde{X}$ and $\bar{X}_t^{(i)}$ denote the testing data and the template data from the target subject respectively. $\bar{\bar{X}}_s^{(i)}$ is the template data form the surrogate source subject (i.e., the set of N_s source subjects). $w_1^{(i)}$, $w_2^{(i)}$ and $w_3^{(i)}$ are the three spatial filters yielded by labeled EEG data from the target subject, the surrogate source subject and both respectively. $\hat{\rho}$ denotes Pearson correlation coefficient between two vectors yielded by spatially filtering the testing data and template data.

optimization problem is formulated as

$$\rho_2 = \underset{w_x, w_y}{\arg\max} \frac{E[w_x^T \bar{X}_s \bar{Y}_s^T w_y]}{\sqrt{E[w_x^T \bar{X}_s \bar{X}_s^T w_x] E[w_y^T \bar{Y}_s \bar{Y}_s^T w_y]}} \tag{12}$$

where $\bar{X}_s = [\bar{X}_s^1, \bar{X}_s^2, \cdots, \bar{X}_s^{N_t}]$ is yielded by concatenating the N_t averaged EEG trials across N_s source subjects, i.e. $\bar{X}_s^h = (1/N_s) \sum_{m=1}^{N_s} \bar{X}_m^h, h = 1, 2, \cdots, N_t$, which we call the EEG trials of the surrogate source subject, and $\bar{Y}_s = [\bar{Y}_s^1, \bar{Y}_s^2, \cdots, \bar{Y}_s^{N_t}] = [\bar{\bar{X}}_s, \bar{\bar{X}}_s, \cdots, \bar{\bar{X}}_s]$, where $\bar{\bar{X}}_s$ is the template signal yielded by averaging the N_t labeled EEG trials from the surrogate source subject, i.e., $\bar{\bar{X}}_s = (1/N_t) \sum_{h=1}^{N_t} \bar{X}_s^h$. The weight vector w_x is used as a spatial filter for feature extraction and renamed as w_2 hereinafter.

2) Inter-Subject CCA: The inter-subject CCA is to find two weight vectors to maximize the correlation between the two variables arisen respectively from a target subject and a set of source subjects. For a stimulus target, the optimization problem is formulated as follows

$$\rho_3 = \underset{w_x, w_y}{\arg\max} \frac{E[w_x^T X_t \bar{Y}_s^T w_y]}{\sqrt{E[w_x^T X_t X_t^T w_x] E[w_y^T \bar{Y}_s \bar{Y}_s^T w_y]}} \tag{13}$$

where X_t is the same as the X_t in TTCCA, and $\bar{Y}_s$ is the same as the $\bar{Y}_s$ in TDCCA. The weight vector w_x is used as a spatial filter for feature extraction and renamed as w_3.

Given a single-trial testing signal $\tilde{X}$, a correlation vector consisting of six correlation coefficients can be achieved by combining the two template signals $\bar{X}_t$ and $\bar{\bar{X}}_s$ with the three spatial filters w_1, w_2 and w_3 for a stimulus target i as follows

$$\hat{r}_i = \begin{bmatrix} \hat{r}_i(1) \\ \hat{r}_i(2) \\ \hat{r}_i(3) \\ \hat{r}_i(4) \\ \hat{r}_i(5) \\ \hat{r}_i(6) \end{bmatrix} = \begin{bmatrix} \hat{\rho}(\tilde{X}^T w_1, \bar{X}_t^T w_1) \\ \hat{\rho}(\tilde{X}^T w_1, \bar{\bar{X}}_s^T w_1) \\ \hat{\rho}(\tilde{X}^T w_2, \bar{X}_t^T w_2) \\ \hat{\rho}(\tilde{X}^T w_2, \bar{\bar{X}}_s^T w_2) \\ \hat{\rho}(\tilde{X}^T w_3, \bar{X}_t^T w_3) \\ \hat{\rho}(\tilde{X}^T w_3, \bar{\bar{X}}_s^T w_3) \end{bmatrix} \tag{14}$$

Then the feature signal for target/frequency recognition is generated as follows

$$r_i = \sum_{n=1}^{6} sign(\hat{r}_i(n)) \cdot (\hat{r}_i(n))^2 \tag{15}$$

where $sign$ function is applied to keep discriminative information from negative correlation coefficients. Finally, the target frequency f_t is recognized using Eq. (7). The proposed IISCCA algorithm is illustrated in Fig. 1.

C. Accuracy-Based Subject Selection (ASS)

TL is an effective approach for reducing the amount of training data from the target subject. However, not EEG data from all source subjects are suitable for transferring because of the variability between subjects. Thus, it is necessary to select a best subset of source subjects for TL, whose EEG data are more similar to those of the target subject. Several methods for subject selection were proposed for motor imagery-based BCIs such as sequential forward floating search (SFFS) algorithm [27], Jensen Shannon ratio (JSR) [28] and Kullback-Leibler (KL) divergence [29], but so far there is no highly efficient method proposed for SSVEP-based BCIs. In this study, we proposed an accuracy-based subject selection (ASS) algorithm, which is adapted from the SFFS algorithm.

The ASS algorithm is shown in Algorithm 1, in which the EEG data from N_s source subjects $\chi_{s,m}^{(i)}$, $m = 1, 2, \cdots, N_s$, are used as training data, whereas the training data from the target subject $\chi_t^{(i)}$ are employed as testing data. The function $acc(m) = trainThenTest(\chi_{s,m}, \chi_t)$ returns the classification accuracy yielded when EEG data from mth source subject $\chi_{s,m}$ are employed to create a spatial filter and a template for each stimulus frequency, which are then utilized to classify the testing data χ_t using the TDCCA algorithm. Then all the source subjects are sorted by their classification accuracies. The first C source subjects are selected as transferred subjects, if their EEG data are pooled together as training data and achieve the highest accuracy among all subsets of source subjects. The number of training trials from the target subject was set as 5 trials per stimulus frequency in the study. The combination of ASS and IISCCA is named ASS-IISCCA.

Algorithm 1 Accuracy-Based Subject Selection (ASS)

Input: χ_t, training data of the target subject
$\chi_{s,m}$, EEG data of N_s source subjects
Output: $\chi_{s,c}$, EEG data of C selected source with high similarity to the target subject
C, the index of selected source subjects

for $m = 1$ to N_s do
  $acc(m) = \text{trainThenTest}(\chi_{s,m}, \chi_t)$
end for
$[\sim, index] = \text{sort}(acc, \text{'}descend\text{'})$
for $n = 1$ to N_s do
  $acc(n) = \text{trainThenTest}(\chi_{s,index(1:n)}, \chi_t)$
end for
$[\sim, maxIndex] = \max(acc)$
$\chi_{s,C} = \chi_{s,index(1:\max Index)}$
$C = index(1 : maxIndex)$

III. Data Acquisition and Preprocessing

The proposed algorithm was evaluated on a benchmark dataset [30] acquired from an SSVEP-based BCI speller experiment. The visual speller contained 40 stimulus targets. The stimulus frequencies ranged from 8 Hz and 15.8 Hz with an interval of 0.2 Hz, whereas the phases ranged between 0 rad and 1.5π rad with an interval of 0.5π rad. Thirty-five healthy subjects (18 males, mean 22 years old) participated the experiment. Eight of them had experience using SSVEP-based BCIs and the others were naive. The experiment consisted of six blocks, each of which included 40 trials corresponding to the 40 targets cued in a random order. Each trial started with a visual cue of length 0.5 s, which was achieved by reddening the target to be gazed at. Subjects were asked to shift their gaze to the target as fast as possible. Then all stimulus targets flickered for 5 s. After visual stimulation, the screen was blank for 0.5 s and then the next trial began. The continuous data were resampled to 250 Hz for offline analysis.

EEG data from the nine channels over occipital lobe (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, O2) were used for this study. All data epochs of these channels were band-pass filtered in the frequency band 8-88 Hz with an IIR filter of Chebyshev type I. To avoid phase distortion, the filtering was performed forward and backward. Considering a latency delay of 0.14 s in the visual system and 0.5 s for gaze shifting, the filtered epochs were intercepted in the window [0.64 s, $(0.64 + d)$ s], where d denotes the data length used for target recognition.

IV. Results

The data set contained 35 subjects, who can be divided into three groups according to their experience with the BCI, namely the experienced (the first eight subjects), naive (the latter 27 subjects), and a mixture of the two groups (all 35 subjects). We used a leave-one subject-out cross validation strategy to divide the subjects from each group into the target subject and source subjects, i.e., each subject of a group acts as the target subject once, and the other subjects act as the source subjects. In addition, all source subjects per group were used as transferred subjects for IISCCA, whereas the selected source subjects by ASS were employed as transferred subjects for ASS-IISCCA. IISCCA and ASS-IISCCA were evaluated by comparing them with TRCA **under condition** 1 **(w/o FB)**, i.e., the ensemble spatial filter was adopted, but the filter bank analysis was not, and **condition 2 (w/ FB)**, i.e., the ensemble spatial filter and the filter bank analysis were adopted jointly.

The performance evaluation is conducted from several aspects such as classification accuracy, simulated ITR and feature distribution. Classification accuracy is defined as the ratio of the number of targets correctly recognized to total number of targets, whereas ITR defined by Gao et al. [3] is calculated as follows

$$ITR = \left(\log_2 M + P\log_2 P + (1-P)\log_2\left[\frac{1-P}{M-1}\right]\right)\left(\frac{60}{T}\right) \tag{16}$$

where M is the number of targets, P is the detection accuracy of targets, and T in seconds/detection is the average time for a detection. For the estimation of ITRs, the 0.5 s used for gaze shifting was included in the time for target detection.

Regarding the calculation of classification accuracy and ITR at different numbers of training trails, a leave-one trial-out cross validation strategy was used for all the three comparison algorithms, i.e., each of the six trials per stimulus frequency from the target subject was used for testing and the remaining five trials were used as the training set. Different numbers of training trials were sequentially extracted from the training set. The cross validation led to six classification experiments and the classification accuracy and ITR were the mean of six classification accuracies and ITRs respectively.

A. Classification Accuracy and ITR

Since the performance of an SSVEP BCI is closely related to the data length, the number of channels and the number of the training blocks used for target recognition, we explored

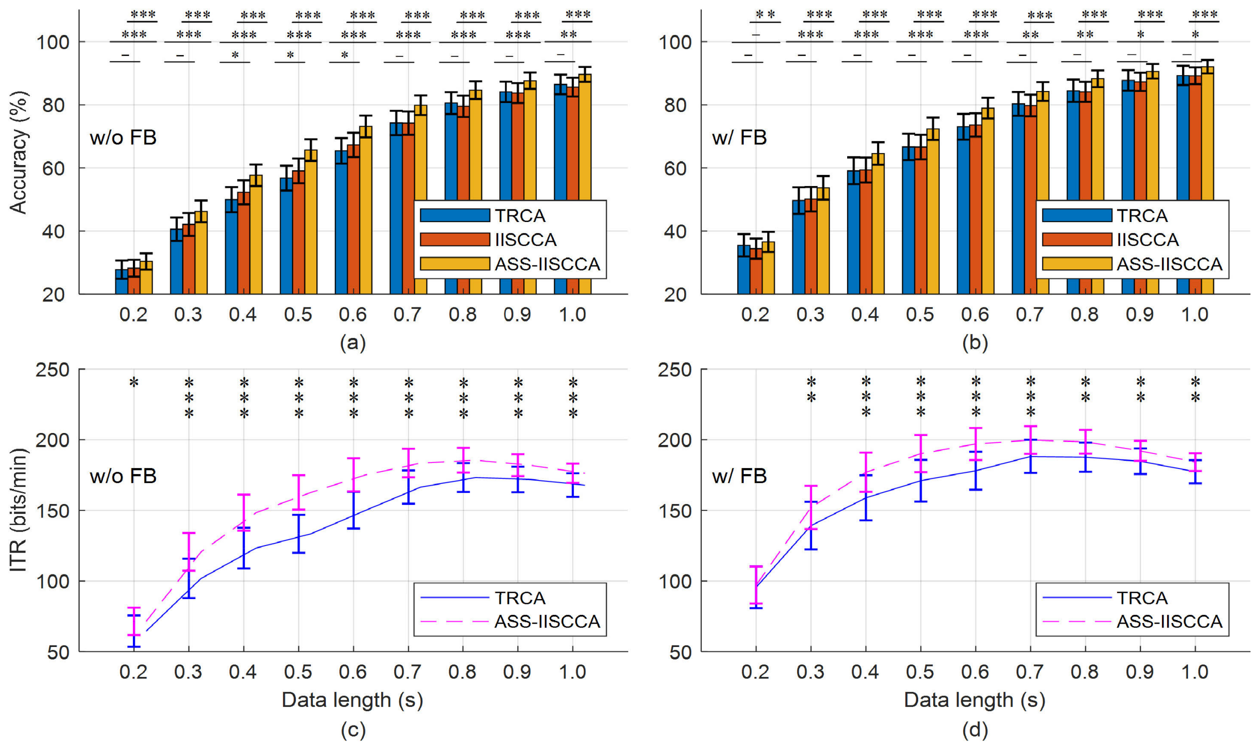

Fig. 2. Averaged classification accuracies of the three algorithms (TRCA, IISCCA and ASS-IISCCA) and averaged ITRs of the two algorithms (TRCA and ASS-IISCCA) across all 35 subjects at nine data lengths, under condition 1 (w/o FB) and condition 2 (w/ FB) respectively. Error bars indicate standard errors. The asterisks indicate the statistically significant difference between two algorithms (i.e., p value yielded by the paired t-test): * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$, whereas the dashed line denotes no statistically significant difference: – $p \geq 0.05$.

the influence of the these parameters on classification accuracy and ITR. The accuracy was compared among the three algorithms, but the ITR was compared between ASS-IISCCA and TRCA for simplicity. For the group of 35 subjects, accuracy and ITR changes with these parameters are illustrated in Fig. 2, 3 and 4 respectively. Fig. 2 shows the averaged accuracies across all subjects at nine different data lengths ranging from 0.2s to 1s with the interval of 0.1s. The number of channels and the number of training blocks were fixed at 9 and 3 respectively. From the figure, it is easily seen that for each of the two conditions, accuracies of these three algorithms increase monotonically with the data length, indicating its important impact on the BCI performance; The ITRs of both algorithms first increase with data length and decrease after reaching the peak. Their highest ITRs are yielded at the data length of 0.8s and 0.7s under condition 1 (w/o FB) and condition 2 (w/ FB respectively. Under condition 1, ASS-IISCCA significantly outperforms TRCA and IISCCA at all data lengths in accuracy, and significantly outperforms TRCA in ITR at all data lengths; IISCCA significantly outperforms TRCA in accuracy at the data lengths of 0.4s, 0.5s and 0.6s, and they do not have significant difference at other data lengths; Under condition 2, ASS-IISCCA significantly outperforms both TRCA and IISCCA in accuracy at all data lengths except for TRCA at data length of 0.2s. IISCCA and TRCA do not have significant difference in accuracy at all data lengths. ASS-IISCCA significantly outperforms TRCA in ITR at all data lengths except for the data length of 0.2s. These results validate that both the algorithm IISCCA for feature extraction and the algorithm ASS for subject selection are effective for improving the performance of SSVEP BCIs.

Fig. 3 shows the averaged classification accuracies and ITRs across subjects at seven different numbers of channels ranging between 3 and 9 with the interval of 1. The data length and the number of training blocks were fixed at 0.7s and 3 respectively. For simplicity, the channels were sequentially selected from the nine channels used for the study as indicated in section III. From the figure, it is revealed that for each of the two conditions, the accuracies of all the three algorithms and ITRs of the two algorithms increase monotonically with the number of channels, indicating its important impact on the BCI performance. Under condition 1, ASS-IISCCCA is superior to TRCA and IISCCA at all numbers of channels except for IISCCA at 3 and 4 channels. IISCCA is superior to TRCA in accuracy at first five numbers of channels and they have no significant difference in accuracy between IISCCA and TRCA at 8 and 9 channels; Under condition 2, ASS-IISCCA is superior to TRCA and IISCCA in accuracy at all numbers of channels except for IISCCA at 3 channels. IISCCA is superior TRCA in accuracy at first 4 numbers of channels and they have no significant difference in accuracy at latter 3 numbers of channels. Under either condition, ASS-IISCCA is superior to TRCA in ITRs at all numbers of channels. These results demonstrate the effectiveness of ASS for improving the performance of IISCCA. Moreover, both ASS-IISCCA and IISCCA show better performance than TRCA in the case of fewer channels, thus improving the convenience of BCI usage.

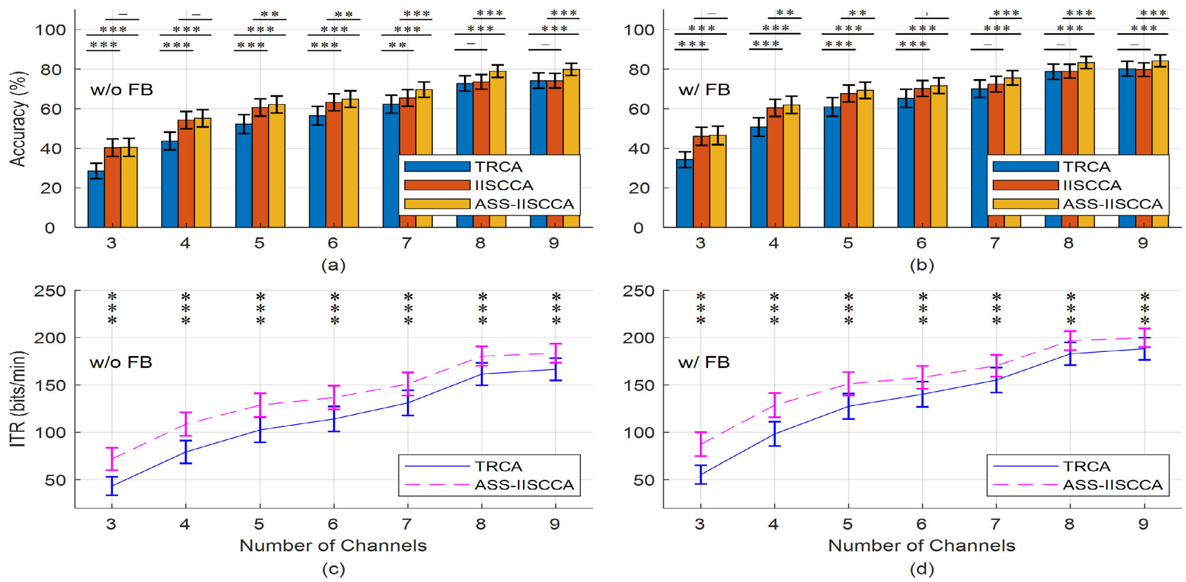

Fig. 3. Averaged classification accuracies of the three algorithms (TRCA, IISCCA and ASS-IISCCA) and averaged ITRs of the two algorithms (TRCA and ASS-IISCCA) across all 35 subjects at seven groups of channels under condition 1 (w/o FB) and condition 2 (w/ FB) respectively.

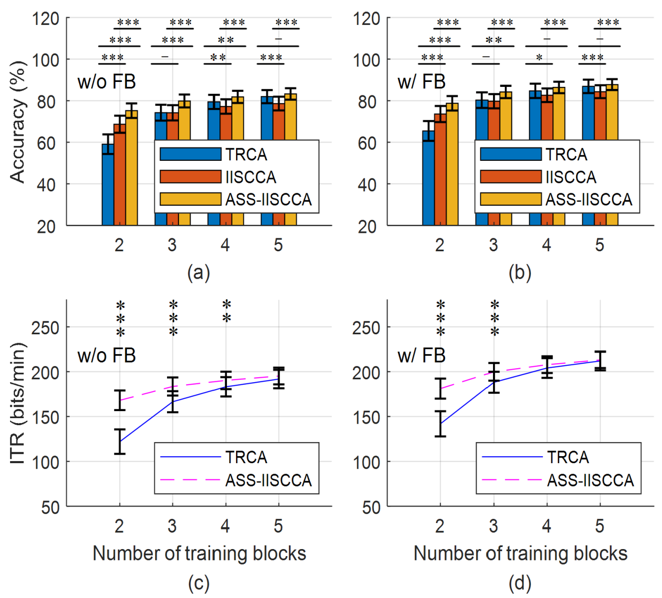

Fig. 4. Averaged classification accuracies of the three algorithms (TRCA, IISCCA and ASS-IISCCA) and averaged ITRs of the two algorithms (TRCA and ASS-IISCCA) across subjects at four numbers of training blocks under condition 1 (w/o FB) and condition 2 (w/ FB) respectively.

Fig. 4 depicts the averaged classification accuracies of the three algorithms and the averaged ITRs of the two algorithms across subjects at four different numbers of training blocks (i.e., training trials per target) ranging from 2 to 5 with the interval of 1. The data length and the number of channels are set as 0.7s and 9 respectively. As shown in the figure, the accuracies of the three algorithms and ITRs of the two algorithms increase with the number of training blocks for the two conditions, indicating its important impact on BCI performance. Under condition 1, ASS-IISCCA outperforms IISCCA and TRCA in accuracy at all numbers of training blocks and the first 3 numbers of training blocks respectively. ASS-IISCCA outperforms TRCA in ITR at first 3 numbers of training blocks; under condition 2, ASS-IISCCA outperforms IISCCA and TRCA at all numbers of training blocks and the first 2 numbers of training blocks respectively. ASS-IISCCA outperforms TRCA at the first 2 numbers of training blocks. These results suggest the validity of the ASS algorithm for improving performance of SSVEP BCIs and the usefulness of ASS-IISCCA for reducing the amount of training data. As an objective comparison of calibration time, TRCA required 3 and 4 training blocks to obtain the accuracies of 79.43% and 84.73% under the condition 1 and 2 respectively, whereas ASS-IISCCA required only 2 and 3 training blocks to achieve two similar accuracies 79.86% and 84.19% under condition 1 and 2 respectively, reducing the calibration time by 1/3 and 1/4 respectively.

Fig. 5 illustrates averaged accuracies of the three algorithms and averaged ITRs of the two algorithms across subjects from three groups of subjects at data length of 0.7s, 9 channels and 3 training blocks. Under condition 1, in accuracy, ASS-IISCCA is significantly better than TRCA for all three groups of subjects, is significantly better than IISCCA for the two groups of naïve subjects and all subjects, and has no significant difference with IISCCA for the group of experienced subjects. In ITR, ASS-IISCCA is significantly better than TRCA for all three groups of subjects; under condition 2; in accuracy, ASS-IISCCA is significantly better than both TRCA and IISCCA for the two groups of naïve subjects and all subjects, and has no significant difference with them for the group of experienced subjects. In ITR, ASS-IISCCA is significantly better than TRCA for the two groups of naïve subjects and all subjects and has no significant difference with it for the group of experienced subjects. These results suggest that for the group of experienced subjects, all the three algorithms can achieve high performance under either condition, whereas for other two groups, ASS-IISCCA can significantly improve the

TABLE I

INDEXES OF TRANSFERRED SOURCE SUBJECTS FOR EACH TARGET SUBJECT IN THE GROUP OF 8 EXPERIENCED SUBJECT

Target subject	1	2	3	4	5	6	7	8
Transferred subjects	2,3,5,7	3,5,6,7	1,5,7	1,6,7	1,2,3	2,3,5,7	1,2,3	2,5,6,7

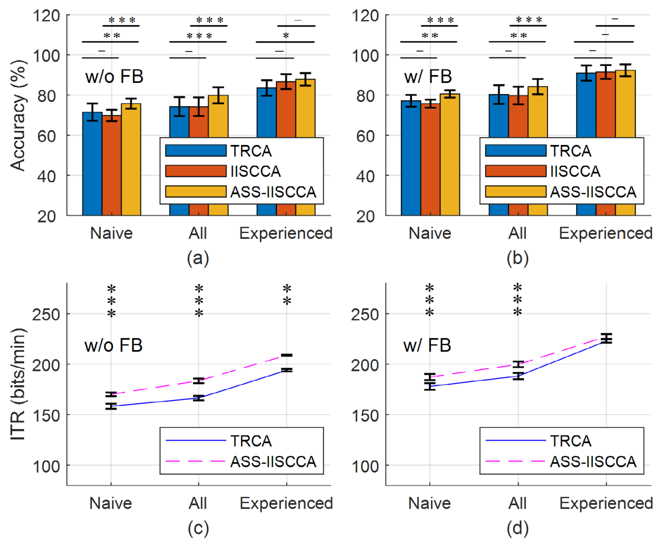

Fig. 5. Averaged accuracies of the three algorithms (TRCA, IISCCA and ASS-IISCCA) and averaged ITRs of the two algorithms (TRCA and ASS-IISCCA) across subjects from three groups (i.e., 27 naive, all 35 subjects and 8 experienced) at data length of 0.7 s, 9 channels and 3 training blocks.

accuracy and ITR under either condition. This is meaningful because with any one classical algorithm, experienced subjects tend to yield high performance, and the focus of BCI studies is how to increase performance of ordinary users especially novice users. Hence, the proposed algorithm is expected to facilitate the practical applications of SSVEP BCIs.

B. Feature Distribution

For IISCCA and ASS-IISCCA, the similarity information between subjects is of prime importance. In order to further investigate the performance of the proposed algorithm, a 2-dimensional t-SNE [31] is used for comparing the 40-dimensional features yielded by the three algorithms. Fig. 6 illustrates the two-dimensional distributions of the 40-dimensional features obtained by TRCA, IISCCA and ASS-IISCCA under condition 1 (w/o FB) and condition 2 (w/ FB) using a 0.7s time window for an example subject (S22). Each point in the figure denotes one of the 240 trials, and each color denotes a stimulation frequency/class. As shown in the figure, for each of the two conditions, from left to right, the latter algorithm produced tighter clusters and more separable classes in the feature space than the former one. The reason is that compared to TRCA, which leverages only subject-specific information, IISCCA and ASS-IISCCA exploit the similarity of other subjects to extract more discriminative features for frequency recognition, and ASS-IISCCA further screens more similar source subjects for feature extraction, thus performing the best among the three algorithms.

C. The Number and Indexes of Transferred Subjects

Fig. 7 shows the number of transferred source subjects for each target subject in the three groups. It is revealed from the figure that for the group of experienced subjects, each target subject selected 3 or 4 source subjects for transferring, whereas for the other two groups of subjects, most of target subjects selected more than 5 source subjects for transferring. The reason may be that the EEG data from experienced subjects are more similar to each other than those from naive and mixed subjects. Taking the group of 8 experienced subjects as an example, the indexes of selected/transferred source subjects for each target subject are reported in Table I.

D. Ablation Experiments

Fig. 8 illustrates the averaged accuracies and ITRs across all 35 subjects yielded by the proposed algorithm ASS-IISCCA and the modified algorithms yielded by removing either the filter w2 yielded by Eq. (13) or the filter w3 yielded by Eq. (14) under the two conditions. Removing the filter w3, accuracies and ITRs decrease at all data lengths except for the data lengths of 0.2s and 0.3s under the condition 1, and at all data lengths under the condition 2; Removing the filter w2, accuracies and ITRs decrease at all data lengths except for 0.7s, 0.9s and 1.0s under condition 1, and decrease at all data lengths except for 0.7s, 0.8s and 0.9s under condition 2. In summary, the filter w3 contributed more to accuracies and ITRs at larger data lengths, whereas w2 contributed more to smaller data lengths. Therefore, the two filters had complementary property at most data lengths and their combination helped to improve the performance and the robustness of ASS-IISCCA.

E. Computational Complexity

A further experiment was performed to explore computational complexity of the three algorithms, TRCA, IISCCA and ASS-IISCCA. Each algorithm includes the two stages of training and testing. In the training stage, these algorithms contain the main procedures of sub-band filtering of training data, estimation of spatial filters, and creation of templates, and additional procedures of source subject selection where applicable; In the testing stage, these algorithms contain the procedures of sub-band filtering of testing data, spatial filtering of both testing and template data, feature extraction and frequency recognition. This experiment was performed under Matlab R2020a on a computer of the brand Hp, configured with Intel (R) Core (TM) i5-6500 CPU @3.20GHz, 4-core, 16 GB RAM, 64-bit Windows10. For TRCA and IISCCA, their training times are 0.28s and 0.38s respectively, whereas their testing times are 0.13s and 0.12s respectively. Thereby, the two algorithms can be directly applied to online experiments without any modification. For ASS-IISCCA,

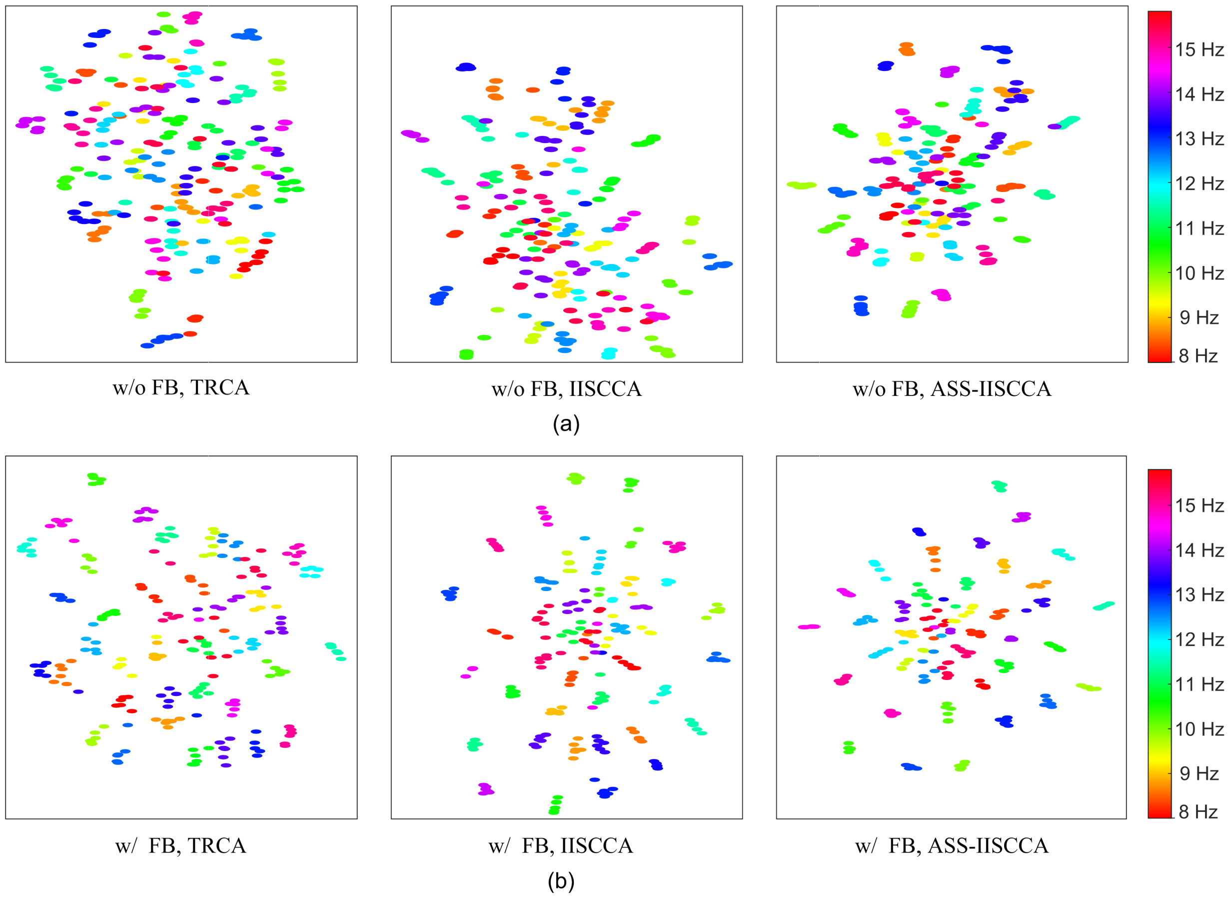

Fig. 6. Two-dimensional t-SNE visualization of the 40-dimensional features obtained by TRCA, IISCCA and ASS-IISCCA under (a) condition 1 (w/o FB) and (b) condition 2 (w/ FB) using a 0.7s time window for an example subject (S22). Each two-dimensional point represents one of the 240 trials (40 (targets)×6 (trials)=240 trials), and each color represents a stimulation frequency/class.

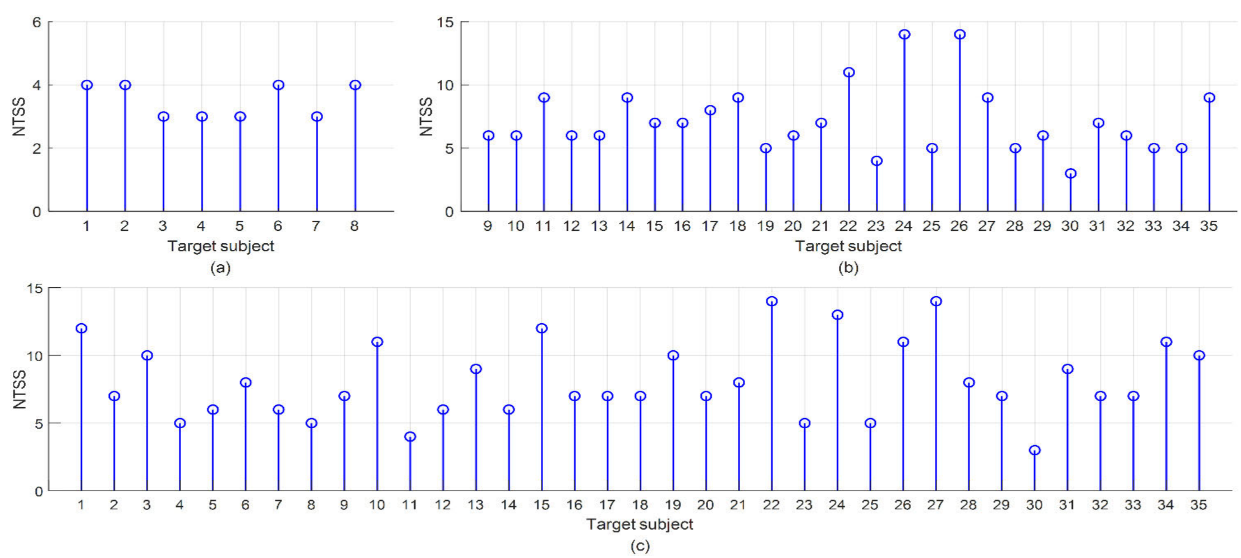

Fig. 7. Number of transferred source subjects for each target subject in the three groups of subjects: (a) 8 experienced, (b) 27 naive, and (c) all 35 subjects. NTSS denotes the number of transferred (or selected) source subjects.

an additional step, the selection of source subjects, is included in the training procedure, and its running time depends upon the number of source subjects. The more the number of source subjects, the longer is the running time. Taking the group of 8 experienced subjects as an example, the averaged time across subjects for the selection of source subjects is about 83 s, or 1.4 minutes. Such a time is too long for online applications. However, having obtained a small amount of training data from the target subject, the training procedure can be done before online testing, and thereby ASS-IISCCA

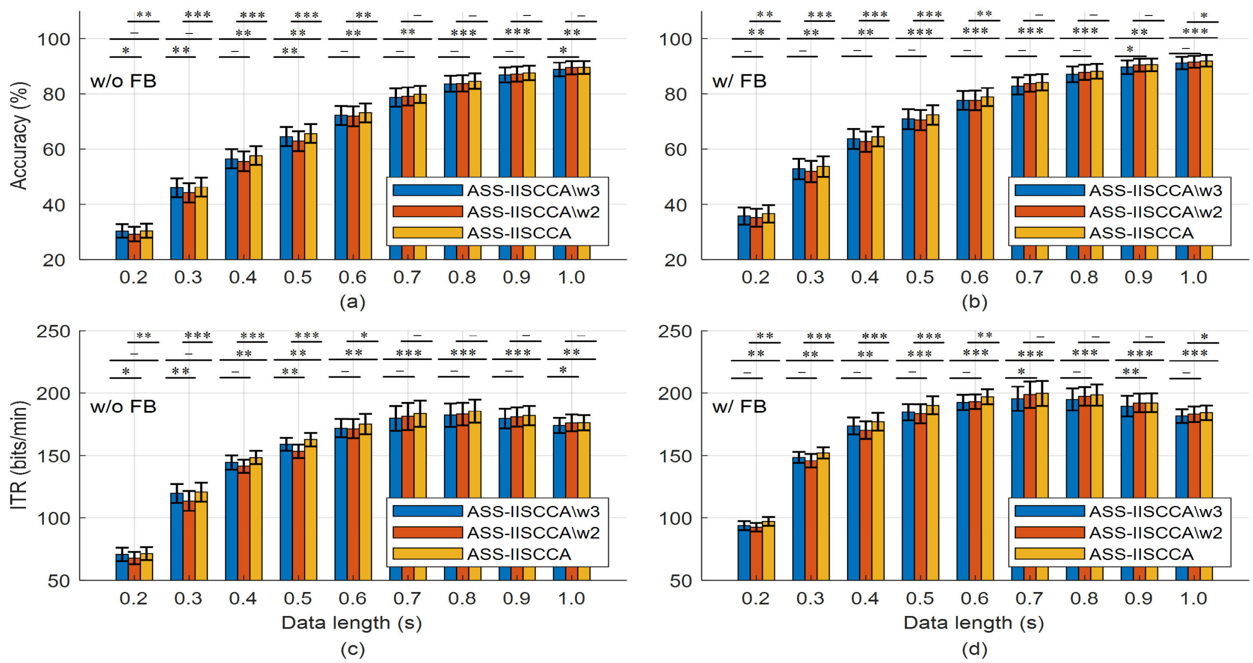

Fig. 8. Averaged accuracies across all 35 subjects yielded by the proposed algorithm ASS-IISCCA and the modified algorithms derived from removing either the filter w2 or w3 from ASS-IISCCA under the two conditions. The two signs '\w2' and '\w3' indicate removing w2 and w3 from ASS-IISCCA respectively.

is still applicable to online experiments. On the other hand, applying parallel computation or object-oriented language like C++ can accelerate this algorithm so that its strict online experiment might be implemented.

F. Comparison With Widely Used TL Algorithms

In order to further verify the performance of the proposed algorithm, we compared it with several widely used TL algorithms. They are combined tCCA (comb-tCCA) [20], transfer template CCA (ttCCA) [19] and least square transformation (LST)-based CCA [21], [32]. Fig. 9 illustrates the averaged classification accuracies and ITRs across all 35 subjects of these algorithms at nine different data lengths under the two conditions. For the sake of comparison, the performance of TRCA is also included in the figure. It is clearly seen that the accuracies and ITRs of both ASS-IISCCA and TRCA are much higher than those of the three most widely used TL algorithms.

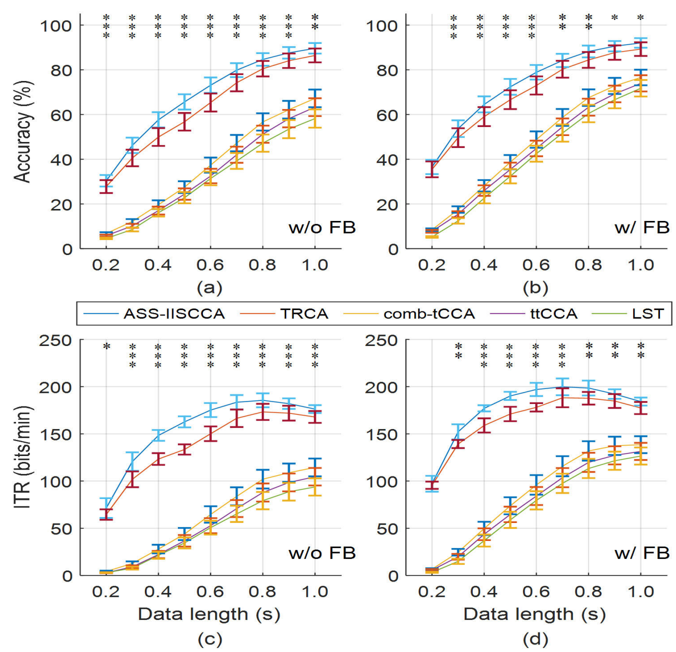

Fig. 9. Averaged classification accuracies and ITRs across all 35 subjects of ASS-IISCCA, TRCA and three widely used TL algorithms at 9 different data lengths under condition 1 (w/o FB) and condition 2 (w/ FB). The pared t-tests are applied for statistical analysis between ASS-IISCCA and TRCA.

G. Classification Results on a Larger Data Set

To verify the robustness of the proposed algorithm, we used it to classify a larger SSVEP data set, the BETA data set [33], which contains 70 subjects participating in four blocks of a cued-spelling task on a 5 × 10 visual board matrix. The other differences between the BETA data set and the Benchmark data set used in this study are that the former was acquired without electromagnetic shielding and its stimulation duration was either 2s or 3s, whereas the latter was acquired with electromagnetic shielding and its stimulation duration was 5s. Thereby, the BETA data set is the one recorded in a more realistic setting. The averaged classification accuracies and ITRs across all 70 subjects of ASS-IISCCA and TRCA at nine data lengths under two conditions are illustrated in Fig. 10. The number of channels and the number of training blocks are fixed at 9 and 3 respectively. Under condition 1, the accuracies and ITRs of ASS-IISCCA were significantly higher than those of TRCA at all data lengths; Under condition 2, the accuracies and ITRs of ASS-IISCCA were significantly higher than those

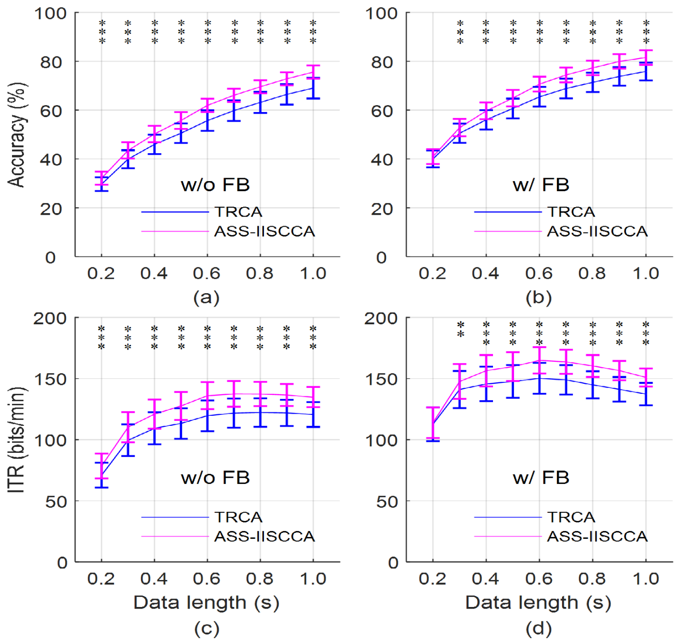

Fig. 10. Averaged classification accuracies and ITRs across all 70 subjects of ASS-ISSCCA and TRCA at nine different data lengths under condition 1 (w/o FB) and condition 2 (w/ FB) achieved on the BETA data set.

of TRCA at all data lengths except for 0.2s. These results fully demonstrate that the proposed algorithm has not only high performance, but also high robustness.

V. Discussion

Most existing BCIs require a lengthy calibration session to collect training data for building user-specific classification models, which seriously affects user experience and the utility of BCIs. In addition, long calibration sessions easily fatigue the users and distract them, thus reducing their performance in testing sessions. TL is a potential method for addressing the problem of shortage of labeled data from the target subject. In this study, we propose a TL-based classification framework for enhancing the performance of SSVEP-BCIs or reducing their calibration time. The framework consists of two main modules: One is the IISCCA algorithm, which is employed to integrate both subject-specific and subject-independent knowledge for frequency recognition and the other is the ASS algorithm, which is used to pick out similar source subjects to the target subject, in order to reduce the variability between subjects.

The main idea of IISCCA is to create three spatial filters with knowledge from within and across subjects and two template signals with knowledge from the target subject and the set of transferred subjects. It works well because the six correlation coefficients forming the feature signal used for frequency recognition are calculated by different combinations of filters and templates, both of which are derived from different domains, and consequently are independent of each other and complementary. Integrating the six coefficients in an intelligent way is conducive to improving the effect of frequency detection. Therefore, it is the complementarity of the six coefficients and their effective integration that enhance the classification performance of this algorithm.

For the group of all 35 subjects, the classification accuracies achieved by ASS-IISCCA were significantly higher than or equal to those achieved by TRCA at all data lengths, all numbers of channels and all numbers of training blocks under the two conditions, as shown in Fig. 2, 3 and 4 respectively. For the data set with fixed number of targets, higher ITRs yielded by ASS-IISCCA were attributed to its higher accuracies and/or the shorter data lengths used for target recognition. At the data length of 0.7s, 9 channels and 3 training blocks, ASS-IISCCA yielded significantly higher accuracies and ITRs than TRCA for all three groups of subjects under condition 1 and for the two groups of all 35 subjects and 27 naive subjects under condition 2, as shown in Fig. 5. This feature contributes to the real-world application of a BCI and its popularization. Thereby, ASS-IISCCA is a superior algorithm for SSVEP-BCIs.

In cases where the number of targets is fixed (40 targets in the data set used in the study), the performance of a BCI system is mainly decided by data length, the number of channels and the number of training blocks. Therefore, we investigate the influences of these parameters on accuracy and ITR. In general, accuracy increases with these three parameters because the larger these parameters, the more accurate the classification model is. ITR, however, is not the case because ITR is inversely proportional to the detection time, which includes gaze shift time and target fixation time (i.e., the data length used for target detection). Thereby, the ITR achieves a maximum value at a particular data length as shown in Fig. 2 and Fig. 10. The number of channels also affects the convenience of BCI usage. For both performance and convenience, usually the nine channels located in the occipital region of the brain are used to establish an SSVEP BCI. As mentioned above, the number of training blocks (i.e., the calibration session) should be minimized or suppressed. As shown in Fig. 4, ASS-IISCCA and TRCA achieved a similar accuracy (84.73% vs 84.19%) under condition 2 with 3 and 4 training blocks respectively. ASS-IISCCA reduced the calibration time by 1/4. According to this study, 0.7s, 9 channels and 3 training blocks can be considered as a good compromise among these parameters. To promote the online performance of the SSVEP BCI, the two spatial filters w_1 and w_3 and the template $\bar{X}_t$ derived from offline analysis could be updated in real time with new data from the target subject. Furthermore, a dynamic stopping strategy [34] could be used to tackle the issue of individual discrepancy in data length and thus enhance the ITR of an SSVEP BCI. Future applications of the proposed classification framework-based SSVEP BCI could be the control of word input systems, wheelchairs, and drones.

Compared to other neurophysiological signals such as those derived from motor imagery and P300 event-related potential, SSVEP signals have higher signal-to-noise ratio (SNR), which makes it possible to develop a BCI system with higher ITRs. SSVEP-based BCI is expected to be the first one that supports a wide range of daily applications. Over the past decades, continuous efforts have been devoted to the goal and the focus of SSVEP-BCI research is placed on the improvement of its performance and practicality. To this end, three recent studies deserve special attention. Liu et al. proposed an algorithm task-discriminant component analysis (TDCA) to enhance the performance of individually calibrated SSVEP-BCI [35]. TDCA estimates a common spatial filter for all stimuli by learning projection directions that are shared by all classes of data and using the temporal information embedded in SSVEP. The former overcomes the problem of the redundancy of spatial filters existed in other algorithms like TRCA, whereas the latter further increases the discriminability of EEG signals from different classes. An offline and online experiment showed that TDCA outperformed the ensemble TRCA and thus validated the effectiveness of the algorithm. Chen et al. presented a spectrally-dense joint frequency-phase modulation (sJFPM) method to extend the number of encodable stimuli [36]. Using the frequency band of 12 Hz-23.9 Hz and the frequency interval of 0.1 Hz, in conjunction with the initial phase of 0 and phase interval of 0.35π, sJFPM achieved the encoding for as many as 120 stimuli. The encoding method evaluated by offline and online experiments indicated its feasibility and effectiveness. The encoding method enables the BCI system to meet complex applications requiring a large number of buttons like a word input system. Liu et al. proposed a TL framework, named align and pool for EEG headset domain adaptation (ALPHA), to facilitate dry electrode-based SSVEP-BCI [37]. By exploiting auxiliary individual wet-electrode EEG data, ALPHA aligns the spatial pattern and the covariance of dry-electrode EEG data. The proposed framework was evaluated on 75 subjects with a 12-target SSVEP-BCI. The results demonstrated that ALPHA significantly outperformed a baseline approach and other two competing TL approaches. ALPHA has methodological and practical implications and pushes the dry electrode-based SSVEP-BCI toward real-world applications.

The limitation of the proposed algorithm is that the accuracy-based algorithm for subject selection (ASS) requires a small amount of labeled data from the target subject and its running time is relatively long. In the future, we will explore more efficient methods for subject selection. The present study only carried out an offline analysis of the proposed algorithm. Future study will focus on its online applications.

VI. Conclusion

In this paper, we proposed a transfer learning framework for frequency recognition of steady-state visual evoked potential (SSVEP) based brain-computer interface (BCI), which consists in an accuracy-based algorithm for the selection of source subjects and an intra- and inter-subject canonical correlation analysis (IISCCA) algorithm for feature extraction. For each stimulus frequency, three spatial filters are created by IISCCA with EEG data from the target subject, the set of selected source subjects and both, and two template signals are estimated by EEG data from the former two respectively. Six coefficients are first extracted by Pearson correlation between the filtered testing signal and each filtered template signal, then integrated into a feature signal and finally classified by a template matching method. The proposed framework was applied to a benchmark data set and evaluated by classification accuracy and information transfer rate. The results suggested that the proposed algorithm outperforms the state-of-the-art algorithm task related component analysis (TRCA) and can significantly decrease the training time of SSVEP-based BCIs.

References

[1] M. Middendorf, G. Mcmillan, G. Calhoun, and K. S. Jones, "Brain–computer interfaces based on the steady-state visual-evoked response," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 211–214, Jun. 2000.

[2] F.-B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives," *Prog. Neurobiol.*, vol. 90, no. 4, pp. 418–438, Apr. 2010.

[3] S. Gao, Y. Wang, X. Gao, and B. Hong, "Visual and auditory brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1436–1447, May 2014.

[4] H. Cecotti, "Adaptive time segment analysis for steady-state visual evoked potential based brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 3, pp. 552–560, Mar. 2020.

[5] G. R. Müller-Putz and G. Pfurtscheller, "Control of an electrical prosthesis with an SSVEP-based BCI," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 1, pp. 361–364, Jan. 2008.

[6] O. Friman, I. Volosyak, and A. Graser, "Multiple channel detection of steady-state visual evoked potentials for brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 742–750, Apr. 2007.

[7] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2610–2614, Dec. 2006.

[8] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method," *J. Neural Eng.*, vol. 6, no. 4, Aug. 2009, Art. no. 046002.

[9] J. Pan, X. Gao, F. Duan, Z. Yan, and S. Gao, "Enhancing the classification accuracy of steady-state visual evoked potential-based brain–computer interfaces using phase constrained canonical correlation analysis," *J. Neural Eng.*, vol. 8, no. 3, Jun. 2011, Art. no. 036027.

[10] Y. Zhang, G. Zhou, J. Jin, M. Wang, X. Wang, and A. Cichocki, "L1-regularized multiway canonical correlation analysis for SSVEP-based BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 6, pp. 887–896, Nov. 2013.

[11] Y. Wang, M. Nakanishi, Y. Wang, and T. Jung, "Enhancing detection of steady-state visual evoked potentials using individual training data," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Chicago, IL, USA Aug. 2014, pp. 3037–3040.

[12] P. Poryzala and A. Materka, "Cluster analysis of CCA coefficients for robust detection of the asynchronous SSVEPs in brain–computer interfaces," *Biomed. Signal Process. Control*, vol. 10, pp. 201–208, Mar. 2014.

[13] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis," *Int. J. Neural Syst.*, vol. 24, no. 4, Jun. 2014, Art. no. 1450013.

[14] M. Nakanishi, Y. Wang, Y.-T. Wang, and T.-P. Jung, "A comparison study of canonical correlation analysis based methods for detecting steady-state visual evoked potentials," *PLoS ONE*, vol. 10, no. 10, Oct. 2015, Art. no. e0140703.

[15] Q. Wei, S. Zhu, Y. Wang, X. Gao, H. Guo, and X. Wu, "A training data-driven canonical correlation analysis algorithm for designing spatial filters to enhance performance of SSVEP-based BCIs," *Int. J. Neural Syst.*, vol. 30, no. 5, May 2020, Art. no. 2050020.

[16] Y. Zhang et al., "Two-stage frequency recognition method based on correlated component analysis for SSVEP-based BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 7, pp. 1314–1323, Jul. 2018.

[17] M. Nakanishi, Y. Wang, X. Chen, Y. Wang, X. Gao, and T. Jung, "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 1, pp. 104–112, Jan. 2018.

[18] R. Zerafa, T. Camilleri, O. Falzon, and K. P. Camilleri, "To train or not to train? A survey on training of feature extraction methods for SSVEP-based BCIs," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 051001.

[19] P. Yuan, X. Chen, Y. Wang, X. Gao, and S. Gao, "Enhancing performances of SSVEP-based brain–computer interfaces via exploiting inter-subject information," *J. Neural Eng.*, vol. 12, no. 4, Aug. 2015, Art. no. 046006.

[20] N. R. Waytowich, J. Faller, J. O. Garcia, J. M. Vettel, and P. Sajda, "Unsupervised adaptive transfer learning for steady-state visual evoked potential brain-computer interfaces," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Budapest, Hungary, Oct. 2016, pp. 004135–004140.

[21] K. Chiang, C. Wei, M. Nakanishi, and T. Jung, "Cross-subject transfer learning improves the practicality of real-world applications of brain–computer interfaces," in *Proc. 9th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, San Francisco, CA, USA, Mar. 2019, pp. 424–427.

[22] H. Wang et al., "Cross-subject assistance: Inter- and intra-subject maximal correlation for enhancing the performance of SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 517–526, 2021.

[23] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, nos. 3–4, pp. 321–377, Dec. 1936.

[24] O. Friman, J. Cedefamn, P. Lundberg, M. Borga, and H. Knutsson, "Detection of neural activity in functional MRI using canonical correlation analysis," *Magn. Reson. Med.*, vol. 45, no. 2, pp. 323–330, Feb. 2001.

[25] X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface," *J. Neural Eng.*, vol. 12, no. 4, Aug. 2015, Art. no. 046008.

[26] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain–computer interface," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 44, Nov. 2015.

[27] F. Lotte and C. Guan, "Learning from other subjects helps reducing brain-computer interface calibration time," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Dallas, TX, USA, Mar. 2010, pp. 614–617.

[28] J. Giles, K. K. Ang, L. S. Mihaylova, and M. Arvaneh, "A subject-to-subject transfer learning framework based on Jensen–Shannon divergence for improving brain–computer interface," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Brighton, U.K., May 2019, pp. 3087–3091.

[29] A. M. Azab, L. Mihaylova, K. K. Ang, and M. Arvaneh, "Weighted transfer learning for improving motor imagery-based brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 7, pp. 1352–1359, Jul. 2019.

[30] Y. Wang, X. Chen, X. Gao, and S. Gao, "A benchmark dataset for SSVEP-based brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1746–1752, Oct. 2017.

[31] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.

[32] M. Nakanishi, Y. Wang, C. Wei, K. Chiang, and T. Jung, "Facilitating calibration in high-speed BCI spellers via leveraging cross-device shared latent responses," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 4, pp. 1105–1113, Apr. 2020.

[33] B. Liu, X. Huang, Y. Wang, X. Chen, and X. Gao, "BETA: A large benchmark database toward SSVEP-BCI application," *Frontiers Neurosci.*, vol. 14, p. 627, Jun. 2020.

[34] E. Yin, Z. Zhou, J. Jiang, Y. Yu, and D. Hu, "A dynamically optimized SSVEP brain–computer interface (BCI) speller," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1447–1456, Jun. 2015.

[35] B. Liu, X. Chen, N. Shi, Y. Wang, S. Gao, and X. Gao, "Improving the performance of individually calibrated SSVEP-BCI by task-discriminant component analysis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1998–2007, 2021.

[36] X. Chen, B. Liu, Y. Wang, and X. Gao, "A spectrally-dense encoding method for designing a high-speed SSVEP-BCI with 120 stimuli," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 2764–2772, 2022.

[37] B. Liu, X. Chen, X. Li, Y. Wang, X. Gao, and S. Gao, "Align and pool for EEG headset domain adaptation (ALPHA) to facilitate dry electrode based SSVEP-BCI," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 2, pp. 795–806, Feb. 2022.