Multi-Stimulus Least-Squares Transformation With Online Adaptation Scheme to Reduce Calibration Effort for SSVEP-Based BCIs

Dandan Li, Xuedong Wang, Mingliang Dou, Yao Zhao, Xiaohong Cui, Jie Xiang, and Bin Wang

***Abstract*— Steady-state visual evoked potential (SSVEP), one of the most popular electroencephalography (EEG)-based brain-computer interface (BCI) paradigms, can achieve high performance using calibration-based recognition algorithms. As calibration-based recognition algorithms are time-consuming to collect calibration data, the least-squares transformation (LST) has been used to reduce the calibration effort for SSVEP-based BCI. However, the transformation matrices constructed by current LST methods are not precise enough, resulting in large differences between the transformed data and the real data of the target subject. This ultimately leads to the constructed spatial filters and reference templates not being effective enough. To address these issues, this paper proposes multi-stimulus LST with online adaptation scheme (ms-LST-OA). *Methods*: The proposed ms-LST-OA consists of two parts. Firstly, to improve the precision of the transformation matrices, we propose the multi-stimulus LST (ms-LST) using cross-stimulus learning scheme as the cross-subject data transformation method. The ms-LST uses the data from neighboring stimuli to construct a higher precision transformation matrix for each stimulus to reduce the differences between transformed data and real data. Secondly, to further optimize the constructed spatial filters and reference templates, we use an online adaptation scheme to learn more features of the EEG signals of the target subject through an iterative process trial-by-trial. *Results*: ms-LST-OA performance was measured for three datasets (Benchmark Dataset, BETA Dataset, and UCSD Dataset). Using few calibration data, the ITR of ms-LST-OA achieved 210.01±10.10 bits/min, 172.31±7.26 bits/min, and 139.04±14.90 bits/min for all three datasets, respectively. *Conclusion*: Using ms-LST-OA can reduce calibration effort for SSVEP-based BCIs.**

***Index Terms*— Brain-computer interface, multi-stimulus least squares transform, online adaptation scheme, steady-state visual evoked potential, transfer learning.**

Manuscript received 24 November 2023; revised 26 February 2024 and 1 April 2024; accepted 7 April 2024. Date of publication 10 April 2024; date of current version 19 April 2024. This work was supported in part by Shanxi Province Basic Research Plan under Grant 202303021211055, in part by the National Key Scientific and Technological Infrastructure Project "Earth System Numerical Simulation Facility" (EarthLab) under Grant 2023-EL-PT-000371, and in part by the National Natural Science Foundation of China under Grant 62176177 and Grant 62376184. *(Corresponding author: Dandan Li.)*

The authors are with the College of Computer Science and Technology (College of Data Science), Taiyuan University of Technology, Taiyuan 030024, China (e-mail: lidandan@tyut.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2024.3387283

I. Introduction

ELECTROENCEPHALOGRAPHY (EEG)-based brain-computer interfaces (BCIs) provide people with disabilities a new approach to interact with the outside world that does not rely on neural and muscular pathways, but rather by decoding the brain activities [1], [2]. Steady-state visual evoked potential (SSVEP) is one of the most popular EEG-based BCI paradigms due to its high information transfer rate (ITR) and signal-to-noise ratio (SNR) [3], [4]. In practice, SSVEP-based BCI indicates different commands by displaying visual stimuli flashing at different frequencies on a monitor. The selected command is recognized by detecting the specific frequency component of the EEG signal [5]. In the past decades, SSVEP has been proven to have great application potential, e.g., speller [6], [7], disability assistance [8], [9], smart home [10], [11], and gaming [12], [13], etc.

In previous studies, calibration-based recognition algorithms have received significant attention. With the continuous development of recognition algorithms, more and more excellent calibration-based recognition algorithms have been proposed, e.g., extended canonical correlation analysis (eCCA) [14], ensemble task-related component analysis (eTRCA) [15], and task-discriminant component analysis (TDCA) [16], etc., all of which have achieved remarkable performance. But calibration-based recognition algorithms require a large amount of calibration effort, which results in a significant amount of time consumption and user fatigue [17].

Transfer learning is used as a common solution to reduce calibration effort, in which cross-subject is a popular scenario for transfer learning [18]. Least-squares transformation (LST), a cross-subject transfer learning approach, uses a small amount of calibration data from the target subject to reduce calibration effort by transforming existing data from the source subjects to calibration data for the target subject [19]. Small data LST (sd-LST), a variant of LST, uses less calibration data than LST. Unlike LST, sd-LST constructs a common transformation matrix for all stimuli for each source subject. The data from all stimuli are transformed using these common transformation matrices [20]. Although these LST methods mentioned above reduce a lot of calibration effort and achieve high performance, there are still some problems. The transformation matrices constructed by current LST methods are not precise enough.

TABLE I
THE MAIN DIFFERENCES OF LST, SD-LST AND MS-LST

Approach	Utilized data	The purpose of transformation matrix
LST	Only the target stimulus	The target stimulus
sd-LST	All available stimuli	All stimuli
ms-LST	A range of neighboring stimuli and the target stimulus	The target stimulus

For LST, the effect of noise leads to imprecisely constructed transformation matrices when the calibration data are small. For sd-LST, the constructed common transformation matrix may not be the most appropriate for each stimulus due to differences between different stimuli. This results in differences between the transformed data and the real data, ultimately leading to the constructed spatial filters and reference templates not being effective enough.

To solve these problems, we propose the multi-stimulus LST with online adaptation scheme (ms-LST-OA) to reduce calibration effort for SSVEP-based BCIs. This approach successfully addresses the problems mentioned above: 1) [21] shows that the SSVEP impulse responses elicited by stimuli of neighboring frequencies are similar. Thus, the cross-stimulus learning scheme using data from neighboring stimuli can improve the transformation matrices constructed by LST and the ability against insufficient calibration [22]. So, we can use the cross-stimulus learning scheme to improve the current LST, called ms-LST. Different from LST and sd-LST, ms-LST not only uses the target stimulus, but also utilizes data from neighboring stimuli (the stimuli which flash at the frequencies nearby that of the target stimulus) to construct a more precise transform matrix for each stimulus. Each transformation matrix is only applied to stimuli of specific frequency. This improvement can result in better transformed data. The differences of LST, sd-LST and ms-LST are displayed in Table I. 2) Previous studies have suggested that an online adaptation scheme can improve BCI performance [23], [24]. In this paper, we modify the recognition algorithm to online learning mode to recognize EEG signals, hence the method is named ms-LST-OA. Using the online learning mode, spatial filters and reference templates are continuously optimized by learning more EEG signal features of the target object each time an EEG signal is recognized. The experiment results show that ms-LST-OA can reduce the calibration effort for SSVEP-based BCI and achieve high ITR.

The remainder of this paper is organized as follows: Section II introduces some preliminaries of SSVEP-based BCIs. Section III presents the ms-LST-OA. Section IV describes the details of the experiments, and the results are presented in section V. Finally, Section VI concludes this paper.

II. PRELIMINARIES

This section introduces some approaches, including CCA-based approaches, TRCA-based approaches, and LST.

A. CCA-Based Approaches

CCA is a popular calibration-free recognition algorithm for SSVEP-based BCIs, first introduced to SSVEP recognition by Lin et al. [25]. For the unlabeled EEG signal $X \in \mathbb{R}^{N_{ch} \times N_s}$ and the reference template $Y_k \in \mathbb{R}^{2N_h \times N_s}$, the CCA calculates the correlation coefficient between them to extract the frequency component of the EEG signal. Where N_{ch} is the number of EEG channels, N_h is the number of harmonics ($N_h = 5$ in this paper), N_s is the number of sampling points, Y_k is a sine-cosine signal of the k-th stimulus, i.e.,

$$Y_k = \begin{bmatrix} \sin(2\pi f_k t + \varphi_k) \\ \cos(2\pi f_k t + \varphi_k) \\ \vdots \\ \sin(2\pi N_h f_k t + N_h \varphi_k) \\ \cos(2\pi N_h f_k t + N_h \varphi_k) \end{bmatrix}, \tag{1}$$

where $t = [1/\mathrm{Fs}, 2/\mathrm{Fs}, \cdots, N_p/\mathrm{Fs}]$, Fs is the sampling rate. f_k and φ_k are the frequency and phase of the k-th stimulus, respectively.

To compute the correlation coefficient between X and Y_k, CCA finds two weight vectors $w_x \in \mathbb{R}^{N_{ch} \times 1}$ and $w_{Y_k} \in \mathbb{R}^{2N_h \times 1}$ by solving the following problem:

$$w_X, w_{Y_k} = \arg\max_{u,v} \frac{E\left[u^{\mathrm{T}} X Y_k^{\mathrm{T}} v\right]}{\sqrt{E\left[u^{\mathrm{T}} X X^{\mathrm{T}} u\right] E\left[v^{\mathrm{T}} Y_k Y_k^{\mathrm{T}} v\right]}}. \tag{2}$$

After that, the correlation coefficient between the EEG signal X and the reference template Y_k can be calculated as:

$$r_k = \mathrm{corr}(X^{\mathrm{T}} w_X, Y_k^{\mathrm{T}} w_{Y_k}). \tag{3}$$

Finally, the frequency of the EEG signal f_t can be determined by the following:

$$f_t = \arg\max_k \ r_k, \ k = 1, \cdots, N_f, \tag{4}$$

where N_f is the number of stimuli.

Chen et al. [26] proposed FBCCA as an extension of CCA to improve the detection of EEG signal. It uses a filter bank to decompose the EEG signal X into N_b sub-bands $(X^1, \cdots, X^{N_b})$ with different passbands. Then, the correlation coefficient r_k^i between the sub-band component X^i and the reference template Y_k can be calculated by CCA. After that, the weighted square sum is applied to the correlation coefficients of all sub-bands to calculate the final correlation coefficient r_k between the EEG signal X and the reference template Y_k:

$$r_k = \sum_{i=1}^{N_b} \left(i^{-1.25} + 0.25\right) \cdot \left(r_k^i\right)^2. \tag{5}$$

B. TRCA-Based Approaches

TRCA is a method to extract task-relevant components by maximizing reproducibility during the task periods, originally proposed by Tanaka et al. [27] and later introduced by Nakanishi et al. [15] for SSVEP-based BCIs. For the k-th stimulus, TRCA extracts spatial filters $w_k \in \mathbb{R}^{N_{ch} \times 1}$ by

maximizing the inter-trial covariance to eliminate the influence of task-unrelated signal as much as possible, i.e.,

$$w_k = \arg\max_u u^{\mathrm{T}} S_k u = \arg\max_u u^{\mathrm{T}} \left(\sum_{\substack{i,j=1 \\ i \neq j}}^{N_{\mathrm{colk}}} \mathrm{Cov}\left(X_{k,i}, X_{k,j}\right) \right) u, \tag{6}$$

where $S_k \in \mathbb{R}^{N_{ch} \times N_{ch}}$ denotes the sum of inter-trial covariance of the calibration data for the k-th stimulus, N_{cali} is the number of calibration trials for each stimulus, $X_{k,i}$ denotes the i-th trial of the k-th stimulus.

In order to have a finite solution to (6), the variable $Q_k \in \mathbb{R}^{N_{ch} \times N_{ch}}$ is defined as:

$$Q_k = \mathrm{Cov}\left(\Lambda_k, \Lambda_k\right), \tag{7}$$

where $\Lambda_k = \left[X_{k,1}, \cdots, X_{k,N_{cali}}\right]$ is the concatenated matrix of all calibration data for the k-th stimulus.

Then the above problem can be solved by the following equation:

$$w_k = \arg\max_u \frac{u^{\mathrm{T}} S_k u}{u^{\mathrm{T}} Q_k u}. \tag{8}$$

After that, TRCA calculates the correlation coefficient based on the EEG signal X and the subject-specific averaged reference template $\overline{X_k} \in \mathbb{R}^{N_{ch} \times N_s}$:

$$r_k = \mathrm{corr}(X^{\mathrm{T}} w_k, \overline{X_k}^{\mathrm{T}} w_k), \tag{9}$$

where

$$\overline{X_k} = \frac{1}{N_{cali}} \sum_{i=1}^{N_{cali}} X_{k,i}. \tag{10}$$

Furthermore, Nakanishi et al. [15] proposed ensemble TRCA as an extension of TRCA to further improve the performance by integrating all spatial filters, i.e.,

$$W = \left[w_1, w_2, \cdots, w_{N_f}\right]. \tag{11}$$

After that (9) is modified as follows:

$$r_k = \mathrm{corr}(X^{\mathrm{T}} W, \overline{X_k}^{\mathrm{T}} W). \tag{12}$$

Finally, the frequency of the EEG signal f_t can be determined by (4).

C. Least-Squares Transformation

Chiang et al. [19] proposed LST used for SSVEP cross-subject transfer learning to reduce calibration effort for the target subject.

LST utilizes least squares regression to construct a transformation matrix ${}^{scr}_{j}P_{k,i} \in \mathbb{R}^{N_{ch} \times N_{ch}}$ for the target subject as follow:

$${}^{scr}_{j}P_{k,i} = \overline{{}^{tar}X_k}\, {}^{scr}_{j}X^{\mathrm{T}}_{k,i} \left({}^{scr}_{j}X_{k,i}\, {}^{scr}_{j}X^{\mathrm{T}}_{k,i} \right)^{-1}, \tag{13}$$

where ${}^{tar}X_{k,i} \in \mathbb{R}^{N_{ch} \times N_s}$ and ${}^{scr}_{j}X_{k,i} \in \mathbb{R}^{N_{ch} \times N_s}$ denote the calibration data for the i-th trial of the k-th stimulus for the target subject and the j-th source subject, respectively. $\overline{{}^{tar}X_k} \in \mathbb{R}^{N_{ch} \times N_s}$ denotes the averaged calibration data for the k-th stimulus of the target subject calculated as:

$$\overline{{}^{tar}X_k} = \frac{1}{{}^{tar}N_{cali}} \sum_{i=1}^{{}^{tar}N_{cali}} {}^{tar}X_{k,i}, \tag{14}$$

${}^{tar}N_{cali}$ is the number of calibration trials for each stimulus of the target subject.

After that, the calibration data of the source subject can be transformed into the calibration data of the target subject by ${}^{scr}_{j}X'_{k,i} = {}^{scr}_{j}P_{k,i}\, {}^{scr}_{j}X_{k,i}$. Finally, the calibration dataset of the target subject ${}^{tar}D'$ consists of the original calibration data of the target subject and the transformed calibration data of other source subjects, i.e.,

$$\begin{aligned} {}^{tar}D' = &\left\{ {}^{tar}X_{k,i} \right\}_{k \in [1,N_f],\ i \in [1,{}^{tar}N_{cali}]} \\ &\oplus \left\{ {}^{scr}_{j}X'_{k,i} \right\}_{j \in [1,N_{sub}],\ k \in [1,N_f],\ i \in \left[1,{}^{scr}_{j}N_{cali}\right]}, \end{aligned} \tag{15}$$

where N_{sub} is the number of the source subjects, ${}^{scr}_{j}N_{cali}$ is the number of calibration trials for each stimulus of the j-th source subject, $\oplus$ indicates data merging operation.

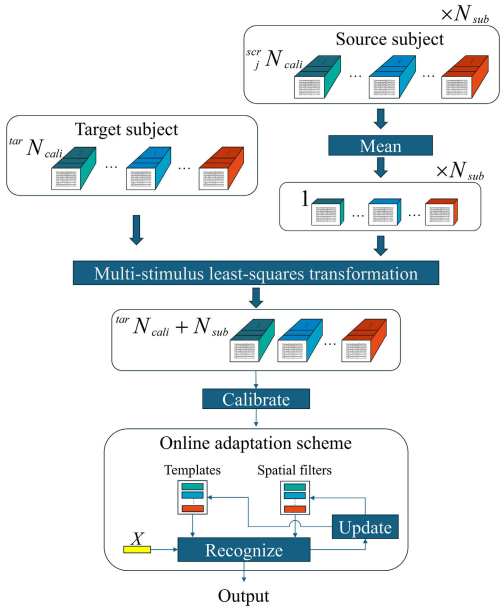

Fig. 1. The overall framework of multi-stimulus LST with online adaptation scheme (ms-LST-OA).

III. Methods

In this study, we propose the ms-LST-OA to reduce the calibration effort for SSVEP-based BCI, as displayed in Fig. 1. This approach consists of two parts: ms-LST and online adaptation scheme.

A. Multi-Stimulus LST

The SSVEPs of subjects elicited by stimuli of neighboring frequencies have similar impulse responses that can assist LST in constructing transformation matrices with higher precision [21]. Furthermore, cross-stimulus learning scheme using neighboring stimuli can also improve the ability against

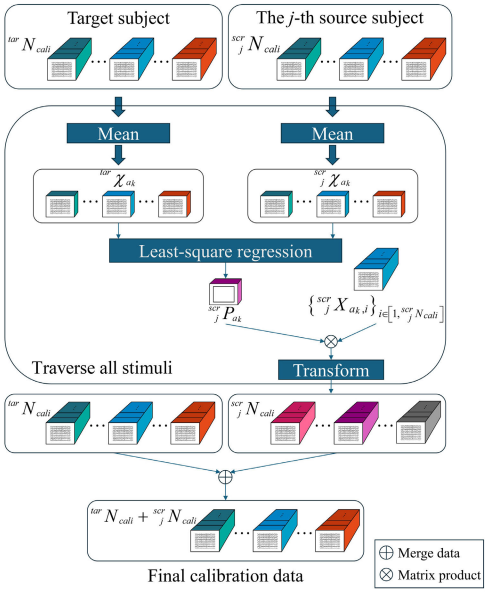

Fig. 2. The overall framework of multi-stimulus LST (ms-LST).

insufficient calibration [22]. Therefore, the proposed ms-LST utilizes the data from neighboring stimuli to construct a transformation matrix for each stimulus, which can improve the precision of the transformation matrix. The overall framework of ms-LST is shown in Fig. 2.

To transform the calibration data of the j-th source subject for the target subject, we need to sort the dataset of the target subject ${}^{tar}D$ and the dataset of the j-th source subjects ${}^{scr}_{j}D$ according to the frequency of the stimuli:

$$ {}^{tar}D = \left\{ {}^{tar}X_{a_1,i}, \cdots {}^{tar}X_{a_{N_f},i} \right\}_{i\in\left[1,{}^{tar}N_{cali}\right]}, \tag{16} $$

$$ {}^{scr}_{j}D = \left\{ {}^{scr}_{j}X_{a_1,i}, \cdots, {}^{scr}_{j}X_{a_{N_f},i} \right\}_{i\in\left[1,{}^{scr}_{j}N_{cali}\right]}, \tag{17} $$

where $a_i \in \left[1, N_f\right]$ is stimulus index and these indexes need to satisfy:

$$ f_{a_1} < f_{a_2} < \cdots < f_{a_{N_f}}. \tag{18} $$

As in [20], the target subject can use data from only K stimuli for calibration with $K \leqslant N_f$. The K stimuli need to be uniformly distributed over N_f frequencies as:

$$ a_i = 1 + \left\lfloor \frac{N_f \times (2i-1)}{2K} \right\rfloor. \tag{19} $$

Then, we can average the data of these K stimuli according to get K subject-specific templates, while the other $N_f - K$ templates are filled with zero matrices.

Unlike LST, the ms-LST utilizes data not only from the target stimulus (which flashes at f_{a_k}) but also from neighboring stimuli (which flash at the frequencies nearby f_{a_k}) to construct transformation matrix for the target stimulus by solving the following problem:

$$ {}^{scr}_{j}P_{a_k} = \arg\min_{P} \left(\left\| P \cdot {}^{scr}_{j}\chi_{a_k} - {}^{tar}\chi_{a_k} \right\|_F^2 \right), \tag{20} $$

where ${}^{tar}\chi_{a_k}$ and ${}^{scr}_{j}\chi_{a_k}$ denote the subject-specific templates for the stimulus of frequency f_{a_k} and its neighboring stimuli for the target subject and the j-th source subject, respectively, which can be calculated as:

$$ {}^{tar}\chi_{a_k} = \left[\overline{{}^{tar}X_{a_k-p}}, \cdots, \overline{{}^{tar}X_{a_k+q}} \right], \tag{21} $$

$$ {}^{scr}_{j}\chi_{a_k} = \left[\overline{{}^{scr}_{j}X_{a_k-p}}, \cdots, \overline{{}^{scr}_{j}X_{a_k+q}} \right], \tag{22} $$

$\overline{{}^{tar}X_{a_k}}$ and $\overline{{}^{scr}_{j}X_{a_k}}$ denote the subject-specific templates of frequency f_{a_k} for the target subject and the j-th source subject, respectively, $\overline{{}^{scr}_{j}X_{a_k}}$ can be calculated as follows:

$$ \overline{{}^{scr}_{j}X_{a_k}} = \frac{1}{{}^{scr}_{j}N_{cali}} \sum_{i=1}^{{}^{scr}_{j}N_{cali}} {}^{scr}_{j}X_{a_k,i}. \tag{23} $$

As in [22], the range from p to q is the range of neighboring stimuli s for the stimulus of frequency f_{a_k}, and $s = p+q+1$. To ensure that at least one template falls within the range of neighboring stimuli, $s \geqslant \lceil N_f/K \rceil$. Setting the parameter $s_0 = s/2$ when s is even and $s_0 = (s-1)/2$ when s is odd, the values of p and q can be determined as:

$$ \begin{cases} p = 1, \\ q = s, \end{cases} \quad a_k \in [1, s_0] $$

$$ \begin{cases} p = k - s_0, \\ q = k + s - s_0 - 1, \end{cases} \quad a_k \in (s_0, N_f - s_0 + 1). $$

$$ \begin{cases} p = N_f - s + 1, \\ q = N_f, \end{cases} \quad a_k \in [N_f - s_0 + 1, N_f] \tag{24} $$

Then the transformation matrix ${}^{scr}_{j}P_{a_k}$ for the a_k-th stimulus of the j-th source subject can be calculated as follows:

$$ {}^{scr}_{j}P_{a_k} = {}^{tar}\chi_{a_k} \, {}^{scr}_{j}\chi^{\mathrm{T}}_{a_k} \left({}^{scr}_{j}\chi_{a_k} \, {}^{scr}_{j}\chi^{\mathrm{T}}_{a_k} \right)^{-1}. \tag{25} $$

Finally, the transformation matrix is applied to each trial data for the a_k-th stimulus of the j-th subject:

$$ {}^{scr}_{j}X'_{a_k,i} = {}^{scr}_{j}P_{a_k} \, {}^{scr}_{j}X_{a_k,i}, \ i = 1, \cdots, {}^{scr}_{j}N_{cali}. \tag{26} $$

After the transformation is completed for all trials for all subjects, the final calibration data for the target subject ${}^{tar}D'$ is expanded as:

$$ {}^{tar}D' = \left\{ {}^{tar}X_{a_k,i} \right\}_{k\in[1,K],\ i\in\left[1,{}^{tar}N_{cali}\right]} \oplus \left\{ {}^{scr}_{j}X'_{k,i} \right\}_{k\in[1,N_f],\ i\in\left[1,{}^{scr}_{j}N_{cali}\right],\ j\in[1,N_{sub}]}. \tag{27} $$

We used the final calibration data ${}^{tar}D'$ to calibrate the eTRCA recognition algorithm. Since previous studies have proven that the ensemble classifier has higher classification performance, and the sine-cosine templates also contribute to SSVEP signal recognition [14]. Therefore, we use the ensemble classifier of eTRCA and FBCCA to determine the classification results by combining the correlation coefficients calculated by eTRCA and FBCCA:

$$ r_k = r_{k,1} + r_{k,2} \tag{28} $$

where $r_{k,1}$ and $r_{k,2}$ denote the correlation coefficient of FBCCA and eTRCA for the k-th stimulus, respectively.

After calculating the correlation coefficient r_k, the frequency of stimuli is determined according to (4).

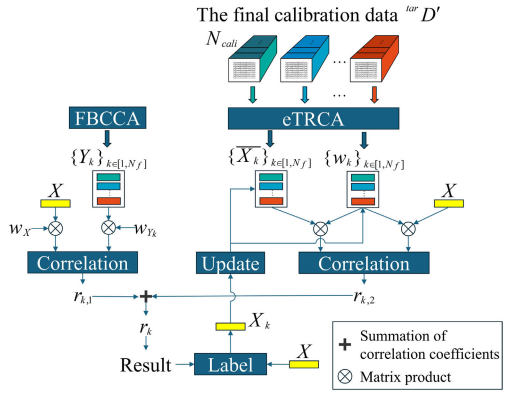

Fig. 3. The overall framework of the online adaptation scheme.

B. Online Adaptation Scheme

To further optimize the spatial filters and reference templates constructed by eTRCA using $^{tar}D'$, as in [24] and [23], this paper uses an online adaptation scheme to modify the eTRCA to online learning mode. The overall framework of the online adaptation scheme is shown in Fig. 3.

We set the state after calibrating the recognition algorithm using transformed data by ms-LST to the 0-th recognition. Suppose that the test data X is recognized as the k-th stimulus in the c-th recognition. So, the reference template and spatial filter for the k-th stimulus need to be updated.

Firstly, the reference template $\overline{X_k}^{[c]}$ and the number of calibration data $N_{cali,k}^{[c]}$ for the k-th stimulus are updated:

$$\overline{X_k}^{[c]} = \frac{N_{cali,k}^{[c-1]} \cdot \overline{X_k}^{[c-1]} + X}{N_{cali,k}^{[c-1]} + 1}, \tag{29}$$

$$N_{cali,k}^{[c]} = N_{cali,k}^{[c-1]} + 1, \tag{30}$$

where $N_{cali,k}^{[c]}$ is the number of calibration trials for k-th stimulus at the c-th recognition, $\overline{X_k}^{[c]}$ denotes the subject-specific averaged reference template for the k-th stimulus at the c-th recognition.

Then, according to (8), in order to update the eTRCA-based spatial filter, the values of the covariance matrices $S_k^{[c]}$ and $Q_k^{[c]}$ at the c-th recognition need to be updated.

According to (6), for the test data X, $S_k^{[c]}$ can be updated as:

$$\begin{aligned} S_k^{[c]} &= S_k^{[c-1]} + \sum_{i=1}^{N_{cali,k}^{[c-1]}} \mathrm{Cov}\ (X_{k,i}, X) + \sum_{i=1}^{N_{cali,k}^{[c-1]}} \mathrm{Cov}\ (X, X_{k,i}) \\ &= S_k^{[c-1]} + \mathrm{Cov}\left(N_{cali,k}^{[c]} \cdot \overline{X_k}^{[c]}, X\right) \\ &\quad + \mathrm{Cov}\left(X, N_{cali,k}^{[c]} \cdot \overline{X_k}^{[c]}\right) - 2 \cdot \mathrm{Cov}\,(X, X) \end{aligned} \tag{31}$$

where $S_k^{[c]}$ denotes the covariance matrix of the k-th stimulus S_k at the c-th recognition.

According to (7), $Q_k^{[c]}$ can be calculated as:

$$Q_k^{[c]} = \mathrm{Cov}\left(\Lambda_k^{[c]}, \Lambda_k^{[c]}\right) \tag{32}$$

where $\Lambda_k^{[c]}$ denotes the set of SSVEP data for the k-th stimulus at the c-th recognition defined by.

$$\Lambda_k^{[c]} = \left[X_{k,1}, \cdots, X_{k,N_{cali,k}^{[c-1]}}, X\right]. \tag{33}$$

To facilitate the use of $\overline{X_k}^{[c]}$ to update $Q_k^{[c]}$, the variable $Z_k^{[c]}$ is defined as:

$$Z_k^{[c]} = \sum_{i=1}^{N_{cali,k}^{[c]}} X_{k,i} \cdot \left(X_{k,i}\right)^{\mathrm{T}}. \tag{34}$$

For the test data X, $Z_k^{[c]}$ can be updated as:

$$Z_k^{[c]} = Z_k^{[c-1]} + X \cdot X^{\mathrm{T}}. \tag{35}$$

Thus, $Q_k^{[c]}$ can be updated as:

$$\begin{aligned} Q_k^{[c]} &= \mathrm{Cov}\left(\Lambda_k^{[c]}, \Lambda_k^{[c]}\right) \\ &= Z_k^{[c]} - \Psi_k^{[c]} \cdot N_{cali,k}^{[c]} \cdot \left(\overline{X_k}^{[c]}\right)^{\mathrm{T}} \\ &\quad - N_{cali,k}^{[c]} \cdot \overline{X_k}^{[c]} \cdot \left(\Psi_k^{[c]}\right)^{\mathrm{T}} + \Psi_k^{[c]} \cdot \left(\Psi_k^{[c]}\right)^{\mathrm{T}}, \end{aligned} \tag{36}$$

where $\Psi_k^{[c]} = \mathrm{mean}\left(\overline{X_k}^{[c]}, 2\right) \cdot \mathrm{I}_{1\times N_s}$, $\mathrm{I}_{1\times N_s}$ denotes a vector of $1 \times N_s$ filled by 1. The mean $\left(\overline{X_k}^{[c]}, 2\right)$ function represents the mean value on the matrix $\overline{X_k}^{[c]}$ in the dimension of the sampling points and results in a vector of $\mathbb{R}^{N_{ch}\times 1}$.

Finally, we can update the spatial filter $w_k^{[c]}$:

$$w_k^{[c]} = \arg\max_u \frac{u^{\mathrm{T}} S_k^{[c]} u}{u^{\mathrm{T}} Q_k^{[c]} u}. \tag{37}$$

In this way, the c-th update of the spatial filter and the reference template is completed.

Additionally, in order to reduce the effect of noise and improve the performance of the online adaptation scheme, the data from source subjects is averaged across trials before transforming the data using ms-LST.

IV. Experiments

To evaluate the performance of ms-LST-OA, we conduct experiments for three datasets using the leave-one-out cross-validation approach, i.e., one subject is used as the target subject while the others are used as source subjects. Furthermore, part of the data of the target subject was used as calibration data and the other part as test data.

A. SSVEP Datasets

A total of three SSVEP datasets are used in this paper for method performance testing:

1) Dataset I: Benchmark Dataset was presented by Wang et al. [28]. This dataset contains EEG data from 35 healthy subjects. For each subject, the data contains 6 blocks. Each block contains 40 trials corresponding to 40 stimuli from 8.0 Hz to 15.8 Hz with an interval of 0.2 Hz.

TABLE II
THE PARTIAL PARAMETERS OF THESE THREE DATASETS

Dataset	Channels	Stimuli	Blocks	Subjects
I	64	40	6	35
II	64	40	4	70
III	8	12	15	10

2) Dataset II: BETA Dataset was presented by Liu et al. [29]. This dataset contains EEG data from 70 healthy subjects. For each subject, the data contains 4 blocks. Each block contains 40 trials corresponding to 40 stimuli from 8.0 Hz to 15.8 Hz with an interval of 0.2 Hz.
3) Dataset III: UCSD Dataset was presented by Nakanishi et al. [30]. This dataset contains EEG data from 10 healthy subjects. For each subject, the data contains 15 blocks. Each block contains 12 trials corresponding to 12 stimuli from 9.25 Hz to 14.75 Hz with an interval of 0.5 Hz.

The partial parameters of these three datasets are presented in Table II, which denotes the number of channels, the number of stimuli, the number of trials for each stimulus, and the number of subjects in the dataset, respectively.

B. Data Preprocessing

Considering the latency in SSVEP response, we used data segment $[T_l, T_l + T_w]$ after stimulus onset for the analysis, where T_l is the SSVEP latency ($T_l = 0.14$ s in this paper) and T_w denotes time windows ($T_w \in \{0.4\text{ s},\ 0.5\text{ s}, \cdots,\ 1.0\text{ s}\}$). In addition, for Dataset I and II, nine electrodes (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, and O2) were used for analysis. For Dataset III, all electrodes were selected for data analysis. After data formatting, the EEG data for each trial was removed from the power-line noise using a 50 Hz notch filter. Finally, the EEG data for each trial was decomposed into $N_b = 5$ sub-bands using different bandpass filters, where the upper and lower cut-off frequencies of the i-th sub-band were set to $(i \times 8 - 2)$Hz and 90 Hz, respectively.

C. Performance Evaluation

We use the averaged classification accuracy (ACC) and averaged ITR to measure the performance of the involved methods. In particular, the ITR is calculated as:

$$\begin{aligned}\text{ITR} = [&\log_2\left(N_f\right) + P\log_2(P) \\ &+ \ (1-P)\log_2\left(\frac{1-P}{N_f-1}\right)] \times \frac{60}{T}, \end{aligned} \tag{38}$$

where P is the ACC, T denotes the time required to complete a recognition, and $T = T_w + 0.5$s, 0.5s is the gaze shifting time.

D. Data Analysis

The following experiments were conducted to test the performance of ms-LST-OA. All LST-based methods use the final calibration data ${}^{tar}D'$ to train the ensemble classifier of eTRCA and FBCCA to recognize SSVEP data based on (29).

1) Parameter Exploration of ms-LST: According to (25) there are many parameters that affect the performance of ms-LST, such as the range of neighboring stimuli s, the number of data channels N_{ch}, the time window T_w, the amount of calibration data ${}^{tar}N_{cali}$, and so on. As the LST-based methods are used to reduce the calibration effort, only a very small amount of calibration data was used in the experiment (${}^{tar}N_{cali} = 1$ and $K = N_f$). So, this section explores the influence of other parameters for the performance of ms-LST (s, N_{ch}, and T_w). Where, since Dataset III has only 8 channels, the experiment was conducted only for Dataset I and II when exploring the effect of the number of data channels.

2) Effectiveness of ms-LST-OA: To verify the effectiveness of ms-LST-OA, the following experiments are conducted in this section: 1) To demonstrate that the cross-stimulus learning scheme can improve the performance of LST, ms-LST is compared with LST and sd-LST. 2) To prove that the online adaptation scheme can further optimize the spatial filters and reference templates, ms-LST-OA is compared with ms-LST. All experiments were conducted at ${}^{tar}N_{cali} = 1$ and $K \in [1, N_f]$.

3) Ablation Experiment: Compared to the LST method, the proposed ms-LST-OA adopts cross-stimulus learning scheme that uses data from neighboring stimuli to construct the transformation matrices and utilizes an online adaptation scheme to optimize spatial filters and reference templates. So, the ablation experiment was conducted to explore the effects of these improvements. In the experiment, only one calibration data was used per stimulus, i.e., ${}^{tar}N_{cali} = 1$ and $K = N_f$.

4) The Ability of ms-LST-OA to Reduce Calibration Effort: In this section, to explore the ability of ms-LST-OA to reduce calibration effort, only a small amount of data is used for calibration. The calibration-based approach uses as much data as possible and makes its ITR comparable to ms-LST-OA. The ability of ms-LST-OA to reduce the calibration effort is reflected by comparing the ITR and the number of calibration data at different time windows $T_w \in \{0.4\text{s}, 0.5\text{s}, \cdots, 1.0\text{s}\}$. Among the approaches that participate in the comparison besides ms-LST-OA are OACCA [24], eCCA [14], eTRCA [15], TDCA [16], stCCA [31] and sd-LST [20]. Where OACCA is the state-of-the-art calibration-free algorithm, eCCA and eTRCA are the classical calibration-based algorithms, TDCA is the state-of-the-art calibration-based algorithm, and stCCA and sd-LST are the state-of-the-art transfer learning methods.

V. RESULTS

A. Parameter Exploration of ms-LST

The ITR of ms-LST for Dataset I, II, and III at the number of templates $K = N_f$ with different ranges of neighboring stimuli s are shown in Fig. 4. The ITR peaked at $s = 11$ (213.15 ± 10.86 bits/min), $s = 14$ (181.33 ± 7.51 bits/min), and $s = 7$ (137.20 ± 13.45 bits/min) for the three datasets, respectively. In addition, at $N_f = 40$ and $K = N_f$, we averaged the ITR of ms-LST for Dataset I and II at each s and obtained the optimal $s = 11$ (which maximizes the ITR), ITR $= 197.05 \pm 9.23$ bits/min. At $N_f = 12$ and $K = N_f$, the optimal s is 7 according to the results of Dataset III.

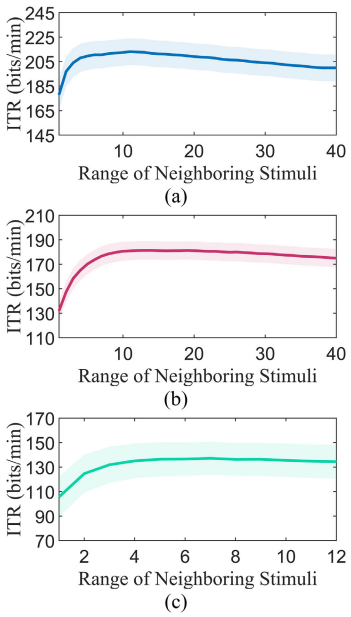

Fig. 4. The ITR of ms-LST-OA was tested for Dataset I (a), II (b) and III (c) under different ranges of neighboring frequencies at $K = N_f$. In the plots, the curves indicate the average ITR of ms-LST across subjects, and the shaded areas indicate standard errors.

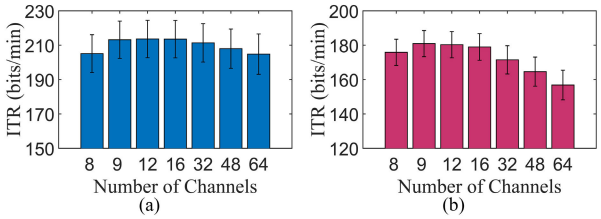

Fig. 5. Averaged ITR across subjects for Dataset I (a) and II (b) for ms-LST at different number of channels. Error bars indicate standard errors.

To get a more general scenario, we repeated the experiment at different numbers of templates K and ranges of neighboring stimuli s. Then, we obtained the optimal s at different K for $N_f = 40$ and $N_f = 12$, respectively. After that, polynomial fitting was used to fit the optimal s and K. Finally, the fitting result showed that the optimal s can be obtained according to for Dataset I and II, and according to for Dataset III.

$$s = \left\lfloor 40 \times K^{-0.34} + 1/2 \right\rfloor, \tag{39}$$

$$s = \left\lfloor 12 \times K^{-0.2} + 1/2 \right\rfloor, \tag{40}$$

The averaged ITR across subjects of ms-LST for Dataset I and II at different numbers of channels is shown in Fig. 5. In this part of the experiment, the experiment was performed only for Dataset I and II because Dataset III has only 8 channels. In the experiment, we measured the ITR of the ms-LST method when only the SSVEP-related channels were used (when the number of channels is 8 and 9) and when the SSVEP-unrelated channels were used (when the number of channels is greater than 9). When only SSVEP-related channels were utilized, the average ITR was maximized with all SSVEP-related channels being used (213.15 ± 10.86 bits/min for Dataset I and 180.94 ± 7.60 bits/min for Dataset II at $N_{ch} = 9$). When a small number of SSVEP-unrelated channels

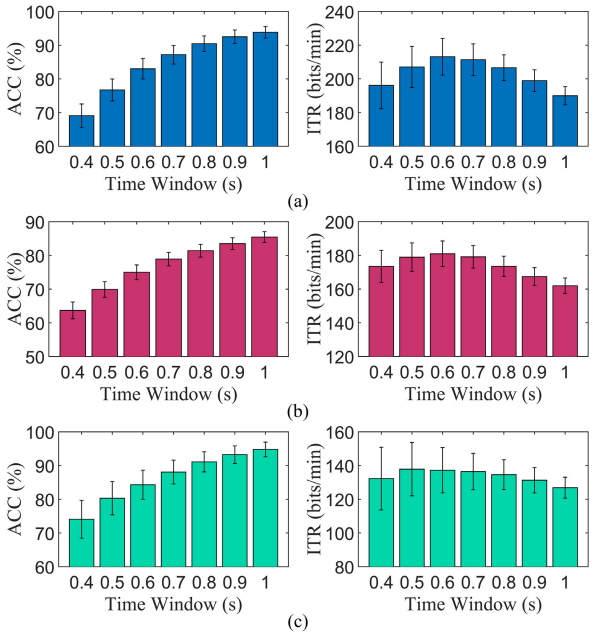

Fig. 6. Averaged ACC and ITR across subjects for ms-LST for Dataset I (a), II (b) and III (c). Error bars indicate standard errors.

were used, the ITR was not far from the ITR when all SSVEP-related channels were used (213.60 ± 10.86 bits/min for Dataset I and 180.30 ± 7.64 bits/min for Dataset II at $N_{ch} = 12$, only a very slight difference from $N_{ch} = 9$). However, as the number of SSVEP-unrelated channels continued to increase, the ITR decreased (204.80 ± 11.74 bits/min for Dataset I and 156.83 ± 8.64 bits/min for Dataset II at $N_{ch} = 64$). There is a large decrease compared to that at $N_{ch} = 9$. Therefore, utilizing the SSVEP-unrelated channels was not helpful for ms-LST, so only all SSVEP-related channels were used in the following experiments.

Fig. 6 illustrates the performance of the ms-LST method under different time windows. Generally, high ITR can be achieved by getting a higher ACC in a shorter time. For Dataset I and II, the ms-LST method reaches the highest ITR at 0.6s (213.15 ± 10.86 bits/min and 180.94 ± 7.60 bits/min), and at 0.5s for Dataset III (137.85 ± 15.82 bits/min). This shows that shorter time windows (e.g., 0.6s or 0.7s) can be chosen to get a higher performance.

B. Effectiveness of ms-LST-OA

To demonstrate that the cross-stimulus learning scheme can improve the performance of LST, we compared the ITR of ms-LST with that of LST and sd-LST at $T_w = 0.6$s with different K. The results of the comparison are displayed in Fig. 7. For all three datasets, ms-LST was able to obtain an ITR that exceeded that of LST using data from only a small number of templates. Furthermore, when the number of templates was small, the ITR of ms-LST was comparable to that of sd-LST. But as the number of templates increased, the ITR of sd-LST increased only slightly, whereas that of ms-LST increased more significantly. In addition, we calculated significant differences between the ITR of ms-LST and sd-LST at different templates numbers using paired $t-$test and corrected the $t-$test result using the Bonferroni method.

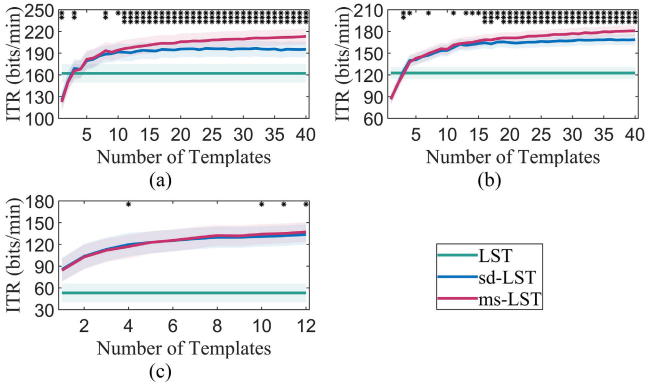

Fig. 7. The averaged ITR of the ms-LST compared to LST and sd-LST at different numbers of templates (K) for Dataset I (a), II (b) and III (c). Shaded areas indicate standard errors. Significant differences between ms-LST and sd-LST are also calculated using paired $t-$test and the results are corrected using the Bonferroni method. Asterisks indicate significant difference between ms-LST and sd-LST (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$).

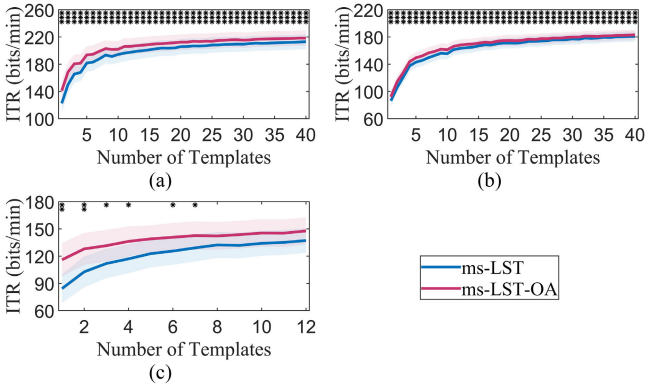

Fig. 8. The averaged ITR of ms-LST-OA compared to ms-LST at different numbers of templates (K) for Dataset I (a), II (b) and III (c). Shaded areas indicate standard errors. Significant differences are also calculated using paired $t-$test and the results are corrected using the Bonferroni method. Asterisks indicate significant difference between ms-LST-OA and ms-LST (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$).

The results show that the difference between the performance of ms-LST significantly outperforms sd-LST as K increases for all three datasets (for Dataset I, $p < 0.05$ when $K \geqslant 10$, for Dataset II, $p < 0.05$ when $K \geqslant 13$ and $p < 0.05$ when $K \geqslant 10$ for Dataset III). Thus, as the number of templates increases, ms-LST can achieve a higher ITR than sd-LST, and the performance difference becomes increasingly significant. These results suggest that the cross-stimulus learning scheme can greatly improve LST performance.

To prove that the online adaptation scheme can further optimize the spatial filters and reference templates, we compared the ITR of ms-LST-OA and ms-LST at $T_w = 0.6$s with different K. The results of comparison are displayed in Fig. 8. For all three datasets, ITR was higher for ms-LST-OA than for ms-LST at different numbers of templates. Furthermore, we also calculated the significant difference between ms-LST-OA and ms-LST at different numbers of templates using paired $t-$test and corrected the $t-$test result using the Bonferroni method. For Dataset I and II, the ITR of ms-LST-OA was significantly higher than that of ms-LST for all $K \in [1, 40]$ ($p < 0.001$). For Dataset III, the performance

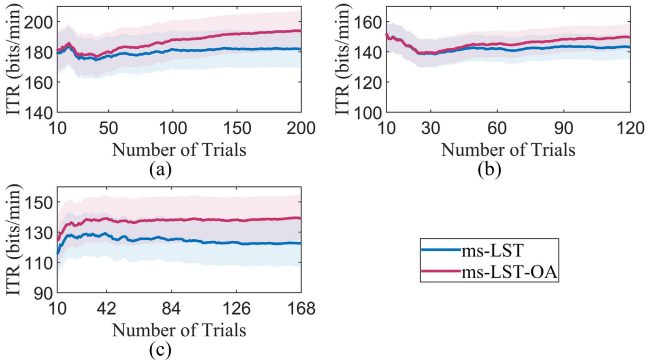

Fig. 9. The averaged performance changes of the ms-LST-OA with the number of trials at K=5 for Dataset I (a), II (b) and III (c). The shaded areas indicate the standard error across subjects.

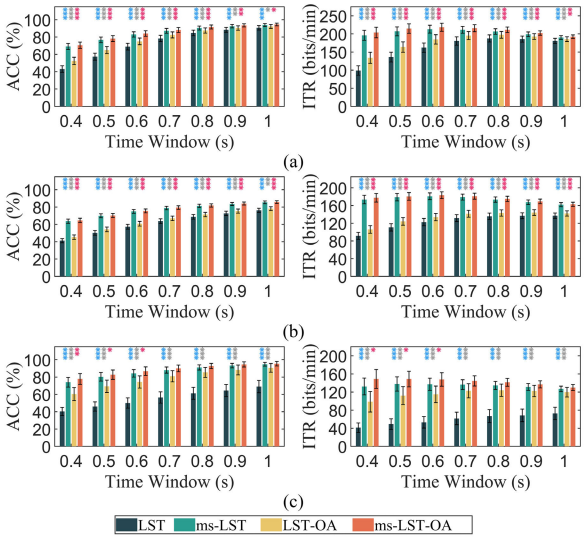

Fig. 10. Results of ablation experiment for Dataset I (a), II (b) and III (c). Error bars indicate standard errors. Significant differences are also calculated using paired $t-$test and the paired $t-$test results are corrected using the Bonferroni method. Asterisks indicate significant difference (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$). Blue, grey and pink indicate the significant differences between LST and ms-LST, LST and LST with online adaptation scheme (LST-OA), ms-LST and LST-OA, respectively.

difference between ms-LST-OA and ms-LST decreased as K increased ($p < 0.01$ at $K = 1$, $p < 0.05$ at $K = 6$, and when $K = 12$, there is no significant difference). Fig. 9 illustrates the averaged performance changes of the ms-LST-OA with the number of trials at $K = 5$. The result shows that the performance difference between ms-LST-OA and ms-LST became gradually larger. Therefore, these experiment results demonstrate that the online adaptive scheme can further optimize the spatial filters and reference templates constructed from ms-LST transformed data to improve performance.

Consequently, the ms-LST-OA, which uses cross-stimulus learning scheme and online adaptation scheme, can achieve higher performance than the current LST method.

C. Ablation Experiment

In the ablation experiment, we compared the performance of LST, ms-LST, LST with online adaptation scheme (LST-OA), and ms-LST-OA at $K = 40$ with different time windows. The results are displayed in Fig. 10. These methods achieve

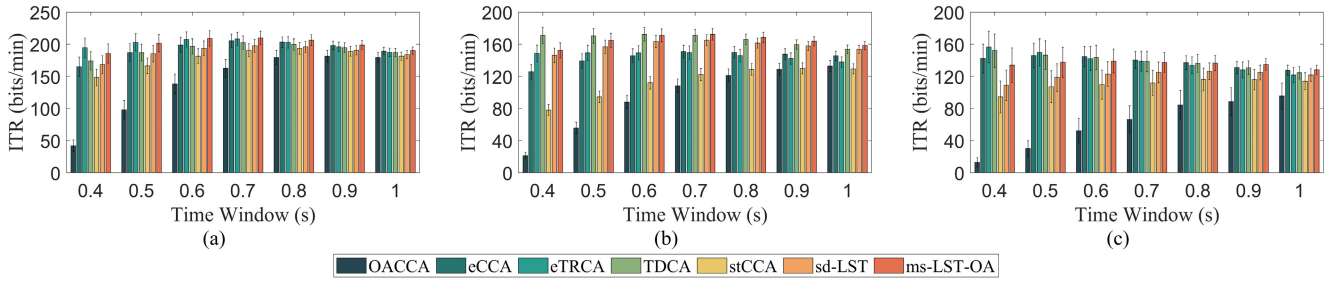

Fig. 11. The averaged ITR across subjects of the different approaches for Dataset I (a), II (b) and III (c) at different time windows. Error bars indicate standard errors.

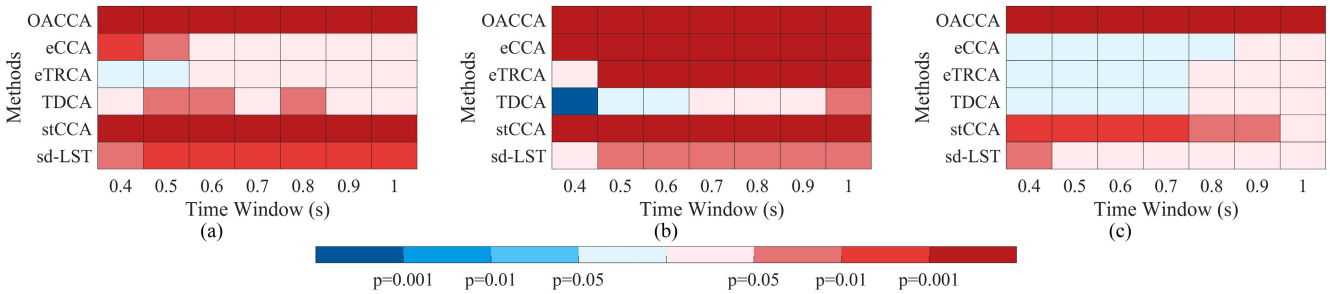

Fig. 12. The paired $t-$ test results of ms-LST-OA and other methods corrected by the Bonferroni method. Where blue (red) color indicates that the ITR of ms-LST-OA is lower (higher) than that of the compared method. And the color depth indicates the level of significant difference between ms-LST-OA and the compared method, with the darker color indicating the more significant difference.

the highest averaged ITR at different T_w (LST: 187.65 $\pm$ 9.73 bits/min at 0.8s; ms-LST: 213.15 $\pm$ 10.86 bits/min at 0.6s; LST-OA: 197.44 $\pm$ 9.30 bits/min at 0.8s; ms-LST-OA: 218.33 $\pm$ 11.23 bits/min at 0.6s). Furthermore, the significant differences were calculated using paired $t-$test and corrected by the Bonferroni method (between LST and ms-LST, LST and LST-OA, ms-LST and LST-OA). Firstly, the averaged ACC and ITR of ms-LST and LST-OA were significantly higher than those of LST for all three datasets ($p < 0.001$ for all three datasets). The results show that the performance of the LST method can be improved using both the cross-stimulus learning scheme and the online adaptation scheme. Secondly, ms-LST had a higher averaged ACC and ITR than LST-OA for all three datasets ($p < 0.001$ at most time windows for Dataset I and II, $p < 0.05$ at time windows from 0.4s to 0.6s for Dataset III, and no significant difference at other time windows). This demonstrates that using the cross-stimulus learning scheme leads to a larger performance improvement than using the online adaptation scheme. Moreover, ms-LST-OA uses online adaptation scheme on the base of ms-LST to gradually optimize the spatial filters and reference templates, which is able to enhance ITR further.

TABLE III
THE HIGHEST AVERAGED ITR OF THE DIFFERENT APPROACHES

Dataset	Approach	T_w (s)	K	${}^{tar}N_{cali}$	ITR (bits/min)
I	OACCA	0.9	0	0	181.41±9.13
	eCCA	0.7	40	5	205.48±10.20
	eTRCA	0.7	40	4	208.10±10.36
	TDCA	0.7	40	2	202.45±10.43
	stCCA	0.8	15	1	193.46±8.95
	sd-LST	0.7	15	1	197.59±10.16
	ms-LST-OA	0.7	15	1	**210.01±10.10**
II	OACCA	1.0	0	0	132.89±6.77
	eCCA	0.7	40	3	150.79±7.98
	eTRCA	0.7	40	3	149.59±8.33
	TDCA	0.6	40	3	172.29±8.26
	stCCA	0.9	15	1	129.88±6.74
	sd-LST	0.7	15	1	164.96±6.87
	ms-LST-OA	0.7	15	1	**172.31±7.26**
III	OACCA	1.0	0	0	95.50±16.10
	eCCA	0.5	12	14	145.96±15.35
	eTRCA	0.4	12	4	156.54±19.61
	TDCA	0.4	12	3	151.88±20.79
	stCCA	0.8	5	1	116.10±14.27
	sd-LST	0.8	5	1	126.20±10.45
	ms-LST-OA	0.6	5	1	**139.04±14.90**

D. The Ability of ms-LST-OA to Reduce Calibration Effort

To measure the ability of ms-LST-OA to reduce the calibration effort, the experiment was conducted to compare the performance of ms-LST-OA with other methods. The averaged ITR across subjects for the different methods are displayed in Fig. 11. Table III shows the highest averaged ITR of the different approaches. It can be found that: 1) In contrast to OACCA, a calibration-free online adaptive method, ms-LST-OA required only a small amount of calibration data to achieve a higher ITR. 2) Comparing eCCA, eTRCA, and TDCA, which are calibration-based methods that require a large amount of calibration data, ms-LST-OA achieved high performance for all three datasets using only a smaller amount of calibration data. 3) Compared to stCCA, sd-LST, which are transfer learning methods, ms-LST-OA achieved a higher ITR for all three datasets using the same amount of data.

Fig. 12 shows the paired $t-$test results of ms-LST-OA and other methods corrected by the Bonferroni method. It can be known that: 1) ms-LST-OA significantly outperformed

OACCA for the three datasets ($p < 0.001$). 2) For all three datasets, ms-LST-OA used less calibration data, but in most cases, there were no significant differences compared to eCCA, eTRCA, and TDCA, which use a large amount of data. This demonstrates the ability of ms-LST-OA to reduce a large calibration effort. 3) For the three datasets, ms-LST-OA was superior to stCCA ($p < 0.001$ for Dataset I and II as well as $p < 0.05$ for Dataset III in most cases). And ms-LST-OA performed better than sd-LST for Dataset I and II, whereas there was essentially no significant difference for Dataset III ($p < 0.01$ mostly for Dataset I and $p < 0.05$ mostly for Dataset II). This indicates that ms-LST-OA is an excellent method for transfer learning.

All results indicate that ms-LST-OA can significantly reduce calibration effort and achieves higher ITR.

VI. Conclusion

To reduce the calibration effort for high-speed SSVEP-based BCIs, this study proposes the ms-LST-OA, which uses cross-stimulus learning scheme and online adaptation scheme to make improvements to the LST. Our experiment results indicate that ms-LST-OA requires only a small amount of calibration data to achieve comparable performance to calibration-based recognition algorithms. This means that ms-LST-OA can reduce a large amount of calibration and achieve a high ITR, which is helpful for the application of SSVEP-based BCI.

Furthermore, the proposed method could be improved. Due to the restricted specifications of the public datasets, more online experiments are needed to determine the relationship between the range of neighboring stimuli (s) and the number of templates (K) at different numbers of stimuli (N_f). Future research will focus on improving the applicability of the proposed method.

References

[1] 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.

[2] T. Yan et al., "Effects of microstate dynamic brain network disruption in different stages of schizophrenia," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 2688–2697, 2023.

[3] 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.

[4] A. S. Albahri et al., "A systematic review of using deep learning technology in the steady-state visually evoked potential-based brain-computer interface applications: Current trends and future trust methodology," *Int. J. Telemedicine Appl.*, vol. 2023, pp. 1–24, Apr. 2023.

[5] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical VEP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 234–240, Jun. 2006.

[6] N. Shi et al., "Steady-state visual evoked potential (SSVEP)-based brain–computer interface (BCI) of Chinese speller for a patient with amyotrophic lateral sclerosis: A case report," *J. Neurorestoratol.*, vol. 8, no. 1, pp. 40–52, Mar. 2020.

[7] Z. Zhang et al., "A flexible speller based on time-space frequency conversion SSVEP stimulation paradigm under dry electrode," *Frontiers Comput. Neurosci.*, vol. 17, pp. 1–11, Feb. 2023.

[8] Z. Tang, X. Wang, J. Wu, Y. Ping, X. Guo, and Z. Cui, "A BCI painting system using a hybrid control approach based on SSVEP and P300," *Comput. Biol. Med.*, vol. 150, Nov. 2022, Art. no. 106118.

[9] D. Zhang et al., "Brain-controlled 2D navigation robot based on a spatial gradient controller and predictive environmental coordinator," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 12, pp. 6138–6149, Dec. 2022.

[10] F. Putze, D. Weiß, L.-M. Vortmann, and T. Schultz, "Augmented reality interface for smart home control using SSVEP-BCI and eye gaze," in *Proc. IEEE Int. Conf. Syst., Man Cybern. (SMC)*, Oct. 2019, pp. 2812–2817.

[11] H. Si-Mohammed et al., "Designing functional prototypes combining BCI and AR for home automation," in *Virtual Reality and Mixed Reality* (Lecture Notes in Computer Science), vol. 13484. Springer, 2022, pp. 3–21.

[12] M. A. Lopez-Gordo, E. Perez, and J. Minguillon, "Gaming the attention with a SSVEP-based brain–computer interface," in *Understanding the Brain Function and Emotions* (Lecture Notes in Computer Science), vol. 11486. Springer, 2019, pp. 51–59.

[13] E. Perez-Valero, M. A. Lopez-Gordo, and M. A. Vaquero-Blasco, "An attention-driven videogame based on steady-state motion visual evoked potentials," *Expert Syst.*, vol. 38, no. 4, pp. 1–12, Jun. 2021.

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

[15] M. Nakanishi, Y. Wang, X. Chen, Y.-T. 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.

[16] 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.

[17] 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.

[18] D. Wu, Y. Xu, and B.-L. Lu, "Transfer learning for EEG-based brain–computer interfaces: A review of progress made since 2016," *IEEE Trans. Cogn. Develop. Syst.*, vol. 14, no. 1, pp. 4–19, Mar. 2022.

[19] K.-J. Chiang, C.-S. Wei, M. Nakanishi, and T.-P. Jung, "Boosting template-based SSVEP decoding by cross-domain transfer learning," *J. Neural Eng.*, vol. 18, no. 1, Feb. 2021, Art. no. 016002.

[20] R. Bian, H. Wu, B. Liu, and D. Wu, "Small data least-squares transformation (sd-LST) for fast calibration of SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 446–455, 2023.

[21] C. M. Wong et al., "Transferring subject-specific knowledge across stimulus frequencies in SSVEP-based BCIs," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 2, pp. 552–563, Apr. 2021.

[22] 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.

[23] 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.

[24] C. M. Wong et al., "Online adaptation boosts SSVEP-based BCI performance," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 6, pp. 2018–2028, Jun. 2022.

[25] 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. 54, no. 6, pp. 1172–1176, Jun. 2007.

[26] 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.

[27] H. Tanaka, T. Katura, and H. Sato, "Task-related component analysis for functional neuroimaging and application to near-infrared spectroscopy data," *NeuroImage*, vol. 64, pp. 308–327, Jan. 2013.

[28] 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.

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

[30] 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.

[31] C. M. Wong et al., "Inter- and intra-subject transfer reduces calibration effort for high-speed SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 10, pp. 2123–2135, Oct. 2020.