Improving the Performance of Individually Calibrated SSVEP-BCI by Task-Discriminant Component Analysis

Bingchuan Liu, *Student Member, IEEE*, Xiaogang Chen, *Member, IEEE*, Nanlin Shi, *Student Member, IEEE*, Yijun Wang, *Member, IEEE*, Shangkai Gao, *Life Fellow, IEEE*, and Xiaorong Gao, *Member, IEEE*

***Abstract*— A brain-computer interface (BCI) provides a direct communication channel between a brain and an external device. Steady-state visual evoked potential based BCI (SSVEP-BCI) has received increasing attention due to its high information transfer rate, which is accomplished by individual calibration for frequency recognition. Task-related component analysis (TRCA) is a recent and state-of-the-art method for individually calibrated SSVEP-BCIs. However, in TRCA, the spatial filter learned from each stimulus may be redundant and temporal information is not fully utilized. To address this issue, this paper proposes a novel method, i.e., task-discriminant component analysis (TDCA), to further improve the performance of individually calibrated SSVEP-BCI. The performance of TDCA was evaluated by two publicly available benchmark datasets, and the results demonstrated that TDCA outperformed ensemble TRCA and other competing methods by a significant margin. An offline and online experiment testing 12 subjects further validated the effectiveness of TDCA. The present study provides a new perspective for designing decoding methods in individually calibrated SSVEP-BCI and presents insight for its implementation in high-speed brain speller applications.**

Manuscript received July 14, 2021; revised September 11, 2021; accepted September 17, 2021. Date of publication September 20, 2021; date of current version October 1, 2021. This work was supported in part by the Key-Area Research and Development Program of Guangdong Province under Grant 2018B030339001, in part by the Doctoral Brain+X Seed Grant Program of Tsinghua University, in part by the Strategic Priority Research Program of the Chinese Academy of Sciences under Grant XDB32040200, in part by Beijing Science and Technology Program under Grant Z201100004420015, in part by the National Natural Science Foundation of China under Grant 62171473 and Grant 61431007, and in part by the National Key Research and Development Program of China under Grant 2017YFB1002505. *(Corresponding author: Xiaorong Gao.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board of Tsinghua University under Application No. 20200020.

Bingchuan Liu, Nanlin Shi, Shangkai Gao, and Xiaorong Gao are with the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China (e-mail: louislbc@hotmail.com; snl18@mails.tsinghua.edu.cn; gsk-dea@tsinghua.edu.cn; gxr-dea@mail.tsinghua.edu.cn).

Xiaogang Chen is with the Institute of Biomedical Engineering, Chinese Academy of Medical Sciences and Peking Union Medical College, Tianjin 300192, China (e-mail: chenxg@bme.cams.cn).

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

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSRE.2021.3114340, provided by the authors.

Digital Object Identifier 10.1109/TNSRE.2021.3114340

***Index Terms*— Brain-computer interface (BCI), steady-state visual evoked potential (SSVEP), electroencephalography (EEG), task-related component analysis (TRCA), task-discriminant component analysis (TDCA).**

I. Introduction

A BRAIN-COMPUTER interface (BCI) enables people to interface and interact with the outside world by leveraging brain signals related to sensation, perception or high-level cognitive activities [1]. BCI technology generally falls into the categories of invasive and noninvasive paradigms. For the noninvasive paradigm, steady-state visual evoked potential based BCI (SSVEP-BCI) [2] has enjoyed wide-spread adoption due to its advantages in high information transfer rate (ITR), low cost, and ease of use. The merit of high ITR is attributable to the high signal-to-noise ratio (SNR) profile of SSVEP, which is a frequency-tagged occipital brain signal that can be evoked by flickers, moving gratings, and reversible checkerboards. The relatively high performance of SSVEP-BCI helps with the development of practical applications related to dialing [3] and spelling [2], as well as other end-user applications, e.g., wheelchair control [4], and smart home applications [5].

The past decades have witnessed rapid progress in the development of frequency recognition methods for improving the performance of SSVEP-BCI. Based on whether the supervision of calibration or training data are required, frequency recognition methods can be categorized into training-free methods, and supervised or training-based methods. Training-free methods recognize stimulus frequencies in a plug-and-play manner, including canonical correlation analysis (CCA) [6] and its filter-bank extension [2], minimum energy combination (MEC) [7], multivariate synchronization index (MSI) [8], and Ramanujan periodicity transforms (RPT) [9], etc. However, training-free methods work at the expense of ITR and long stimulation duration. To circumvent the problems, training-based methods conventionally leverage individual calibration data to construct individual templates and spatial filters, which yield a substantial improvement in ITR and a reduction in the required data length. In general, the literature reports two main categories of training-based methods, i.e., CCA-derived methods and task-related component analysis (TRCA)-derived methods. Specifically, CCA-derived

methods optimize spatial filters by maximizing the correlation coefficient in the projection subspace, and the orthogonal subspace corresponding to the maximal correlation coefficient is used for frequency recognition. This line of work, to name a few, the extended CCA [10], modified extended CCA (m-extended CCA) [2] and L1-regularized multiway CCA (L1MCCA) [11] usually utilize a sine-cosine reference signal. Distinct from these CCA-derived methods, TRCA [12] is a recently proposed and state-of-the-art method that maximizes the reproducibility between trials in SSVEP-BCI. Follow-up studies of TRCA facilitate detection by exploiting additional information, e.g., neighboring stimuli [13], reference signals [14], or similarity constraints [15].

The motivation of the present study derives from the redundancy of spatial filters in TRCA. Previous studies imply that different classes of stimuli share a common spatial pattern or spatial filter in SSVEP-BCI [13], [16], [17]. Thus, to enhance performance, TRCA usually employs an ensemble technique by concatenating the spatial filters from each class. However, the concatenation of spatial filters yields redundancy, and the mechanism of how to determine the optimal redundancy is not well understood to date. A variant of TRCA [13] further increased this redundancy by learning across neighboring stimuli, which improved the ITR. But, the issue of the ensemble technique has yet been addressed. As the spatial patterns are similar across stimuli, learning the spatial filter from each stimulus is not necessary. Therefore, this study proposes to learn projection directions that are shared by all classes of data. Different from the TRCA, which is based on a generative model, the proposed method tackles the problem of frequency recognition with a discriminative model. Furthermore, it is postulated that the temporal information of SSVEP is not fully utilized and has the potential to benefit frequency recognition. Based on these intuitions, a task-discriminant component analysis (TDCA) method is proposed for individually calibrated SSVEP-BCI in this study. The effectiveness of the proposed method is verified by two benchmark datasets, an offline experiment and an online experiment.

The remainder of this paper is structured as follows: Section 2 introduces an overview of TRCA and the proposed TDCA, along with the datasets, experimental setup, performance evaluation, and feature evaluation in the technical validation. The result of the technical validation is presented in Section 3. The methodological and practical benefits of TDCA are discussed in Section 4. And finally this paper is concluded in Section 5.

II. Methods

A. Task-Related Component Analysis

TRCA [12] is a recently developed state-of-the-art method to boost the performance of individually calibrated SSVEP-BCI at short data lengths. The intuition of TRCA is to maximize the reproducibility of task-related components after spatial filtering. The underlying assumption of TRCA is a generative model and its signal model is given as follows:

$$x_j(t) = a_{1,j}s(t) + a_{2,j}n(t), \quad j = 1, 2, \ldots, N_{ch} \tag{1}$$

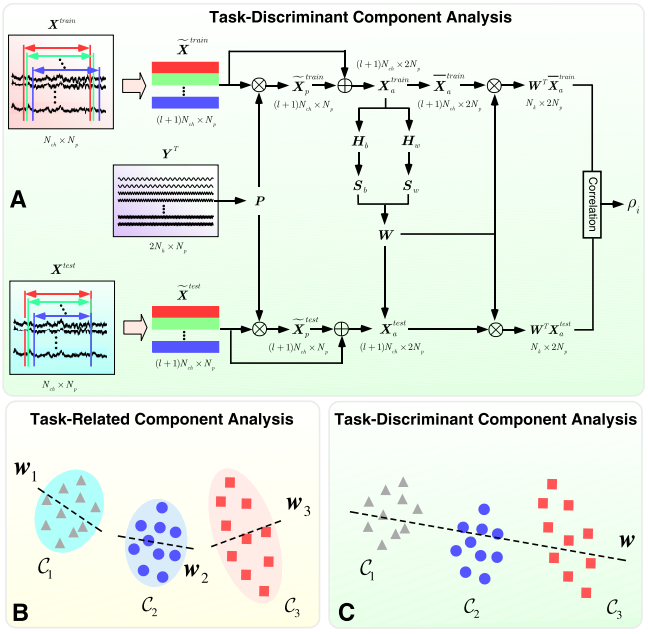

Fig. 1. Flow chart of task-discriminant component analysis (TDCA) (A) and a schematic comparison of the difference between task-related component analysis (TRCA) and TDCA (B and C). In the flow chart (A), the augmented data for both the training sample and test sample are initially formed by temporally delayed copies. Then, the data are projected onto the subspace spanned by a sinusoidal reference signal, and finally, a discriminant analysis is performed to derive spatio-temporal filters. The detection statistics are obtained by template matching. In the comparison (B and C), the triangles, dots, and squares denote the single-trial SSVEP samples from different classes. For illustrative purposes, only three classes are shown. The shaded areas covering the samples in B denotes the probability distributions the generative model assumes. The dashed line denotes the spatial filter or projection direction. Different from the generative model of TRCA, TDCA utilizes a discriminative model to learn the projection direction for all classes.

where $x(t)$ is the observation, $s(t)$ is the task-related component, and $n(t)$ is the task-unrelated component. $a_{1,j}$ and $a_{2,j}$ are mixing coefficients for task-related components. N_{ch} is the number of channels. $s(t)$ can be estimated by maximizing the inter-trial covariance after spatial filtering:

$$\boldsymbol{w}^T \boldsymbol{S}\boldsymbol{w} = \sum_{\substack{h_1,h_2=1 \\ h_1 \neq h_2}}^{N_t} \sum_{j_1,j_2=1}^{N_{ch}} w_{j_1} w_{j_2} \mathrm{Cov}\left(x_{j_1}^{(h_1)}(t), x_{j_2}^{(h_2)}(t)\right) \tag{2}$$

To solve the optimization problem, a quadratic constraint of finite variance is added

$$\boldsymbol{w}^T \boldsymbol{Q}\boldsymbol{w} = \sum_{j_1,j_2=1}^{N_{ch}} w_{j_1} w_{j_2} \mathrm{Cov}\left(x_{j_1}(t), x_{j_2}(t)\right) \tag{3}$$

Then, the TRCA spatial filter can be derived from the generalized Rayleigh quotient problem

$$\widehat{\boldsymbol{w}} = \arg\max_{\boldsymbol{w}} \frac{\boldsymbol{w}^T \boldsymbol{S}\boldsymbol{w}}{\boldsymbol{w}^T \boldsymbol{Q}\boldsymbol{w}} \tag{4}$$

It is worth noting that ensemble TRCA employs an ensemble technique to constitute the final spatial filter:

$$\boldsymbol{W}^{(m)} = \left[\boldsymbol{w}_1^{(m)}, \boldsymbol{w}_2^{(m)}, \cdots, \boldsymbol{w}_{N_c}^{(m)}\right] \tag{5}$$

where m denotes the filter bank in [18] and N_c denotes the number of classes.

B. Task-Discriminant Component Analysis

The flow chart of TDCA is illustrated in Figure 1 **A**. Figure 1 **B** and **C** illustrate the distinction between TRCA and the proposed TDCA in terms of underlying assumptions and spatial filters. Different from the generative model of TRCA, this study utilizes a discriminative model. In contrast to the strategy that optimizes the spatial filter for each class, here, discriminant analysis is employed to seek the spatio-temporal filter with respect to all classes. The details of the proposed TDCA are presented in the followings.

In an individually calibrated SSVEP-BCI, consider $\boldsymbol{X}^{(i)} \in \mathbb{R}^{N_{ch}\times N_p}$ is the i-th training trial, $i = 1, 2, \cdots, N_t$, where N_t is the number of trials, and N_p is the number of sampling point. For each training trial, the dimensionality of the EEG data is elevated first, as follows:

$$\widetilde{\boldsymbol{X}} = \left[\boldsymbol{X}^T, \boldsymbol{X}_1^T, \cdots, \boldsymbol{X}_l^T\right]^T \tag{6}$$

where $\widetilde{\boldsymbol{X}} \in \mathbb{R}^{(l+1)N_{ch}\times N_p}$ is the augmented EEG trial. $\boldsymbol{X}_l \in \mathbb{R}^{N_{ch}\times N_p}$ denotes the EEG trial delayed by l points, i.e., the data copy from time (in data points) $l+1$ to time $N_p + l$. In a similar manner, this procedure is applied to the test trial. To guarantee that data points greater than N_p are not tested, different from the training trial, data points exceeding N_p are padded with zeros.

$$\boldsymbol{X}_l = \left[\boldsymbol{X}_l', \mathbf{O}^{N_{ch}\times l}\right] \tag{7}$$

where $\boldsymbol{X}_l' \in \mathbb{R}^{N_{ch}\times(N_p-l)}$ denotes the data copy from time $l+1$ to time N_p. The augmented EEG trial is then projected onto the subspace spanned by the reference signal.

$$\widetilde{\boldsymbol{X}}_p = \widetilde{\boldsymbol{X}}\boldsymbol{P}_i \tag{8}$$

Here $\boldsymbol{P}$ is an orthogonal projection matrix of the i-th class, which is formed by [7], [14]:

$$\boldsymbol{P}_i = \boldsymbol{Q}\boldsymbol{Q}^T \tag{9}$$

$\boldsymbol{Q}$ derives from the QR decomposition of the sine-cosine reference signal $\boldsymbol{Y}$ [6], [19], which corresponds to the i-th stimulus frequency f_i.

$$\boldsymbol{Y}_i = \boldsymbol{Q}\boldsymbol{R} \tag{10}$$

$$\boldsymbol{Y}_i = \begin{bmatrix} \sin\left(2\pi f_i \boldsymbol{t}^T\right) \\ \cos\left(2\pi f_i \boldsymbol{t}^T\right) \\ \vdots \\ \sin\left(2\pi N_h f_i \boldsymbol{t}^T\right) \\ \cos\left(2\pi N_h f_i \boldsymbol{t}^T\right) \end{bmatrix}^T, \quad \boldsymbol{t} = [1/f_s, \cdots, N_p/f_s]^T \tag{11}$$

where N_h is the number of harmonics, and f_s is the sampling rate.

For training and test trials, a further secondary augmented EEG trial is constructed.

$$\boldsymbol{X}_a = \left[\widetilde{\boldsymbol{X}}, \widetilde{\boldsymbol{X}}_p\right] \tag{12}$$

Then, for training trials, two-dimensional linear discriminant analysis [20]–[22] is performed to seek projection directions to discriminate trials from all classes. The between-class difference matrix $\boldsymbol{H}_b \in \mathbb{R}^{(l+1)N_{ch}\times 2N_c N_p}$ and the within-class difference matrix $\boldsymbol{H}_w \in \mathbb{R}^{(l+1)N_{ch}\times 2N_t N_p}$ are defined as

$$\begin{aligned} \boldsymbol{H}_b &= \frac{1}{\sqrt{N_c}}[\bar{\boldsymbol{X}}_a^1 - \bar{\boldsymbol{X}}_a^a, \cdots, \bar{\boldsymbol{X}}_a^{N_c} - \bar{\boldsymbol{X}}_a^a] \\ \boldsymbol{H}_w &= \frac{1}{\sqrt{N_t}}[\boldsymbol{X}_a^{(1)} - \bar{\boldsymbol{X}}_a^{(1)}, \cdots, \boldsymbol{X}_a^{(N_t)} - \bar{\boldsymbol{X}}_a^{(N_t)}] \end{aligned} \tag{13}$$

where $\bar{\boldsymbol{X}}^i$ and $\bar{\boldsymbol{X}}^{(i)}$ denote the two-dimensional class centers of the i-th class and the i-th sample, respectively. The superscript a denotes all classes, and $\bar{\boldsymbol{X}}_a^a$ is obtained by Eq. (14):

$$\bar{\boldsymbol{X}}_a^a = \frac{1}{N_t}\sum_{i=1}^{N_t}\boldsymbol{X}_a^{(i)} \tag{14}$$

Then, the following Fisher criterion is used to derive the projection directions

$$\underset{\boldsymbol{W}}{\text{maximize}} \ \frac{tr(\boldsymbol{W}^T\boldsymbol{S}_b\boldsymbol{W})}{tr(\boldsymbol{W}^T\boldsymbol{S}_w\boldsymbol{W})} \tag{15}$$

where the scatter matrices $\boldsymbol{S}_b$ and $\boldsymbol{S}_w$ have the form

$$\begin{aligned} \boldsymbol{S}_b &= \boldsymbol{H}_b\boldsymbol{H}_b^T \\ \boldsymbol{S}_w &= \boldsymbol{H}_w\boldsymbol{H}_w^T \end{aligned} \tag{16}$$

By utilizing the idempotent property of the projection matrix, i.e., $\boldsymbol{P}^2 = \boldsymbol{P}$, the Eq. (15) can be written as

$$\underset{\boldsymbol{W}}{\text{maximize}} \ \frac{tr\left[\boldsymbol{W}^T\boldsymbol{H}_b'(\boldsymbol{P}_b + \mathbf{I})\boldsymbol{H}_b'^T\boldsymbol{W}\right]}{tr\left[\boldsymbol{W}^T\boldsymbol{H}_w'(\boldsymbol{P}_w + \mathbf{I})\boldsymbol{H}_w'^T\boldsymbol{W}\right]} \tag{17}$$

where $\boldsymbol{H}_b'$ and $\boldsymbol{H}_w'$ are formed by $\widetilde{\boldsymbol{X}}$ in Eq. (13). For the projection matrices, $\boldsymbol{P}_b = \oplus_{i=1}^{N_c}\boldsymbol{P}_i$ and $\boldsymbol{P}_w = \oplus_{i=1}^{N_c}\oplus_{j=1}^{N_b}\boldsymbol{P}_i$, where $\oplus$ denotes the direct sum and N_b denotes the number of blocks.

C. Evaluation

1) Public Datasets: The performance of TDCA was initially evaluated on two public SSVEP-BCI datasets, i.e., the Benchmark dataset [23] and the BETA dataset [24]. For the Benchmark dataset, 35 subjects participated in six blocks of a cued-spelling task on a 5×8 matrix of a virtual keyboard. For the BETA dataset, 70 subjects participated in four blocks of a cued-spelling task on a QWERTY virtual keyboard. Both datasets contained 40 targets, including 10 digits, 26 letters of the English alphabets, and four non-alphanumeric signs. Each target was encoded by one of the stimulus frequencies ranging from 8 Hz to 15.8 Hz (frequency interval: 0.2 Hz) using a joint frequency and phase modulation (JFPM) method [2]. The stimulation duration was 5 s in the Benchmark dataset and either 2 s or 3 s in the BETA dataset. Sixty-four channels of EEG data were collected by SynAmps2 (Neuroscan Inc.) at a sampling rate of 1000 Hz, which were then downsampled to 250 Hz. The two datasets also differed in the experimental condition, which was with and without electromagnetic shielding for the Benchmark and the BETA dataset, respectively.

Four performance evaluations were conducted on the public datasets. *A) Performance evaluation on classical montage:* The performances of TDCA and ensemble TRCA were compared in a classical montage [2], i.e., nine channels in the

occipital region (Pz, PO3/4, PO5/6, POz, Oz, and O1/2). *B) Performance evaluation on other montages:* In addition to the classical montage, other configurations of montage were evaluated for TDCA and ensemble TRCA. Here, five subsets of montage configuration were evaluated, including central occipital montage ($N_{ch} = 3$, Oz, O1, and O2), classical occipital montage ($N_{ch} = 9$, Pz, POz, PO3/4, PO5/6, Oz and O1/2), occipital montage ($N_{ch} = 21$, Pz, P1/2, P3/4, P5/6, P7/8, POz, PO3/4, PO5/6, PO7/8, Oz, O1/2, and CB1/2), parietal-occipital montage ($N_{ch} = 30$, CPz, CP1/2, CP3/4, CP5/6, TP7/8, Pz, P1/2, P3/4, P5/6, P7/8, POz, PO3/4, PO5/6, PO7/8, Oz, O1/2, and CB1/2), and all channels ($N_{ch} = 64$). *C) Performance evaluation with insufficient training data:* The performance of TDCA and ensemble TRCA were evaluated with a varying number of training blocks (N_b). Specifically, for the Benchmark dataset N_b varied from 1 to 5, while for the BETA dataset N_b varied from 1 to 3. *D) Performance comparison with other methods:* In the same setting as *A)*, TDCA was compared with other competitive frequency recognition methods including multi-stimulus TRCA [13], TRCA with sine-cosine signal (TRCA-R) [14], similarity-constrained TRCA (scTRCA) [15], and extended CCA [10]. For multiple comparisons in *B)*, *C)* and *D)*, a repeated measures analysis of variance (RMANOVA) was conducted. When sphericity was violated assessed by Mauchly's test of sphericity, the adjustment of Greenhouse-Geisser correction was applied. When a significant main effect was found ($p < .05$), post hoc comparisons of t-test were then performed with Bonferroni correction. The statistical analyses were conducted in SPSS Statistics 26 (IBM, Armonk, NY, USA).

In the performance evaluation, a k-fold cross-validation was performed, i.e., $k = 6$ for the Benchmark dataset and $k = 4$ for the BETA dataset. Specifically, for each subject, one block of EEG data was used as test set and the remaining data were used as training set in a fold. As short data length is critical for individually calibrated SSVEP-BCI, a sliding window with data length N_p (ranging from 0.1 s to 1 s with an interval of 0.1 s) was employed to trim the epoch for performance evaluation. The onset of the sliding window was set at $[t_s + d, t_s + d + N_p]$, and t_s was the time point when visual stimulation began. d is the latency estimated from the dataset, with $d = 140$ ms for the Benchmark dataset [23] and $d = 130$ ms for the BETA dataset [24]. The metrics of accuracy and ITR were calculated for each data length. The ITR in bits per min (bpm) is defined as [25]:

$$ITR = 60 \cdot (\log_2 M + P\log_2 P + (1-P)\log_2 \frac{1-P}{M-1})/T \tag{18}$$

where M denotes the number of classes, P denotes the accuracy, and T (in s) denotes the selection time including gaze time and gaze shift time. A gaze shift time of 0.5 s is used for analysis [18], [23].

A template matching approach was employed for frequency recognition in the test set. Different from the scheme applied in ensemble TRCA, more than one subspace (N_k) was used in TDCA. Empirically, $N_k = 8$, $l = 5$ for the Benchmark dataset and $N_k = 9$, $l = 3$ for the BETA dataset. Filter bank analysis was performed for all methods, where the number of filter banks $N_{fb} = 5$ and the combining weights were set according to prior literature [12], [18]. The filtering procedure was performed for each data length. The ensemble technique [12] was employed in TRCA and its derived methods, i.e., msTRCA, TRCA-R, and scTRCA, termed as ensemble TRCA, ensemble msTRCA, ensemble TRCA-R, and ensemble scTRCA, respectively.

The profile of latent features for TDCA and ensemble TRCA was further compared. To characterize the discriminability of classes in the feature space, the R-squared statistics of the correlation coefficients associated with the target and nontarget stimuli were computed [12], [26] using the data of the offline experiment. The R-squared statistics were compared for TDCA and ensemble TRCA at data lengths from 0.1 s to 1 s with an interval of 0.1 s. In addition, the activation pattern [27] was explored to delineate the underlying spatial pattern the model learned from the data. The activation pattern is given by

$$\boldsymbol{A} = \boldsymbol{\Sigma}_{\boldsymbol{x}} \boldsymbol{W} \boldsymbol{\Sigma}_{\hat{\boldsymbol{s}}}^{-1} \tag{19}$$

where $\boldsymbol{W}$ is the spatial filter. $\boldsymbol{\Sigma}_{\boldsymbol{x}}$ and $\boldsymbol{\Sigma}_{\hat{\boldsymbol{s}}}$ are the covariance of the original EEG and the estimated source EEG, respectively. The activation patterns from 64-channel EEG were computed at a data length of 1 s. For TDCA, the activation patterns corresponding to the original EEG ($l = 0$) were computed. To ensure that the polarities of the activation pattern were consistent, the activation patterns of each subject were normalized by the value of the Oz channel. For ensemble TRCA, as the common activation pattern is shared across stimuli [13], [16], [17], the activation patterns were averaged across stimuli. Then, the activation patterns were averaged across subjects for both the Benchmark dataset and the BETA dataset.

2) Offline Experiments: An offline and online experiment was conducted to validate the effectiveness of the proposed method on new subjects. Twelve healthy subjects (six males and six females) with a mean age of 23.1 ± 1.2 (mean ± standard error, range from 18 to 31) participated in the study. All subjects had normal or corrected-to-normal vision. The experiment was approved by the institutional review board of Tsinghua University (NO.20200020), and subjects provided full written consent before the experiment.

For this study, a 40-target virtual speller was designed for offline and online experiments. As illustrated in Figure 2, 40 targets on the virtual speller (i.e., 10 digits, 26 letters of the English alphabet, and four symbols) were encoded as flickers by JFPM [2]. The stimulus frequency of the target ranged from 8 Hz to 15.8 Hz (frequency interval: 0.2 Hz) and the initial phase ranged from 0 to 1.5 π (phase interval: 0.5 π). The virtual speller was presented on a 24.5-inch LED monitor (refresh rate: 60 Hz) using the sampled sinusoidal stimulation method [28]. Note that the layout and encoding parameters differed from the Benchmark and the BETA dataset.

In the offline experiment, subjects performed a cued-spelling task and 64-channel EEG were recorded for offline analysis. In line with the Benchmark dataset, the number of blocks was set to six. In each block, subjects were instructed to direct their attention to the center of each target, which

Fig. 2. Virtual keyboard for the 40-target SSVEP-BCI speller used in both the offline and online experiments. (A) Layout of a character input speller with 10 digits, 26 letters of the English alphabet, and four non-alphanumeric signs. The topmost rectangle is designed to show online feedback. (B) The tagged frequency and initial phase corresponding to each target are encoded by the joint frequency and phase modulation, respectively.

cued for 1 s and flickered for 3 s. During flickering, subjects were asked to avoid blinking and body movement. The order of cued targets was randomized. There was a break of 2 min between two consecutive blocks to avoid visual fatigue.

3) Online Experiment: Following the offline experiment, an online experiment was performed on a different day. Specifically, each subject performed 10 blocks of the cued spelling task, including five blocks of training data and five blocks of test data. In each block, the target cued for 0.5 s and flickered for 0.4 s. For the five blocks of test data, a resulting feedback was presented on the speller after online classification by TDCA during the 0.5-s cue time. The classical montage of nine occipital channels (Pz, POz, PO3/4, PO5/6, Oz and O1/2) were recorded and used to compute the online classification accuracy and ITR.

The EEG data in the offline and online experiments were acquired by a wireless amplifier (Neuracle, China) in an electromagnetic shielding room and triggers were synchronized by a parallel port. An infinite impulse response (IIR) notch filter was applied to remove interference from the power line, and data were downsampled to 250 Hz for the offline and online analyses. The procedure of performance evaluation was in accord with the Benchmark dataset, except for the parameters of $N_k = 7$, and $l = 5$. The visual presentation was developed in MATLAB (MathWorks, Inc.) via Psychophysics Toolbox Version 3 [29].

III. Results

Figure 3 illustrates the performance of average classification accuracy and ITR for TDCA and ensemble TRCA evaluated with public datasets (**A** and **B**: the Benchmark dataset; **C** and **D**: the BETA dataset). The result showed that TDCA outperformed ensemble TRCA at all data lengths. The highest ITR for TDCA was 244.34 ± 10.84 bpm at 0.5 s, while for ensemble TRCA, the highest ITR was achieved at 0.5 s with 219.67 ± 12.47 bpm. Paired t-test revealed that the difference for accuracy between TRCA and TDCA was statistically significant ($p < .05$) for all data lengths, and for data lengths greater than 0.1 s on the ITR. For a data length of 0.5 s, the accuracies for TDCA and ensemble TRCA were 0.850 ± 0.027 and 0.794 ± 0.033, respectively, with a difference of 0.057 ($p < .001$). Specifically, TDCA was advantageous over TRCA for 94.3% of the subjects, and the results for

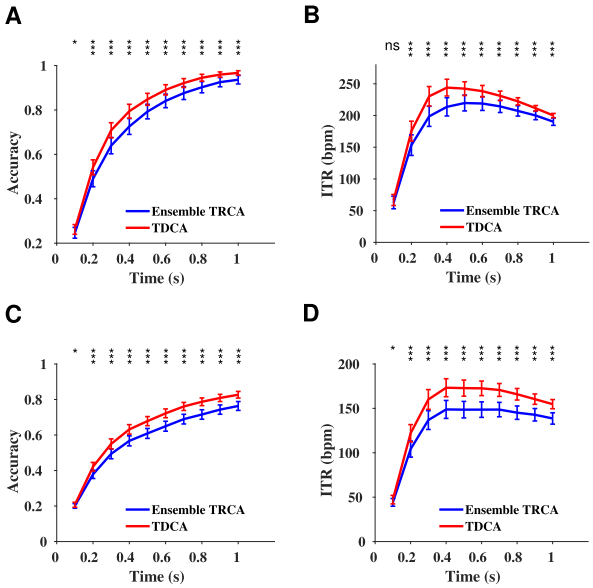

Fig. 3. Average classification accuracy and ITR for TDCA and ensemble TRCA on public datasets. (A) Accuracy on the Benchmark dataset. (B) ITR on the Benchmark dataset. (C) Accuracy on the BETA dataset. (D) ITR on the BETA dataset. Data lengths from 0.1 s to 1 s (with an interval of 0.1 s) are used for evaluation. The asterisks indicate significant differences between TDCA and ensemble TRCA (*$p < .05$, **$p < .01$, ***$p < .001$).

some representative subjects were detailed in the Supplemental Table I. For a data length of 1 s, the accuracies for TDCA and ensemble TRCA were 0.968 ± 0.010 and 0.936 ± 0.019, respectively, with a difference of 0.032 ($p = .002$).

For the BETA dataset, TDCA outperformed ensemble TRCA by a significant margin (c.f. Figure 3 **C** and **D**). The highest ITRs for TDCA and ensemble TRCA were 174.88 ± 10.19 bpm (0.4 s) and 148.81 ± 10.13 bpm (0.4 s), respectively. For data lengths exceeding 0.1 s, paired t-test revealed that the difference between TDCA and ensemble TRCA was statistically significant ($p < .05$), both for the accuracy and ITR. For a data length of 0.5 s, the accuracies for TDCA and ensemble TRCA were 0.680 ± 0.028 and 0.609 ± 0.028, respectively, with a difference of 0.070 ($p < .001$). Specifically, TDCA was advantageous over TRCA for 91.4% of the subjects, and the results for some representative subjects were detailed in the Supplemental Table II. For a data length of 1 s, the accuracies for TDCA

Algorithm 1. TDCA for individually calibrated SSVEP-BCI

Input: The i-th training trial $\boldsymbol{X}^{(i)} \in \mathbb{R}^{N_{ch} \times N_p}$, $i = 1, 2, \cdots, N_t$
Process:
I. Elevate the dimensionality of EEG data.
(a) For training trials,
$$\widetilde{\boldsymbol{X}} = \left[\boldsymbol{X}^T, \boldsymbol{X}_1^T, \cdots, \boldsymbol{X}_l^T\right]^T$$
(b) For test trials,
$$\widetilde{\boldsymbol{X}} = \left[\boldsymbol{X}^T, \boldsymbol{X}_1^T, \cdots, \boldsymbol{X}_l^T\right]^T, \boldsymbol{X}_l = \left[\boldsymbol{X}_l', \mathbf{O}^{N_{ch} \times l}\right]$$
II. Project the augmented EEG trials onto the subspace spanned by the reference signal.
(a) QR decomposition, $\boldsymbol{Y} = \boldsymbol{Q}\boldsymbol{R}$
(b) Form projection matrix, $\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{Q}^T$
(c) Orthogonal projection, $\widetilde{\boldsymbol{X}}_p = \widetilde{\boldsymbol{X}}\boldsymbol{P}$
III. Form secondary augmented EEG trials, $\boldsymbol{X}_a = \left[\widetilde{\boldsymbol{X}}, \widetilde{\boldsymbol{X}}_p\right]$
IV. Seek discriminative projection directions.
(a) Form difference matrices.
· Between-class difference matrix
$$\boldsymbol{H}_b = \frac{1}{\sqrt{N_c}}[\bar{\boldsymbol{X}}_a^1 - \bar{\boldsymbol{X}}_a^a, \cdots, \bar{\boldsymbol{X}}_a^{N_c} - \bar{\boldsymbol{X}}_a^a]$$
· Within-class difference matrix
$$\boldsymbol{H}_w = \frac{1}{\sqrt{N_t}}[\boldsymbol{X}_a^{(1)} - \bar{\boldsymbol{X}}_a{}^{(1)}, \cdots, \boldsymbol{X}_a^{(N_t)} - \bar{\boldsymbol{X}}_a^{(N_t)}]$$
(b) Form scatter matrices
· Between-class scatter matrix $\boldsymbol{S}_b = \boldsymbol{H}_b\boldsymbol{H}_b^T$
· Within-class scatter matrix $\boldsymbol{S}_w = \boldsymbol{H}_w\boldsymbol{H}_w^T$
(c) Derive projection directions from Fisher criterion
$$\underset{\boldsymbol{W}}{\text{maximize}} \quad \frac{tr(\boldsymbol{W}^T\boldsymbol{S}_b\boldsymbol{W})}{tr(\boldsymbol{W}^T\boldsymbol{S}_w\boldsymbol{W})} = \frac{tr\left[\boldsymbol{W}^T\boldsymbol{H}_b'(\boldsymbol{P}_b+\mathbf{I})\boldsymbol{H}_b'^T\boldsymbol{W}\right]}{tr\left[\boldsymbol{W}^T\boldsymbol{H}_w'(\boldsymbol{P}_w+\mathbf{I})\boldsymbol{H}_w'^T\boldsymbol{W}\right]}$$
· where $\boldsymbol{P}_b = \oplus_{i=1}^{N_c}\boldsymbol{P}_i$ and $\boldsymbol{P}_w = \oplus_{i=1}^{N_c}\oplus_{j=1}^{N_b}\boldsymbol{P}_i$
Output: Spatiotemporal filters $\boldsymbol{W}$

and ensemble TRCA were 0.826 ± 0.019 and 0.764 ± 0.025, respectively, with a difference of 0.063 ($p < .001$).

The effects of various montages on TDCA and ensemble TRCA are illustrated in Figure 4 (**A**: the Benchmark dataset; **B**: the BETA dataset). Here, the performance of each method was measured by the maximum average ITR across data lengths. As assessed by two-way (method × montage) repeated measures analysis of variance (RMANOVA), the interaction between the method and montage was statistically significant for both the Benchmark dataset, $F(2.399, 81.556) = 15.335$, $p < .001$, Greenhouse-Geisser corrected, and for the BETA dataset, $F(2.905, 200.441) = 30.097$, $p < .001$, Greenhouse-Geisser corrected. Paired t-test revealed that TDCA significantly outperformed ensemble TRCA on various montages ($p < .001$) for both datasets. It is worth noting that in both TDCA and ensemble TRCA, the classical montage was not the optimal montage that yielded the highest ITR. In TDCA, the optimal montage was parietal-occipital montage for the Benchmark dataset (TDCA: 280.47 ± 10.64 bpm; ensemble TRCA: 239.23 ± 14.52 bpm; $p < .001$), and occipital montage for the BETA dataset (TDCA: 197.72 ± 10.24 bpm; ensemble TRCA: 164.91 ± 10.91 bpm; $p < .001$). When the number of channels exceeded that of the optimal montage, the ITR significantly declined for TDCA and ensemble TRCA, especially for ensemble TRCA.

Figure 5 illustrates the maximum average ITR of TDCA and ensemble TRCA with a varying number of training blocks. Two-way (method × block) RMANOVA identified that there was a statistically significant main effect of blocks for the Benchmark dataset, $F(1.29, 43.859) = 324.455$, $p < .001$, Greenhouse-Geisser corrected, and for the BETA dataset, $F(1.067, 73.629) = 272.329$, $p < .001$, Greenhouse-Geisser corrected. With increasing number of training blocks, the performances of TDCA and ensemble TRCA improved, and TDCA consistently and significantly outperformed ensemble TRCA on both datasets ($p < .001$). For instance, with insufficient training data of three blocks, the performance on the Benchmark dataset was 219.7 ± 8.75 bpm for TDCA and 191.64 ± 11.06 bpm for ensemble TRCA; on the BETA dataset, the performance was 173.31 ± 10.13 bpm for TDCA and 148.81 ± 10.11 bpm for ensemble TRCA.

The performance of various competing methods compared with TDCA is illustrated in Figure 6 (**A**: the Benchmark dataset; **B**: the BETA dataset). Here the performance of each method was also characterized by the maximum average ITR

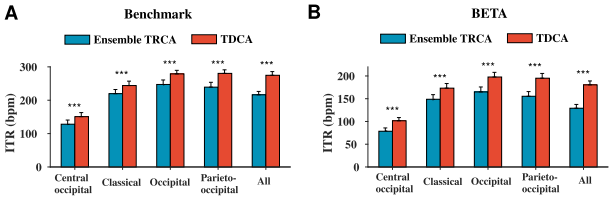

Fig. 4. Maximum average ITR with varying montage for TDCA and ensemble TRCA on public datasets (**A**: the Benchmark dataset; **B**: the BETA dataset). The number of channels is shown in ascending order, i.e., $N_{ch} = 3$ for central occipital montage, $N_{ch} = 9$ for classical montage, $N_{ch} = 21$ for occipital montage, $N_{ch} = 30$ for parietal-occipital montage, $N_{ch} = 64$ for all channels. The asterisks indicate significant differences between TDCA and ensemble TRCA (*$p < .05$, **$p < .01$, ***$p < .001$).

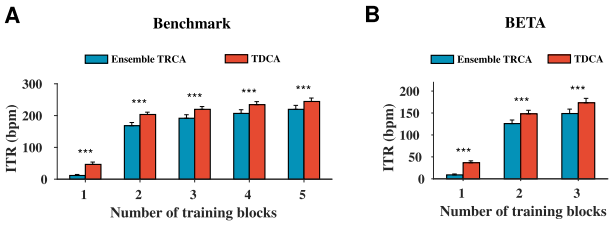

Fig. 5. Maximum average ITR with varying number of training blocks for TDCA and ensemble TRCA on public datasets (**A**: the Benchmark dataset; **B**: the BETA dataset). The asterisks indicate significant differences between TDCA and ensemble TRCA (*$p < .05$, **$p < .01$, ***$p < .001$).

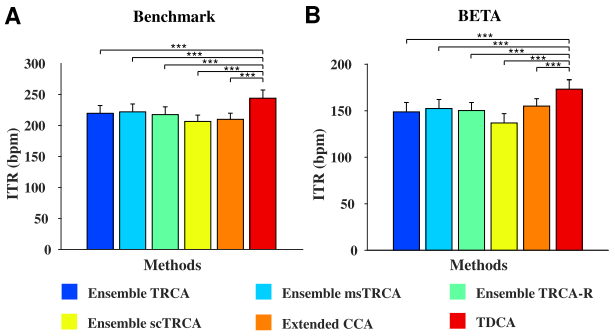

Fig. 6. Maximum average ITR of various competing methods evaluated on public datasets (**A**: the Benchmark dataset; **B**: the BETA dataset). The asterisks indicate significant differences between TDCA and ensemble TRCA (*$p < .05$, **$p < .01$, ***$p < .001$).

TABLE I
COMPARISON OF TDCA AND ENSEMBLE TRCA ON THE INDIVIDUAL ITR IN THE ONLINE EXPERIMENT

	Subject	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Mean±sem
ITR (bpm)	Ensemble TRCA	185.0	283.0	93.6	72.7	157.6	256.3	346.8	132.3	275.8	232.5	151.8	301.9	207.4±25.3
	TDCA	216.2	303.1	158.0	157.5	254.9	297.3	346.8	195.6	281.0	289.6	215.4	306.4	251.8±17.9

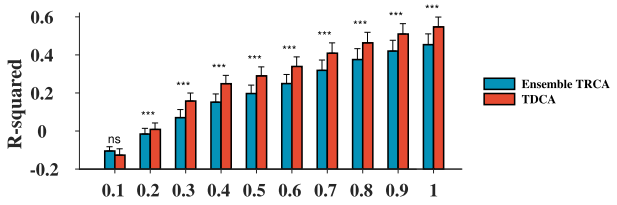

Fig. 7. R-squared feature for TDCA and ensemble TRCA on the data of the offline experiment. Data lengths from 0.2 s to 1 s (with an interval of 0.1 s) are used for evaluation. The asterisks indicate significant differences between TDCA and ensemble TRCA (*$p < .05$, **$p < .01$, ***$p < .001$).

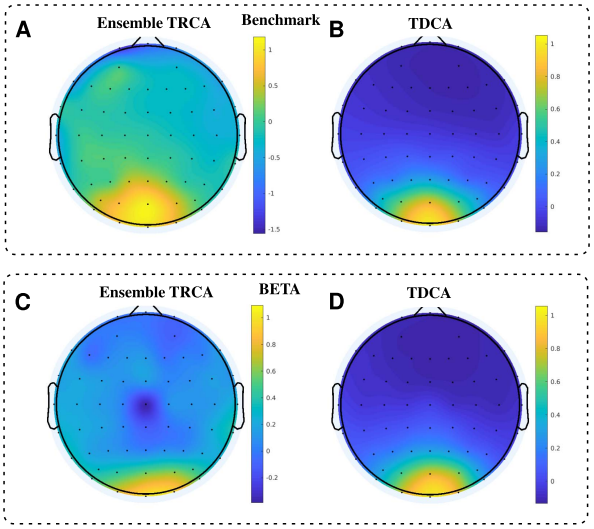

Fig. 8. Average activation pattern for TDCA and ensemble TRCA on public datasets (**A** and **B**: the benchmark dataset; **C** and **D**: the BETA dataset). The activation patterns were computed and normalized from 64-channel EEG at a data length of 1 s.

across all data lengths. In accord with the comparison shown in Figure 3, TDCA achieved the highest performance among all compared methods, both with the Benchmark dataset and the BETA dataset. One-way RMANOVA revealed that there was a statistically significant difference between the compared methods on the Benchmark dataset, $F(1.822, 61.939) = 22.408$, $p < .001$, Greenhouse-Geisser corrected, and on the BETA dataset, $F(3.365, 232.169) = 31.217$, $p < .001$, Greenhouse-Geisser corrected. Paired t-test with Bonferroni correction found a statistically significant difference between TDCA and each other method on both datasets ($p < .001$). The best performing method among the other compared methods was msTRCA for the Benchmark dataset, with a maximum average ITR of 222.00 ± 12.68 bpm, and extended CCA for the BETA dataset, with a maximum average ITR of 155.04 ± 7.94 bpm.

The profiles of the feature space are further presented in the following. As illustrated in Figure 7, the R-squared in TDCA significantly exceeded that in ensemble TRCA for data lengths exceeding 0.1 s ($p < .001$). For instance, at a data length of 1 s, R-squared increased from 0.454 ± 0.057 in ensemble TRCA to 0.547 ± 0.052 in TDCA, which boosted classification accuracy and ITR [16], [17]. Additionally, Figure 8 depicts the average activation pattern of TDCA and ensemble TRCA learned from public datasets, in which the first component of the activation pattern was delineated. Here, the average was performed across all classes and subjects for TRCA and across all subjects for TDCA. Compared with ensemble TRCA, visual inspection indicates that the activation pattern of TDCA is more densely and symmetrically distributed in the occipital region for both the Benchmark dataset (Figure 8**B**) and the BETA dataset (Figure 8**D**). For ensemble TRCA, scattered activations with high weights could be found in both frontal and temporal regions, whereas for TDCA, activations were suppressed in these regions.

For the offline experiment, the average classification accuracy and ITR of 12 subjects are illustrated in Figure 9. Consistent with the result on public datasets, TDCA yielded higher performance than ensemble TRCA on the accuracy and ITR. Paired t-test found a statistically significant difference ($p < .01$) for data lengths exceeding 0.1 s. At the data length of 0.4 s, the highest ITRs were achieved by both methods, i.e., 232.13 ± 20.05 bpm for TDCA and 200.82 ± 21.58 bpm for ensemble TRCA, $p < .001$. At a data length of 1 s, the accuracies for TDCA and ensemble TRCA were 0.917 ± 0.026 and 0.875 ± 0.037, respectively, with a difference of 0.042 ($p = .006$).

Figure 10 illustrates the performance of TDCA compared with ensemble TRCA in the online experiment. The average online accuracy of TDCA was 0.82 ± 0.04, which significantly exceeded the accuracy of ensemble TRCA, 0.72 ± 0.06 ($p = .001$). Consequently, TDCA significantly outperformed ensemble TRCA regarding the online ITR (TDCA: 251.8 ± 17.9 bpm; ensemble TRCA: 207.4 ± 25.3 bpm; $p = 6.14 \times 10^{-4}$). The details of the individual ITR are summarized in Table I. At the individual level, the results showed that TDCA yielded greater improvements for subjects with poor BCI performance, e.g., S3 and S4. However, for subjects with excellent BCI performance, e.g., S12 and S7, the improvement was only marginal and a ceiling effect was observed.

IV. DISCUSSION

This study proposes a task-discriminant component analysis (TDCA) for individually calibrated SSVEP-BCI. We first evaluated the performance of TDCA on two public datasets and the results suggest that TDCA is superior over ensemble

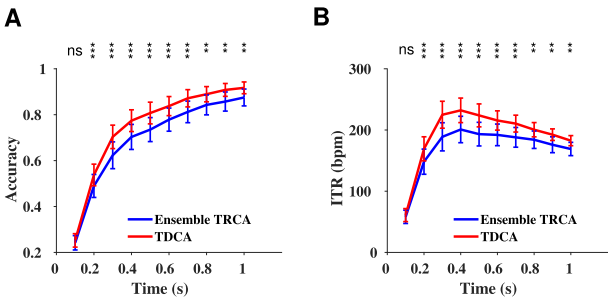

Fig. 9. Average classification accuracy (A) and ITR (B) for TDCA and ensemble TRCA in the offline experiment. Data lengths from 0.1 s to 1 s (with an interval of 0.1 s) are used for evaluation. The asterisks indicate significant differences between TDCA and ensemble TRCA (*$p < .05$, **$p < .01$, ***$p < .001$).

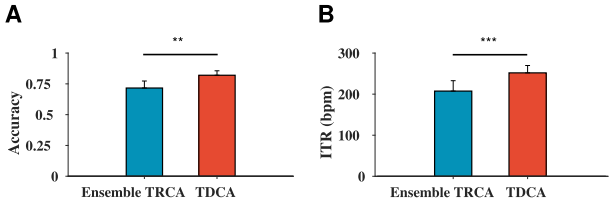

Fig. 10. Online classification accuracy (A) and ITR (B) for TDCA and ensemble TRCA. The asterisks indicate significant differences between TDCA and ensemble TRCA (*$p < .05$, **$p < .01$, ***$p < .001$).

TRCA for various data lengths, montage configurations and the number of training blocks. More importantly, TDCA significantly outperformed other competing frequency recognition methods. To validate the performance of TDCA on new subjects, an offline experiment was conducted with 12 subjects and TDCA was implemented in a subsequent online experiment. The results of both the offline and online experiments suggest that TDCA consistently performs better than ensemble TRCA. This is in line with the result of the two public datasets and demonstrates the effectiveness of the proposed method. The elevated performance of the accuracy and ITR is attributable to the boosted discriminability of different classes in the feature space, as indicated by a comparison of the R-squared statistics in TDCA and ensemble TRCA. In more detail, the increment in the discriminability is due in part to the merit of spatiotemporal filters and their activation pattern. The activation pattern of TDCA showcases its interpretability, and implies a suppression of noise and an extraction of the signal of interest as the underpinning for TDCA.

From a methodological perspective, TDCA is closely related to ensemble TRCA and there are striking distinctions between the two methods as well. First, when solving the generalized eigenvalue problem, ensemble TRCA only optimizes the spatial filter, while TDCA optimizes the spatio-temporal filter instead. Usage of the spatio-temporal filter is widely adopted in other BCI paradigms, e.g., the motor imagery BCI to address the issue of the unknown latencies from the onset of a cue to the onset of task-related brain signals [30]–[33]. Similar issues also affect SSVEP-BCI, where the latencies vary with subjects [23], [34] and the location of targets [35]. Thus, the subtlety in latency or the temporal information (apart from filter bank [18]) should be taken into account in SSVEP-BCI, which is neglected by ensemble TRCA and has received little attention in the literature [36]. Second, the two methods share a common ground in the form of the optimization objective and models from both methods differ in the specificity to classes. To identify both commonalities and differences, in TDCA, we set $l = 0$ and use the data input from TRCA, and show the relationship between TDCA and TRCA spatial filter, as presented in the Appendix. The comparison suggests that TDCA and TRCA share a fundamental form, which is a problem of maximizing the SNR similar to the formulation of maximum signal fraction analysis [37]. Regarding the difference, TDCA simultaneously optimizes the common spatial filter that maximizes the SNR of all stimulus frequencies, whereas TRCA maximizes the SNR of each stimulus frequency in turn and the spatial filter is specific to each target. As TDCA solves the common spatial filter, the computation iteration is reduced to $1/40$ compared with ensemble TRCA and TDCA does not require an ensemble technique. Third, prior information of the stimulus frequency is utilized in TDCA, whereas there are no priors in TRCA. The utilization of the stimulus frequency by orthogonal projection facilitates classification, as evidenced by previous studies [7], [14], [38].

The methodological improvement of TDCA is beneficial in practical applications. First, the number of required channels and the corresponding cost can be reduced by using TDCA. This study demonstrates that marked performance improvements can be achieved by using a number of channels exceeding the classical nine channels. However, in a real-world BCI system, configuring more channels for acquisition would substantially increase the associated cost, which decreases the affordability of the system. Thus, this study uses the classical nine channels for the online experiment. As shown in Figure 4, the use of nine channels (i.e., classical montage) in TDCA yielded a performance comparable to the use of 21 channels in ensemble TRCA, both for the benchmark dataset (TDCA in classical montage: 244.00 ± 13.21 bpm; ensemble TRCA in occipital montage: 247.25 ± 13.45 bpm, $p = .57$) and for the BETA dataset (TDCA in classical montage: 173.24 ± 10.13 bpm; ensemble TRCA in occipital montage: 164.91 ± 10.91 bpm, $p = .05$). This suggests that the application of TDCA is cost-effective in terms of channel configuration. Second, the number of calibration blocks could be decreased, as shown in Figure 5. For instance, regarding the ITR on the benchmark dataset, the use of two blocks of training data in TDCA is comparable to four blocks of data in ensemble TRCA (TDCA: 203.45 ± 7.28 bpm; ensemble TRCA: 206.84 ± 11.57 bpm; $p = .24$); the use of three blocks of training data in TDCA is comparable to five blocks of data in ensemble TRCA (TDCA: 219.66 ± 8.75 bpm; ensemble TRCA: 219.67 ± 12.43 bpm; $p = .99$). This suggests that TDCA is conducive in mitigating the calibration burden, which is time-consuming and impairs the user experience by fatigue. Third, the proposed method is promising for the enhancement of dry electrode based systems. Currently, dry electrode based systems are inferior in performance compared with gel based systems because of the system noise from measurement [39], [40]. As the dry electrode based system is easy to use and preferable

in practical applications, the enhancement of this system is important and TDCA has significant implications in boosting performance of this system. As validated by the online experiment, TDCA is computationally tractable for single-trial real-time decoding in implementation. For instance, for the 1-s data length on the BETA dataset of 70 subjects, the average computational cost was 31.8 ms on a desktop computer with a 3.6 GHz CPU (32 GB RAM). Yet, the computational overhead of TDCA is slightly higher than that of ensemble TRCA due to matrix multiplication in the projection operation, as compared in the Supplemental Table III.

The present study proposes TDCA as a proof of concept, which can be tailored by future studies to meet the demands of real-world applications. For applications that require robust control (e.g., BCI control of a drone), a dynamic stopping strategy [41] could be adopted to increase the classification accuracy via decisions of rejection. For applications where synchronization is not required or satisfying, TDCA could be integrated with an asynchronous strategy [42] to discriminate between the intentional control state and the non-control state. In addition, TDCA has the potential to be employed in other SSVEP-BCI paradigms, e.g., for flexible spatial information decoding [43].

V. Conclusion

This study proposes a novel frequency recognition method, i.e., task-discriminant component analysis (TDCA) to improve the performance of SSVEP-BCI. In TDCA, common spatio-temporal filters are learned from the data in a discriminative manner. Thus, there is no need for the model to learn the projection direction class by class, and different from ensemble TRCA, no ensemble technique is needed. The effectiveness of TDCA is validated by extensive validations on two benchmark datasets, an offline experiment, and an online experiment. The result of these validations demonstrates that TDCA is superior over ensemble TRCA and other competing methods. The present study lays a foundation for methodological extensions of TDCA and indicates promising potential applications in high-speed brain spellers.

Appendix
Relationship Between TDCA and TRCA Spatial Filters

Suppose the SSVEP observations $\boldsymbol{X}$ have the following signal model:

$$\boldsymbol{X} = \boldsymbol{S} + \boldsymbol{N} \tag{20}$$

where $\boldsymbol{S} \in \mathbb{R}^{N_{ch} \times N_p}$ is a multichannel deterministic signal, and $\boldsymbol{N} \in \mathbb{R}^{N_{ch} \times N_p}$ is random noise. The scatter matrix in Eq. (16) can be reformulated as:

$$\begin{aligned} \boldsymbol{S}_b &= \sum_i (\boldsymbol{S}_i - \frac{1}{N_c}\sum_j \boldsymbol{S}_j)(\boldsymbol{S}_i - \frac{1}{N_c}\sum_j \boldsymbol{S}_j)^T \\ \boldsymbol{S}_w &= \sum_i (\boldsymbol{X}_i - \boldsymbol{S}_i)(\boldsymbol{X}_i - \boldsymbol{S}_i)^T \end{aligned} \tag{21}$$

where $\boldsymbol{S}_i$ and $\boldsymbol{X}_i$ denote the $\boldsymbol{S}$ and $\boldsymbol{X}$ associated with the i-th stimulus frequency, respectively. The between-class scatter matrix has the form of

$$\begin{aligned} \boldsymbol{S}_b = \sum_i (\boldsymbol{S}_i \boldsymbol{S}_i^T - \frac{1}{N_c}\sum_j \boldsymbol{S}_j \boldsymbol{S}_i^T - \frac{1}{N_c}\boldsymbol{S}_i \sum_j \boldsymbol{S}_j^T \\ + \frac{1}{N_c^2}\sum_j \boldsymbol{S}_j \sum_k \boldsymbol{S}_k^T) \end{aligned} \tag{22}$$

Assume that the deterministic signals are uncorrelated, i.e., $\boldsymbol{S}_i \boldsymbol{S}_j = 0, i \neq j$. Then,

$$\begin{aligned} \boldsymbol{S}_b &= \sum_i (\boldsymbol{S}_i \boldsymbol{S}_i^T - \frac{2}{N_c}\boldsymbol{S}_i \boldsymbol{S}_i^T + \frac{1}{N_c^2}\sum_j \boldsymbol{S}_j \boldsymbol{S}_j^T) \\ &= \alpha \sum_i \boldsymbol{S}_i \boldsymbol{S}_i^T \end{aligned} \tag{23}$$

where α is a scaling parameter. In parallel, the within-class scatter matrix $\boldsymbol{S}_w$ has the form of

$$\boldsymbol{S}_w = \sum_i \boldsymbol{N}\boldsymbol{N}^T = \beta \boldsymbol{N}\boldsymbol{N}^T \tag{24}$$

where β is a scaling parameter. Thus the Eq. (15) can be written as

$$\underset{\boldsymbol{W}}{\text{maximize}} \quad \frac{tr\left[\boldsymbol{W}^T (\sum_i \boldsymbol{S}_i \boldsymbol{S}_i^T) \boldsymbol{W}\right]}{tr\left[\boldsymbol{W}^T \boldsymbol{N}\boldsymbol{N}^T \boldsymbol{W}\right]} \tag{25}$$

The optimization objective of TRCA has the form of [44]

$$\underset{\boldsymbol{W}}{\text{maximize}} \quad \frac{tr\left[\boldsymbol{W}^T \boldsymbol{S}_i \boldsymbol{S}_i^T \boldsymbol{W}\right]}{tr\left[\boldsymbol{W}^T \boldsymbol{N}\boldsymbol{N}^T \boldsymbol{W}\right]} \tag{26}$$

The comparison indicates that the TDCA spatial filter simultaneously maximizes the SNR of all stimulus frequencies, while the TRCA spatial filter maximizes the SNR of each stimulus frequency in sequence.

Acknowledgment

The authors would like to thank Z. Li from The University of Hong Kong for the assistance in manuscript preparation and B. Li from Tsinghua University for the helpful discussion.

References

[1] X. Gao, Y. Wang, X. Chen, and S. Gao, "Interface, interaction, and intelligence in generalized brain–computer interfaces," *Trends Cogn. Sci.*, vol. 25, no. 8, pp. 671–684, Aug. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364661321000966

[2] 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, pp. E6058–E6067, 2015.

[3] F. Zhu, L. Jiang, G. Dong, X. Gao, and Y. Wang, "An open dataset for wearable SSVEP-based brain-computer interfaces," *Sensors*, vol. 21, no. 4, p. 1256, Feb. 2021.

[4] M. Mahmood *et al.*, "Fully portable and wireless universal brain–machine interfaces enabled by flexible scalp electronics and deep learning algorithm," *Nature Mach. Intell.*, vol. 1, no. 9, pp. 412–422, Sep. 2019.

[5] M. Adams *et al.*, "Towards an SSVEP-BCI controlled smart home," in *Proc. IEEE Int. Conf. Syst., Man Cybern. (SMC)*, Oct. 2019, pp. 2737–2742.

[6] 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.

[7] 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.

[8] Y. Zhang, P. Xu, K. Cheng, and D. Yao, "Multivariate synchronization index for frequency recognition of SSVEP-based brain–computer interface," *J. Neurosci. Meth.*, vol. 221, pp. 32–40, Jan. 2014.

[9] P. Saidi, A. Vosoughi, and G. Atia, "Detection of brain stimuli using Ramanujan periodicity transforms," *J. Neural Eng.*, vol. 16, no. 3, Jun. 2019, Art. no. 036021.

[10] M. Nakanishi, Y. Wang, Y. Wang, Y. Mitsukura, and T. Jung, "A high-speed brain speller using steady-state visual evoked potentials," *Int. J. Neural Syst.*, vol. 24, no. 6, 2014, Art. no. 1450019.

[11] 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.

[12] M. Nakanishi, Y. Wang, X. Chen, Y. Wang, X. Gao, and T.-P. 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.

[13] C. M. Wong *et al.*, "Learning across multi-stimulus enhances target recognition methods in SSVEP-based BCIs," *J. Neural Eng.*, vol. 17, no. 1, Jan. 2020, Art. no. 016026.

[14] C. M. Wong, B. Wang, Z. Wang, K. F. Lao, A. Rosa, and F. Wan, "Spatial filtering in SSVEP-based BCIs: Unified framework and new improvements," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 11, pp. 3057–3072, Nov. 2020.

[15] Q. Sun, M. Chen, L. Zhang, C. Li, and W. Kang, "Similarity-constrained task-related component analysis for enhancing SSVEP detection," *J. Neural Eng.*, vol. 18, no. 4, Aug. 2021, Art. no. 046080.

[16] G. G. Molina and V. Mihajlovic, "Spatial filters to detect steady-state visual evoked potentials elicited by high frequency stimulation: BCI application," *Biomed. Tech./Biomed. Eng.*, vol. 55, no. 3, pp. 173–182, Jan. 2010.

[17] M. A. Pastor, J. Artieda, J. Arbizu, M. Valencia, and J. C. Masdeu, "Human cerebral activation during steady-state visual-evoked responses," *J. Neurosci.*, vol. 23, no. 37, pp. 11621–11627, 2003.

[18] 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.

[19] 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.

[20] J. Ye, R. Janardan, and Q. Li, "Two-dimensional linear discriminant analysis," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 17, 2004, pp. 1569–1576.

[21] Y. Zhang, G. Zhou, Q. Zhao, J. Jin, X. Wang, and A. Cichocki, "Spatial-temporal discriminant analysis for ERP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 2, pp. 233–243, Mar. 2013.

[22] 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.*, early access, Aug. 18, 2021, doi: 10.1109/TBME.2021.3105331.

[23] 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.

[24] 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.

[25] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.

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

[27] S. Haufe *et al.*, "On the interpretation of weight vectors of linear models in multivariate neuroimaging," *NeuroImage*, vol. 87, pp. 96–110, Feb. 2014.

[28] X. Chen, Z. Chen, S. Gao, and X. Gao, "A high-ITR SSVEP-based BCI speller," *Brain-Comput. Interfaces*, vol. 1, nos. 3–4, pp. 181–191, 2014.

[29] D. H. Brainard, "The psychophysics toolbox," *Spatial Vis.*, vol. 10, no. 4, pp. 433–436, 1997.

[30] S. Lemm, B. Blankertz, G. Curio, and K. R. Müller, "Spatio-spectral filters for improving the classification of single trial EEG," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 9, pp. 1541–1548, Sep. 2005.

[31] G. Dornhege, B. Blankertz, M. Krauledat, F. Losch, G. Curio, and K. R. Müller, "Combined optimization of spatial and temporal filters for improving brain-computer interfacing," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 11, pp. 2274–2281, Nov. 2006.

[32] W. Wu, X. Gao, B. Hong, and S. Gao, "Classifying single-trial EEG during motor imagery by iterative spatio-spectral patterns learning (ISSPL)," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 6, pp. 1733–1743, Jun. 2008.

[33] F. Qi, Y. Li, and W. Wu, "RSTFC: A novel algorithm for spatio-temporal filtering and classification of single-trial EEG," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 12, pp. 3070–3082, Dec. 2015.

[34] 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.

[35] M. Nakanishi, J. T.-P. Wang, and Y. Wang, "Vertical target locations modulate the latency of steady-state visual evoked potentials," in *Proc. 21th Annu. Meeting Org. Hum. Brain Mapping (OHBM)*, 2015.

[36] Y. Zhang, D. Guo, D. Yao, and P. Xu, "The extension of multivariate synchronization index method for SSVEP-based BCI," *Neurocomputing*, vol. 269, pp. 226–231, Dec. 2017.

[37] Q. Wei, S. Zhu, Y. Wang, X. Gao, H. Guo, and X. Wu, "Maximum signal fraction analysis for enhancing signal-to-noise ratio of EEG signals in SSVEP-based BCIs," *IEEE Access*, vol. 7, pp. 85452–85461, 2019.

[38] M. Abu-Alqumsan and A. Peer, "Advancing the detection of steady-state visual evoked potentials in brain–computer interfaces," *J. Neural Eng.*, vol. 13, no. 3, Jun. 2016, Art. no. 036005.

[39] V. Mihajlović, G. Garcia-Molina, and J. Peuscher, "Dry and water-based EEG electrodes in SSVEP-based BCI applications," in *Proc. Int. Joint Conf. Biomed. Eng. Syst. Technol.* Berlin, Germany: Springer, 2012, pp. 23–40.

[40] X. Xing *et al.*, "A high-speed SSVEP-based BCI using dry EEG electrodes," *Sci. Rep.*, vol. 8, no. 1, pp. 1–10, Dec. 2018.

[41] 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.

[42] X. Han, K. Lin, S. Gao, and X. Gao, "A novel system of SSVEP-based human–robot coordination," *J. Neural Eng.*, vol. 16, no. 1, Feb. 2019, Art. no. 016006.

[43] J. Chen *et al.*, "A spatially-coded visual brain-computer interface for flexible visual spatial information decoding," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 926–933, 2021.

[44] 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.