Temporally Local Weighting-Based Phase-Locked Time-Shift Data Augmentation Method for Fast-Calibration SSVEP-BCI

Jiayang Huang, Yidan Lv, Zhi-Qiang Zhang, *Member, IEEE*, Bang Xiong, Quan Wang, Bo Wan, and Pengfei Yang, *Member, IEEE*

***Abstract*— Various training-based spatial filtering methods have been proposed to decode steady-state visual evoked potentials (SSVEPs) efficiently. However, these methods require extensive calibration data to obtain valid spatial filters and temporal templates. The time-consuming data collection and calibration process would reduce the practicality of SSVEP-based brain-computer interfaces (BCIs). Therefore, we propose a temporally local weighting-based phase-locked time-shift (TLW-PLTS) data augmentation method to augment training data for calculating valid spatial filters and temporal templates. In this method, the sliding window strategy using the SSVEP response period as a time-shift step is to generate the augmented data, and the time filter which maximises the temporally local covariance between the original template signal and the sine-cosine reference signal is used to suppress the temporal noise in the augmented data. For the performance evaluation, the TLW-PLTS method was incorporated with state-of-the-art training-based spatial filtering methods to calculate classification accuracies and information transfer rates (ITRs) using three SSVEP datasets. Compared with state-of-the-art training-based spatial filtering methods and other data augmentation methods, the proposed TLW-PLTS method demonstrates superior decoding performance with fewer calibration data, which is promising for the development of fast-calibration BCIs.**

***Index Terms*— Brain–computer interfaces (BCIs), data augmentation, phase-locked time-shift (PLTS), steady-state visual evoked potential (SSVEP), temporally local weighting (TLW).**

Manuscript received 7 February 2024; revised 28 May 2024; accepted 22 June 2024. Date of publication 26 June 2024; date of current version 3 July 2024. This work was supported in part by the Key Research and Development Program of Shaanxi under Grant 2024GX-YBXM-039 and Grant 2024GX-ZDCYL-02-15 and in part by the Research and Innovation Funds of Xidian University under Grant YJSJ24015. *(Corresponding authors: Zhi-Qiang Zhang; Pengfei Yang.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Research Ethics Committee of Xidian University.

Jiayang Huang, Yidan Lv, Bang Xiong, Quan Wang, Bo Wan, and Pengfei Yang are with the School of Computer Science and Technology, Xidian University, Xi'an 710126, China (e-mail: huangjiayang@xidian.edu.cn; LvYidan@stu.xidian.edu.cn; bxiong@stu.xidian.edu.cn; qwang@xidian.edu.cn; wanbo@xidian.edu.cn; pfyang@xidian.edu.cn).

Zhi-Qiang Zhang is with the School of Electronic and Electrical Engineering, Institute of Robotics, Autonomous Systems and Sensing, University of Leeds, LS2 9JT Leeds, U.K. (e-mail: z.zhang3@leeds.ac.uk).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSRE.2024.3419013, provided by the authors.

Digital Object Identifier 10.1109/TNSRE.2024.3419013

I. Introduction

BRAIN-COMPUTER interfaces (BCIs) enable individuals to interact with the external environment through direct neural activity [1]. BCIs based on steady-state visual evoked potentials (SSVEPs) are characterised by fast information transfer rates (ITRs) and ease of use [2]. As a result, SSVEP-driven BCIs have found applications in diverse fields, including rehabilitation exoskeleton [3], spelling assistance [4], and robotic operations [5].

For the implementation of SSVEP-based BCI, SSVEP decoding is an essential step to translate human intention into actionable commands [6], [7]. Spatial filtering is one of the efficient decoding methods with low computational complexity and strong robustness [8], which can be classified into two types, training-free and training-based methods. On the one hand, methods such as canonical correlation analysis (CCA) [9], filter bank CCA (FBCCA) [10], multivariate variational mode decomposition CCA (MVMD-CCA) [11], and spatio-spectral CCA (SS-CCA) [12] decode SSVEPs via the maximum similarity between EEG signals and sine-cosine reference signals, are called training-free spatial filtering methods. These methods do not require individual calibration data but achieve a limited information transfer rate. On the other hand, training-based spatial filtering methods, which incorporate individual calibration data, effectively suppress irrelevant brain activity and noise, thus improving SSVEP decoding performance. For example, Wang et al. proposed the extended CCA (eCCA) [13], which combines individual templates and reference signals to compute spatial filters. Nakanishi et al. introduced task-related component analysis (TRCA) into SSVEP-based BCI and proposed ensemble TRCA (eTRCA) [14], which maximise inter-trial covariance to compute spatial filters. In addition, TRCA-based extensions improve decoding performance by exploiting additional information, such as phase information [15], neighbouring stimuli [16], and similarity constraints [17]. In contrast to the TRCA-based methods, Liu et al. proposed the task discriminant component analysis (TDCA) method [18], which decodes

© 2024 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License.
For more information, see https://creativecommons.org/licenses/by/4.0/

SSVEP with a discriminative model. The training-based spatial filtering methods significantly improve SSVEP decoding performance, but such methods require massive individual calibration data to compute valid spatial filters and temporal templates. The tedious collection and calibration session would result in a high time cost and severe visual fatigue in subjects, thus reducing the practicality of SSVEP-based BCIs.

To address the above problem of high calibration cost, transfer learning emerges as a viable approach [19], which facilitates the transfer of knowledge across domains to explore the feasibility of SSVEP decoding with less SSVEP calibration data [20]. According to different transfer domains, the transfer learning methods can be categorised into cross-subject transfer methods [21], [22], cross-device transfer methods [23], [24], and cross-stimulus transfer methods [25], [26]. However, the decoding performance of these transfer learning methods would show limited or even negative improvement when the source data is dissimilar to the target data [27], [28]. Moreover, the calibration data from source domains hold the potential to be further reduced. Therefore, existing transfer learning methods do not provide a fundamental solution to the problem of high calibration costs in SSVEP-based BCIs.

In addition to transfer learning methods, data augmentation is a promising solution to the above problem, augmenting the training data with artificial signals [29]. For example, Fahimi et al. introduced a generative adversarial network (GAN) to the BCI system for training data augmentation [30]. Kwon and Im proposed a starGAN method to generate SSVEP signals [31]. However, such methods based on neural-based models also require massive calibration data to optimise the model parameters. Other data augmentation methods do not require massive calibration data. One way is to generate augmented data epochs by performing data transformation. For example, Li et al. proposed three sample-based data augmentation methods based on electroencephalogram (EEG) signal features [32]. Luo et al. proposed the source aliasing matrix estimation (SAME) method to augment the SSVEP calibration data [33]. Moreover, an easier way to generate training data is to obtain augmented data using the sliding window strategy, called the time-shift data augmentation method. Previous time-shifting methods set the time-shift step to a fixed length of 50 or 100 ms [34], [35], which violates the time-locked feature of SSVEPs, resulting in low SSVEP decoding performance. To preserve the temporal characteristics of SSVEPs, Mai et al. proposed the phase-locked time-shift (PLTS) method [36], which sets the time-shift step to the SSVEP oscillation period. Xiao et al. proposed a data expansion method called cyclic shift trials (CST), which cyclically shifts the SSVEP oscillation period to generate new trials [37]. However, the performance improvement achieved by these methods mainly depends on valid spatial filters rather than valid temporal templates. A valid temporal template should be further investigated using high SNR augmented data epochs. Temporally local covariances can extract the discriminative temporal structure from EEG signals [38]. Therefore, the signal-to-noise ratio (SNR) of the augmented data is expected to improve when the temporally local information of SSVEPs is taken into account [39], [40], [41].

In this paper, a temporally local weighting-based phase-locked time-shift (TLW-PLTS) data augmentation method is proposed. First, a sliding window strategy using the SSVEP response period as a time-shift step is used to generate the augmented data. Second, a time filter is computed that maximises the temporally local correlation between the template and the reference signals to suppress the temporal noise in the augmented signal. Finally, the time-filtered augmented data is added to the original training data set to train the spatial filter and compute the new template signal. The TLW-PLTS method was verified on the state-of-the-art training-based spatial filtering methods using the benchmark and the self-recorded dataset.

The rest of the article is arranged as follows: The materials and methods are introduced in Section II. In section III, the experimental results are presented. Section IV presents the discussion part. Finally, the conclusion is presented in Section V.

II. Materials and Methods

A. Data Description

1) Dataset I: The first dataset is the benchmark dataset collected by Tsinghua University [42]. Dataset I consists of 64-channel data corresponding to 40 targets from 35 healthy subjects. Using the joint frequency and phase modulation (JFPM) approach [43], 40 targets are stimulated at frequencies of 8-15.8 Hz with an interval of 0.2 Hz and 0-1.5π phases with an interval of 0.5π. The collected SSVEP data were downsampled to 250 Hz. The data collection includes 6 data blocks for each subject. Each block has 40 trials, which corresponds to the 40 targets. A 0.5-s gaze shift cue, a 5-s stimulus, and a 0.5-s rest are included in each trial. For the SSVEP signal analysis in this study, 9 channels (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, and O2) were selected.

2) Dataset II: Another publicly available dataset collected by the University of California San Diego (UCSD) is also utilized in this study [44]. The dataset contains 15-block SSVEP data from 10 subjects. For each block, 12-trial data correspond to 12 stimuli, each comprising a 1-s gaze shifting cue and a 4-s gazing time. Each stimulus is modulated by the JFPM method, where frequency ranges from 9.25 Hz to 14.75 Hz, with a 0.5 Hz interval, and the phases range from 0π to 1.5π, with a 0.5π interval. The 8-electrode settings are as PO7, PO3, POz, PO4, PO8, O1, Oz, and O2.

3) Dataset III: In addition, a 12-stimuli SSVEP dataset is collected by the offline BCI experiment. The Xidian University research ethics committee gave its approval to the BCI experiment. Before the experiment began, each subject completed an informed permission form after being informed about the procedures. In the experiment, 9-channel (Pz, PO7, PO3, POz, PO4, PO8, O1, Oz, and O2) EEG data were collected from 12 healthy participants (5 women and 7 men, aged 24-27 years). The JFPM method is used to code the 12 visual stimuli. The frequency ranges from 9.25 Hz to 14.75 Hz, with a 0.5 Hz interval, and the phases range from 0π to 1.5π, with a 0.5π interval. 10 blocks of EEG data were collected from each subject. Each trial contains 0.5 s cue, 5 s stimulation, and

0.5 s rest. The subjects were instructed not to blink while being stimulated. There was a two-minute break in between every two blocks to prevent visual fatigue.

B. Data Preprocessing

The first 0.64 s of the EEG data were eliminated since the data also included a 0.5 s gaze shift cue and a 0.14 s visual latency. As a result, the 9-channel EEG data epochs were recovered as [0.64 0.64+d] s, where d is the length of the calibration data. Each extracted data epoch was then filtered with a 6th order 7-90 Hz Butterworth filter. A 50 Hz notch filter was employed to eliminate power line interference. Upon completion of the pre-processing stage, the subsequent data augmentation and SSVEP decoding procedures were performed.

C. Temporally Local Weighting-Based Phase-Locked Time-Shift (TLW-PLTS) Data Augmentation Method

The temporally local weighting-based phase-locked time-shift (TLW-PLTS) method mainly consists of four steps, including 1) augmented data epoch extraction, 2) time filter calculation, 3) spatial filter training, and 4) SSVEP decoding. The first two steps are the data augmentation process, as shown in Fig. 1, and the last two steps are for data augmentation validation.

1) Augmented Data Epoch Extraction: Denote the calibration data from each subject of the n-th stimulus as $\boldsymbol{\mathcal{X}}_n = \left[\boldsymbol{\mathcal{X}}_n^1, \boldsymbol{\mathcal{X}}_n^2, \cdots, \boldsymbol{\mathcal{X}}_n^{N_t}\right] \in \mathbb{R}^{N_c \times N_d \times N_t}$ $(n = 1, 2, \cdots, N_f)$. Here, N_c represents the number of channels, N_d is the number of sampling points, N_t is the number of calibration trials, and N_f is the number of stimuli. And $N_d = d \times fs$, fs represents the sampling rate. $\boldsymbol{\mathcal{X}}_n^h \in \mathbb{R}^{N_c \times N_d}$ $(h = 1, 2, \cdots, N_t)$ denotes the h-th trial data of $\boldsymbol{\mathcal{X}}_n$. By averaging across trials, the original template signal $\overline{\boldsymbol{\mathcal{X}}}_n \in \mathbb{R}^{N_c \times N_d}$ corresponding to the n-th stimulus is defined as:

$$\overline{\boldsymbol{\mathcal{X}}}_n = \frac{1}{N_t} \sum_{h=1}^{N_t} \boldsymbol{\mathcal{X}}_n^h. \tag{1}$$

First, the data epoch $\boldsymbol{\mathcal{X}}_{\text{org}} \in \mathbb{R}^{N_c \times N_s}$ is extracted as $[0, tw]$ s from $\overline{\boldsymbol{\mathcal{X}}}_n$. Here, tw is the time window, $N_s = tw \times fs$ is the corresponding number of sampling points, and $N_d > N_s$. And then, to ensure retain the time- and phase-locked characteristics of SSVEP signals, the time-shift step Δt_n is set as the response period of n-th stimulus:

$$\Delta t_n = \frac{1}{f_n}. \tag{2}$$

In this way, the onset of the k-th trial augmented data corresponding to the n-th stimulus σ_{aug}^k is defined as:

$$\sigma_{\text{aug}}^k = k \times \Delta t_n, \tag{3}$$

where $k = 1, 2, \cdots, N_k$ and N_k is the number of augmented trials. Therefore, the k-th trial augmented data $\boldsymbol{\mathcal{X}}_{\text{aug}}^k \in \mathbb{R}^{N_c \times N_s}$ is extracted as $[\sigma_{\text{aug}}^k, \sigma_{\text{aug}}^k + tw]$ s from $\overline{\boldsymbol{\mathcal{X}}}_n$.

2) Time Filter Calculation: The calculation of the time filter is based on the sine-cosine reference signals to remove the irrelevant temporal noise [38]. First, the sine-cosine reference signal $\boldsymbol{R}_n \in \mathbb{R}^{2N_H \times N_s}$ corresponding to each stimulus frequency f_n is constructed as:

$$\boldsymbol{R}_n = \begin{pmatrix} \sin(2\pi f_n t/fs) \\ \cos(2\pi f_n t/fs) \\ \vdots \\ \sin(2N_H \pi f_n t/fs) \\ \cos(2N_H \pi f_n t/fs) \end{pmatrix}, t = 1, 2, \ldots, N_s, \tag{4}$$

where N_H represents the number of harmonic frequencies. Here, N_H is set as 5. The time filter $\boldsymbol{\Psi}_n \in \mathbb{R}^{N_c \times N_c}$ is calculated by maximising the temporally local correlation between $\boldsymbol{\mathcal{X}}_{\text{org}}$ and the sine-cosine reference signal $\boldsymbol{R}_n$.

Considering the temporal structure information of $\boldsymbol{\mathcal{X}}_{\text{org}}$ to remove the irrelevant noise in the time domain [40], the definition of $\boldsymbol{\mathcal{X}}_{\text{org}}$ is given as:

$$\boldsymbol{\mathcal{X}}_{\text{org}} = [\boldsymbol{x}(1), \boldsymbol{x}(2), \cdots, \boldsymbol{x}(N_s)], \tag{5}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^{N_c \times 1}$represents the multi-channel signal corresponding to each sampling points t. Define the adjacency matrix $\boldsymbol{A} \in \mathbb{R}^{N_s \times N_s}$ as:

$$\boldsymbol{A}(t_1, t_2) = K\left(\frac{t_1 - t_2}{\varepsilon}\right), \tag{6}$$

where the value of temporally local weight $K\left(\frac{t_1 - t_2}{\varepsilon}\right)$ is determined by the Tukey's tricube weighting function [45]:

$$K\left(\frac{t_1 - t_2}{\varepsilon}\right) = \begin{cases} \left(1 - \left|\frac{t_1 - t_2}{\varepsilon}\right|^3\right)^3, & |\frac{t_1 - t_2}{\varepsilon}| < 1, \\ 0, & \text{else.} \end{cases} \tag{7}$$

Since SSVEPs are stable periodic signals, one period contains complete task-related information. The parameter ε that limits the temporally local range is defined as $\varepsilon = \frac{fs}{2f_n}$. Thus, with Eq. (6), the temporally local covariance matrix $\boldsymbol{C}_{\boldsymbol{xx}}$ of $\boldsymbol{\mathcal{X}}_{\text{org}}$ is calculated as:

$$\begin{aligned} \boldsymbol{C}_{\boldsymbol{xx}} =& \frac{1}{N_s}\left(\sum_{t=1}^{N_s} \boldsymbol{x}(t)^\top \boldsymbol{x}(t) \sum_{t_2=1}^{N_s} \boldsymbol{A}(t_1, t_2) \right. \\ & \left. - \sum_{t_1=1}^{N_s}\sum_{t_2=1}^{N_s} \boldsymbol{A}(t_1, t_2)\boldsymbol{x}(t_1)^\top \boldsymbol{x}(t_2)\right) \\ =& \frac{1}{N_s}\left(\boldsymbol{\mathcal{X}}_{\text{org}}(\boldsymbol{D} - \boldsymbol{A})\boldsymbol{\mathcal{X}}_{\text{org}}^\top\right) \\ =& \frac{1}{N_s}\boldsymbol{\mathcal{X}}_{\text{org}}\boldsymbol{L}\boldsymbol{\mathcal{X}}_{\text{org}}^\top, \end{aligned} \tag{8}$$

where $\boldsymbol{D} \in \mathbb{R}^{N_s \times N_s}$ is a diagnal matrix defined as:

$$\boldsymbol{D} = \text{diag}\left(\sum_{t_2=1}^{N_s} \boldsymbol{A}(t_1, t_2)\right), \tag{9}$$

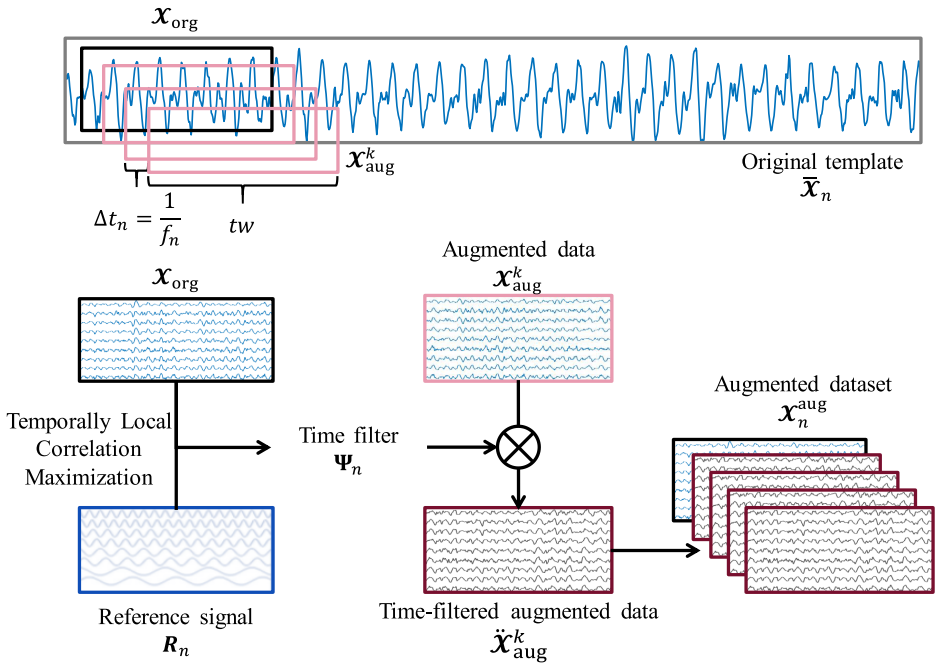


Fig. 1. The illustration of the temporally local weighting-based phase-locked time-shift (TLW-PLTS) data augmentation method, contains two steps: 1) augmented data epoch extraction and 2) time filter calculation.

$\boldsymbol{L} \in \mathbb{R}^{N_s \times N_s}$ is a Laplacian matrix, calculated as $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{A}$. In the same manner, the temporally local covariance matrix $\boldsymbol{C_{rr}}$ of the reference signal $\boldsymbol{R}_n$ is calculated as:

$$\boldsymbol{C_{rr}} = \frac{1}{N_s}\boldsymbol{R}_n\boldsymbol{L}\boldsymbol{R}_n^\top. \tag{10}$$

The temporally local covariance matrices $\boldsymbol{C_{xr}}$ and $\boldsymbol{C_{rx}}$ between $\boldsymbol{\mathcal{X}}_{\text{org}}$ and $\boldsymbol{R}_n$ are calculated as:

$$\boldsymbol{C_{xr}} = \frac{1}{N_s}\boldsymbol{\mathcal{X}}_{\text{org}}\boldsymbol{L}\boldsymbol{R}_n^\top, \tag{11}$$

$$\boldsymbol{C_{rx}} = \frac{1}{N_s}\boldsymbol{R}_n\boldsymbol{L}\boldsymbol{\mathcal{X}}_{\text{org}}^\top. \tag{12}$$

The basis vector $\boldsymbol{\psi} \in \mathbb{R}^{N_c \times N_c}$ is estimated by maximizing the temporally local correlation between $\boldsymbol{\mathcal{X}}_{\text{org}}$ and the reference signal $\boldsymbol{R}_n$. The optimization problem is defined as:

$$\hat{\boldsymbol{\psi}}_n = \arg\max_{\boldsymbol{\psi}} \boldsymbol{\psi}^\top(\boldsymbol{C_{xx}}^{-1}\boldsymbol{C_{xr}}\boldsymbol{C_{rr}}^{-1}\boldsymbol{C_{rx}})\boldsymbol{\psi}, \tag{13}$$

By solving the Eq. (13) via generalized eigenvalue decomposition, the basis vector $\boldsymbol{\psi}$ is obtained. The time filter corresponding to each stimulus $\boldsymbol{\Psi}_n$ is calculated as:

$$\boldsymbol{\Psi}_n = \boldsymbol{\psi}_n\boldsymbol{B}\boldsymbol{\psi}_n^{-1}, \tag{14}$$

where $\boldsymbol{B} \in \mathbb{R}^{N_c \times N_c}$ is a diagonal matrix defined as $\boldsymbol{B}(c,c) = 1$ $(c = 1, 2, \cdots, N_c)$, and c represents the index of each channel. The diagonal element $\boldsymbol{B}(c,c)$ is 1 if the c-th channel of signal needs to be time-filtered and zero otherwise [39]. With the time filter $\boldsymbol{\Psi}_n$, the time-filtered augmented data is computed as:

$$\ddot{\boldsymbol{\mathcal{X}}}^k_{\text{aug}} = \boldsymbol{\Psi}_n^\top\boldsymbol{\mathcal{X}}^k_{\text{aug}}. \tag{15}$$

The final augmented dataset corresponding to each stimulus $\boldsymbol{\mathcal{X}}^{\text{aug}}_n \in \mathbb{R}^{N_c \times N_s \times (N_k+1)}$ is the combination of $\boldsymbol{\mathcal{X}}_{\text{org}}$ and $\ddot{\boldsymbol{\mathcal{X}}}^k_{\text{aug}}$ $(k = 1, 2, \cdots, N_k)$, defined as:

$$\boldsymbol{\mathcal{X}}^{\text{aug}}_n = [\boldsymbol{\mathcal{X}}_{\text{org}}, \ddot{\boldsymbol{\mathcal{X}}}^1_{\text{aug}}, \ddot{\boldsymbol{\mathcal{X}}}^2_{\text{aug}}, \cdots, \ddot{\boldsymbol{\mathcal{X}}}^{N_k}_{\text{aug}}]. \tag{16}$$

Therefore, the augmented training set corresponding to each stimulus $\check{\boldsymbol{\mathcal{X}}}_n \in \mathbb{R}^{N_c \times N_s \times N_{\text{aug}}}$ is obtained as:

$$\check{\boldsymbol{\mathcal{X}}}_n = [\boldsymbol{\mathcal{X}}^{\text{o}}_n, \boldsymbol{\mathcal{X}}^{\text{aug}}_n], \tag{17}$$

where $\boldsymbol{\mathcal{X}}^{\text{o}}_n \in \mathbb{R}^{N_c \times N_s \times N_t}$ is the original training set, which is extracted as $[0, tw]$ s from $\boldsymbol{\mathcal{X}}_n$. N_{aug} is the number of trials in the augmented training set. When $N_t = 1$, $N_{\text{aug}} = N_t + N_k$; when $N_t > 1$, $N_{\text{aug}} = N_t + N_k + 1$.

With the augmented training set, the training-based spatial filtering methods [8] can be conducted to learn spatial filters and then decode SSVEPs. In this study, the spatial filtering method incorporating neighboring stimuli data (termed INS-SF) [16] is used for data augmentation method verification.

3) Spatial Filter Training: The training data corresponding to the n-th stimulus $\check{\boldsymbol{\mathcal{X}}}_n$ is denoted as:

$$\check{\boldsymbol{\mathcal{X}}}_n = \left[\check{\boldsymbol{\mathcal{X}}}^1_n, \check{\boldsymbol{\mathcal{X}}}^2_n, \cdots, \check{\boldsymbol{\mathcal{X}}}^{N_{\text{aug}}}_n\right], \tag{18}$$

where $\check{\boldsymbol{\mathcal{X}}}^\kappa_n$ $(\kappa = 1, 2, \cdots, N_{\text{aug}})$ is the κ-th trial of the augmented training set. Denote the training data corresponding to the neighboring stimuli as:

$$\check{\mathcal{N}}_n = \left\{\check{\boldsymbol{\mathcal{X}}}_{n^1}, \check{\boldsymbol{\mathcal{X}}}_{n^2}, \cdots, \check{\boldsymbol{\mathcal{X}}}_{n^{N_l}}\right\}, \tag{19}$$

which are defined as the horizontal and vertical neighboring stimuli data.

First, the training data $\check{\mathcal{X}}_n^{\kappa}$ are decomposed into N_j subbands using filter bank analysis [10], as $\check{\mathcal{X}}_n^{\kappa,j}$ ($j = 1, 2, \cdots, N_j$). With the training data, the spatial filter $\boldsymbol{w}_n^j \in \mathbb{R}^{N_c}$ can be obtained as:

$$\hat{\boldsymbol{w}}_n^j = \arg\max_{\boldsymbol{w}} \frac{\boldsymbol{w}^\top \boldsymbol{S} \boldsymbol{w}}{\boldsymbol{w}^\top \boldsymbol{Q} \boldsymbol{w}}. \tag{20}$$

The sum of the covariances of $\tilde{\mathcal{X}}_n^j \in \mathbb{R}^{N_c \times N_d}$ and $\tilde{\mathcal{X}}_{n^l}^j \in \mathbb{R}^{N_c \times N_d}$ is calculated as:

$$\boldsymbol{S} = \operatorname{cov}\left(\tilde{\mathcal{X}}_n^j\right) + \sum_{l=1}^{N_l} \operatorname{cov}\left(\tilde{\mathcal{X}}_{n^l}^j\right), \tag{21}$$

where

$$\tilde{\mathcal{X}}_n^j = \frac{1}{N_{\text{aug}}} \sum_{\kappa=1}^{N_{\text{aug}}} \check{\mathcal{X}}_n^{\kappa,j}, \tag{22}$$

$$\tilde{\mathcal{X}}_{n^l}^j = \frac{1}{N_{\text{aug}}} \sum_{\kappa=1}^{N_{\text{aug}}} \check{\mathcal{X}}_{n^l}^{\kappa,j}, \tag{23}$$

The sum of covariances of all-trial data from the target and its neighboring stimuli is denoted as:

$$\boldsymbol{Q} = \sum_{\kappa=1}^{N_{\text{aug}}} \operatorname{cov}\left(\check{\mathcal{X}}_n^{\kappa,j}\right) + \sum_{l=1}^{N_l} \sum_{\kappa=1}^{N_{\text{aug}}} \operatorname{cov}\left(\check{\mathcal{X}}_{n^l}^{\kappa,j}\right). \tag{24}$$

Eq. (20) can be solved by generalised eigendecomposition of $\boldsymbol{Q}^{-1}\boldsymbol{S}$, where the spatial filter $\hat{\boldsymbol{w}}_n^j$ as the eigenvector corresponding to the largest eigenvalue.

4) SSVEP Decoding: With $\boldsymbol{w}_n^j$, the test data $\boldsymbol{Y} \in \mathbb{R}^{N_c \times N_s}$ can be decoded. Firstly, $\boldsymbol{Y} \in \mathbb{R}^{N_c \times N_s}$ is decomposed into N_j subbands with filter bank analysis, as $\boldsymbol{Y}^j$ ($j = 1, 2, \cdots, N_j$). Then $\boldsymbol{Y}^j$ and the n-th template $\tilde{\mathcal{X}}_n^j$ are respectively spatially filtered with $\boldsymbol{w}_n^j$. For each subband j, the correlation coefficient between the spatially filtered test data and template is calculated as:

$$\gamma_n^j = \rho\left(\left(\boldsymbol{Y}^j\right)^\top \boldsymbol{w}_n^j, \left(\tilde{\mathcal{X}}_n^j\right)^\top \boldsymbol{w}_n^j\right), \tag{25}$$

where $\rho(s_1, s_2)$ is the Pearson's correlation coefficient between s_1 and s_2 [46]. The ensemble correlation feature γ_n is calculated as:

$$\gamma_n = \sum_{j=1}^{N_j} \omega(j) \cdot \left(\gamma_n^j\right)^2, \tag{26}$$

where the weights for each subband component $\omega(j) = j^{-1.25} + 0.25$ is to maximise the decoding performance [10]. The target frequency $\hat{f}$ corresponding to the largest correlation coefficient is determined as:

$$\hat{f} = \arg\max_n \gamma_n, n = 1, 2, \cdots, N_f. \tag{27}$$

The pseudo-code for the TLW-PLTS data augmentation method including augmentation, model training, and decoding is described as Algorithm 1.

Algorithm 1 TLW-PLTS Data Augmentation Method

Training

Input: The calibration data $\mathcal{X}_n \in \mathbb{R}^{N_c \times N_d \times N_t}$ and reference signal $\boldsymbol{R}_n \in \mathbb{R}^{N_H \times N_s}$ corresponding to each stimulus n ($n = 1, 2, \cdots, N_f$).

Output: The spatial filter $\boldsymbol{w}_n^j$ and augmented template $\tilde{\mathcal{X}}_n^j$ corresponding to each stimulus n, subband j ($j = 1, 2, \cdots, N_j$).

1. Data augmentation

for $n = 1, 2, \cdots, N_f$ **do**

(1) Average $\mathcal{X}_n$ across trials to obtain template $\overline{\mathcal{X}}_n$;

(2) Extract original template $\mathcal{X}_{\text{org}}$ from $\overline{\mathcal{X}}_n$;

(3) Calculate time filter $\boldsymbol{\Psi}_n$;

for $k = 1, 2, \cdots, N_k$ **do**

(4) Extract augmented data epoch $\mathcal{X}_{\text{aug}}^k$ from $\overline{\mathcal{X}}_n$;

(5) Calculate the time-filtered augmented data $\ddot{\mathcal{X}}_{\text{aug}}^k$;

(6) Append $\ddot{\mathcal{X}}_{\text{aug}}^k$ into $\check{\mathcal{X}}_n$;

end for

end for

2. Spatial filter training

for $n = 1, 2, \cdots, N_f$ **do**

(1) Define neighboring stimuli data $\check{\mathcal{N}}_n$;

(2) Decompose $\check{\mathcal{X}}_n$ into $\check{\mathcal{X}}_n^j$ ($j = 1, 2, \cdots, N_j$) using filter bank analysis;

for $j = 1, 2, \cdots, N_j$ **do**

(3) Average $\check{\mathcal{X}}_n^j$ across trials to obtain augmented template $\tilde{\mathcal{X}}_n^j$;

(4) Calulate spatial filter $\boldsymbol{w}_n^j$;

end for

end for

return $\boldsymbol{w}_n^j$, $\tilde{\mathcal{X}}_n^j$

Testing

Input: The single-trial test data $\boldsymbol{Y} \in \mathbb{R}^{N_c \times N_s}$, the spatial filter $\boldsymbol{w}_n^j \in \mathbb{R}^{N_c}$ and augmented template $\tilde{\mathcal{X}}_n^j \in \mathbb{R}^{N_c \times N_s}$.

Output: Predicted target SSVEP frequency $\hat{f}$.

1. Decompose $\boldsymbol{Y}$ into $\boldsymbol{Y}^j$ using filter bank analysis;

for $j = 1, 2, \cdots, N_j$ **do**

2. Calculate correlation coefficient γ_n^j;

end for

3. Compute the final correlation feature γ_n;

4. Determine target SSVEP frequency $\hat{f}$;

return $\hat{f}$

III. Experimental Results

A. Performance Evaluations

1) Baseline Methods: In this study, the INS-SF method [16] with (w/) and without the different data augmentation methods (CST [37], SAME [33], PLTS [36], and the proposed TLW-PLTS method) were compared to evaluate the performance of TLW-PLTS. The detailed parameter settings of baseline methods are as follows. The detailed description of the parameter selection process for each baseline method is stated in Supplementary Materials.

- INS-SF: For Dataset I, the number of subbands was set as 5. For Dataset II and III, the number of subbands was set as 3.
- INS-SF (w/ CST): Both training and test data were augmented using CST. The numbers of subbands were set in the same way as INS-SF. For Dataset I, the number of augmented trials was set to 5. For Datasets II and III, the numbers of augmented trials were set to 10.
- INS-SF (w/ SAME): The numbers of subbands were set the same as INS-SF. As N_t increased from 1 to 5, the corresponding numbers of augmented trials N_a were set as 3, 4, 5, 5, and 7 respectively for Dataset I; 4, 4, 4, 7, and 7 for Dataset II; 5, 6, 7, 7, and 9 for Dataset III.
- INS-SF (w/ PLTS): The numbers of subbands were set the same as INS-SF. When N_t increased from 1 to 5, the corresponding numbers of augmented trials N_a were set as 7, 5, 3, 1, and 1 respectively for Dataset I; 8, 5, 4, 3, and 3 for Dataset II; 8, 6, 6, 5, and 4 for Dataset III.

2) Evaluation Metrics: In this study, the SSVEP decoding performance of the INS-SF method using the TLW-PLTS data augmentation approach were estimated using classification accuracy and ITR as evaluation metrics. Classification accuracy is the fraction of accurate predictions among all predictions. ITR is the amount of information transferred per minute, which can be expressed as follows:

$$\mathrm{ITR}=\frac{60}{T}\times\left[\log_2 N_f+P\times\log_2 P+(1-P)\times\log_2\left(\frac{1-P}{N_f-1}\right)\right], \tag{28}$$

where P denotes the classification accuracy and T is the total of the time window and the gaze-shifting time of 0.5 s.

3) Validation Method: For the offline experiments, the classification accuracy and ITR estimates were calculated by using leave-out cross-validation. For each dataset, the number of calibration trials N_t varied from 1 to 5, the other trials were used as test trials. The numbers of augmented trials N_k for three datasets were set as 5, 7, and 5 respectively, the settings of which were discussed in subsection E. For all datasets, the time window increased from 0.3 s to 1.0 s, with a step of 0.1 s.

For the online experiment, all the subjects in Dataset III were invited to participate in a cue-guided spelling experiment using a 12-target stimulation interface. They were instructed to spell 30 random characters following the red-dot cue. The spatial filter and template were learned from 5-trial calibration data and 5-trial augmented data according to the optimal parameters selected from the offline experiments using Dataset III. The time window was set to 0.6 s, which corresponded to the highest ITR in the offline experiment.

4) Statistical Analysis: All experimental results were stated as the mean values of the classification accuracies and ITRs unless otherwise stated. The error bars presented in different figures denoted the standard error of the mean (SEM). Furthermore, paired-t tests were conducted between baseline methods and the proposed TLW-PLTS method to show the significance level. The asterisks shown in the figures indicate significant differences between the proposed method and other baseline methods obtained by paired t-tests ($*p<0.05$, $**p<0.01$, $***p<0.001$, $****p<0.0001$).

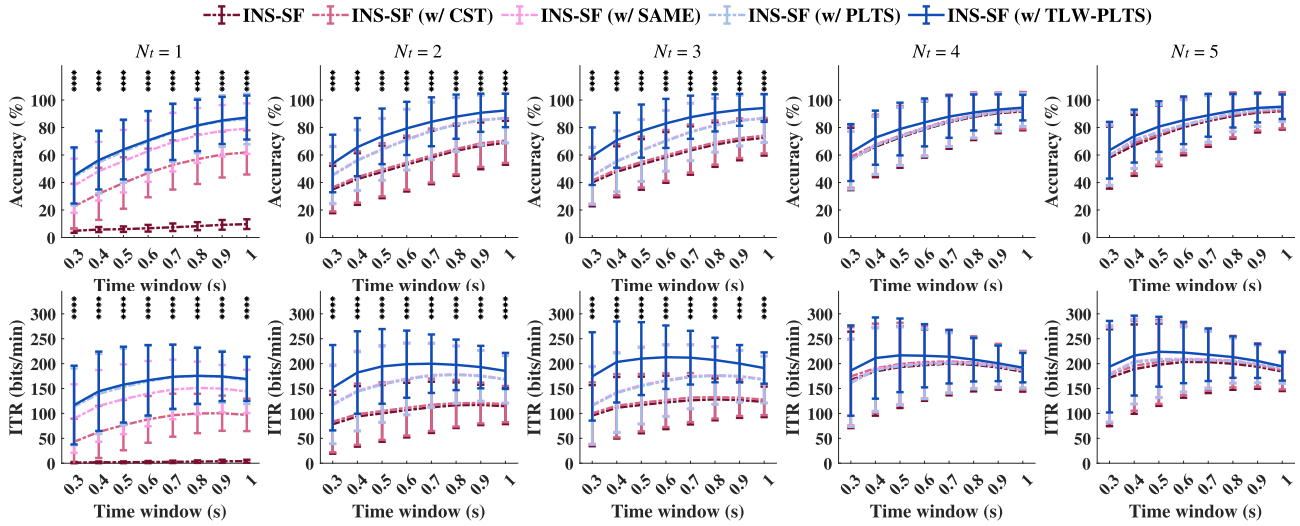
B. The Overall Offline SSVEP Decoding Performance

The classification accuracies and information transfer rates averaged across subjects, calculated using three datasets, are shown in Fig. 2. When $N_t = 1$, compared to the original INS-SF method, the SSVEP decoding performances are significantly improved after using the three data augmentation methods of CST, SAME, PLTS, and TLW-PLTS, among which the CST method has the least improvement, and the TLW-PLTS method has the most improvement. Here there is not much difference between the PLTS and TLW-PLTS methods. When $N_t = 2$, the decoding performance obtained by the original INS-SF method is similar to that obtained by the INS-SF (w/ CST), while INS-SF (w/ SAME) and INS-SF (w/ PLTS) have similar decoding performance. With TLW-PLTS, the classification accuracy and the information transfer rate are significantly higher than the original INS-SF method. Compared to the SAME and PLTS methods, the classification accuracy and information transfer rate of the TLW-PLTS method are also superior. When $N_t = 3$, the decoding performance obtained by INS-SF is also similar to that obtained by using the CST method. INS-SF (w/ SAME) and INS-SF (w/ PLTS) also obtain similar performance. For Dataset I and II, using the TLW-PLTS method improves the classification accuracy and the information transfer rate compared to INS-SF and INS-SF (w/ CST). For Dataset II, there is no significant difference between the five competing methods. When $N_t = 4$ or $N_t = 5$, there is little difference in decoding performance between the four cases for Dataset I and II. For Dataset III, the difference between the other four baseline methods decreases as the number of calibration trials increases, while the proposed method still outperforms the other methods.

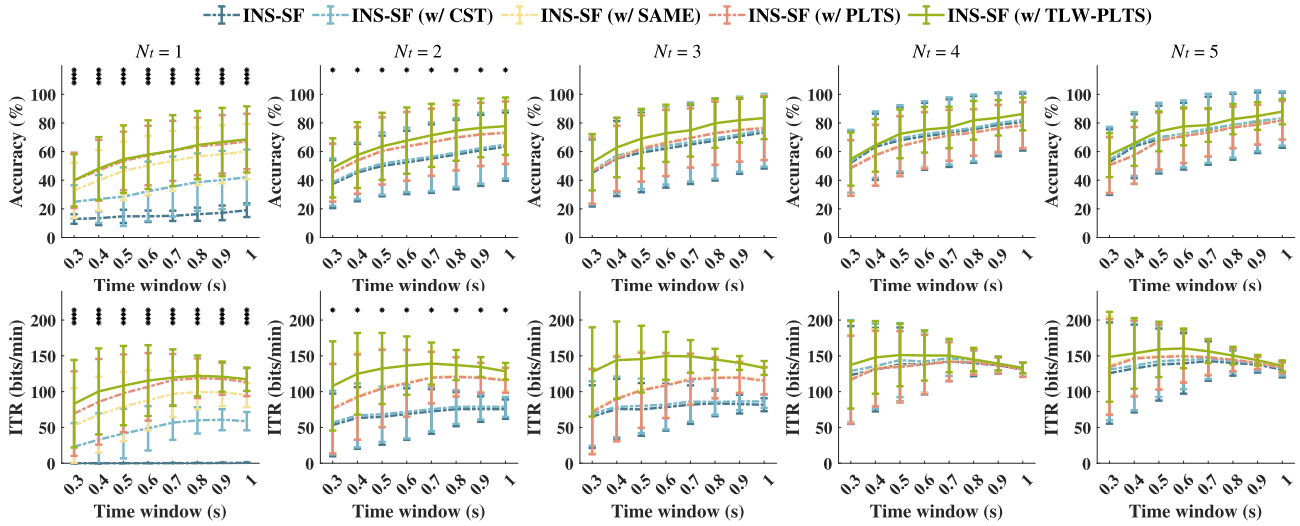
In conclusion, the experimental results estimated by all datasets show that using the TLW-PLTS data augmentation method can effectively suppress the time-domain noise and improve the signal-to-noise of the augmented data, thus effectively improving the decoding performance of SSVEP with small training sets.

C. Offline Performance Under Single-Trial Calibration Data

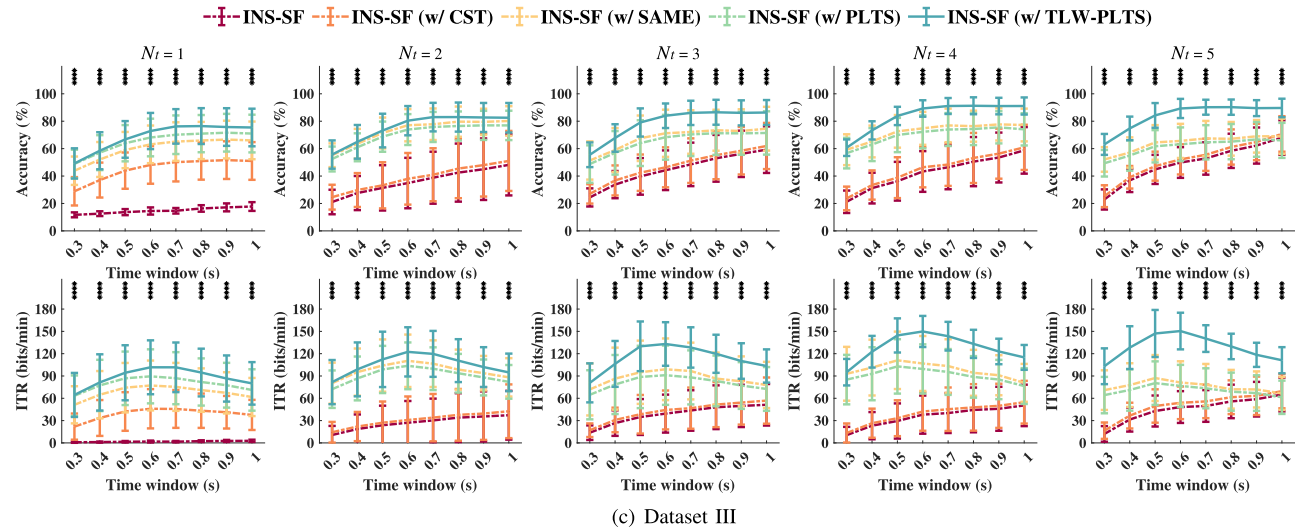
In addition, the decoding performance of the original INS-SF method and that of different data augmentation methods under single-trial calibration data are also compared to verify the superiority of the proposed TLW-PLTS method. Table I reports the maximum averaged ITRs, calibration and training time costs of the five conditions using Dataset I, including INS-SF, INS-SF (w/ CST), INS-SF (w/ SAME), INS-SF (w/ PLTS), INS-SF (w/ TLW-PLTS). Here, 5-trial calibration data were used to obtain the maximum averaged ITR of the INS-SF method. The other four conditions with data augmentation methods used single-trial calibration data. The experiments were implemented using Matlab 2023a on a Lenovo PC equipped with an Intel(R) Xeon(R)



(a) Dataset I



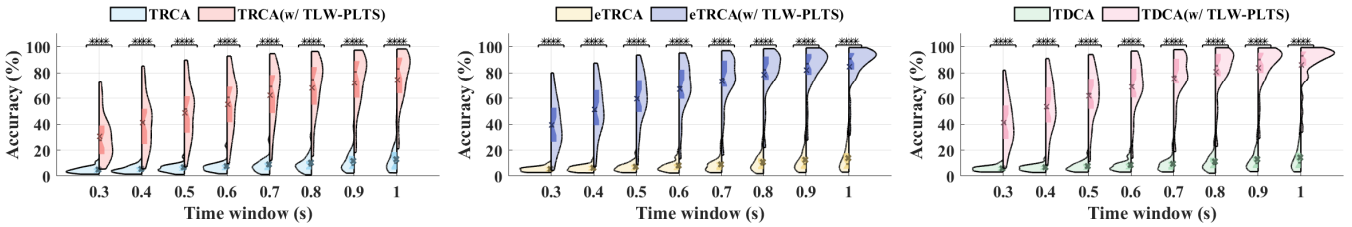
(b) Dataset II



(c) Dataset III

Fig. 2. Averaged accuracies and ITRs across subjects of the INS-SF method with TLW-PLTS method and other baseline methods using different time windows varying from 0.3 s to 1.0 s with a step of 0.1 s on Dataset I (a), Dataset II (b), and Dataset III (c).

Silver 4116 CPU @ 2.10GHz, 32GB RAM, and 64-bit Windows 10 OS. The calibration time for each subject represents the total length of calibration data corresponding to all stimuli. The training time represents the sum of the computational time



(a) TRCA (b) eTRCA (c) TDCA

Fig. 3. SSVEP decoding accuracy comparison between training-based spatial filtering methods with (w/) and without TLW-PLTS method under single-trial calibration data. For the violin, the top and bottom mean the maximum and minimum without outliers, and the width refers to the probability density. For the box in the violin, the top and bottom represent the 25th and 75th percentiles respectively. The centre line and the x-marker mean the median and the mean respectively.

TABLE I
THE PERFORMANCE COMPARISON WITH DIFFERENT DATA AUGMENTATION METHODS UNDER SINGLE-TRIAL CALIBRATION DATA

Methods	Maximum averaged ITR (bits/min)	Calibration time (s)	Training time (s)
INS-SF	205.79	120	5.53
INS-SF (w/ CST)	100.40	40	9.31
INS-SF (w/ SAME)	151.65	36	7.85
INS-SF (w/ PLTS)	173.42	69	9.31
INS-SF (w/ TLW-PLTS)	175.55	48.8	7.75

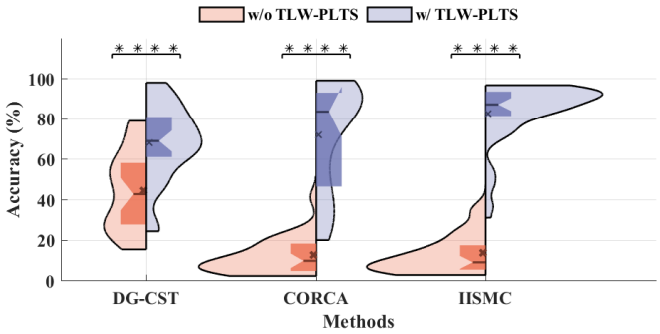


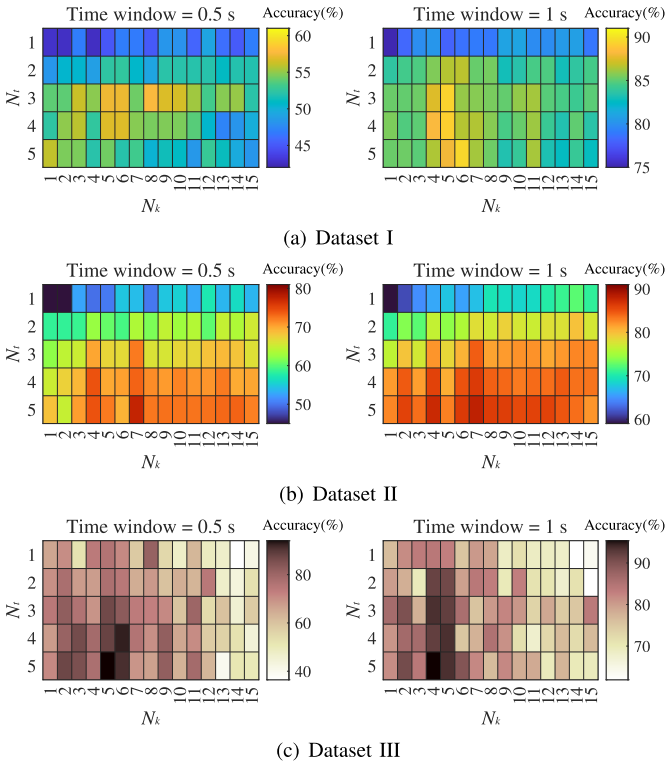
Fig. 4. Comparisons of SSVEP decoding accuracy between training-based transfer learning methods with (w/) and without (w/o) TLW-PLTS method using the least number of calibration trials via half violin plot.

of data augmentation and spatial filter training corresponding to all stimuli. The table shows that INS-SF (w/ TLW-PLTS) does not reach the maximum ITR as high as the original INS-SF, but the calibration time is significantly reduced from 120 s to 48.8 s. Moreover, with the proposed TLW-PLTS method, INS-SF obtained the highest ITR under single-trial calibration data, which is close to that of INS-SF (w/ PLTS), but much higher than that of INS-SF (w/ CST) and INS-SF (w/ SAME). The calibration time and training time of INS-SF (w/ TLW-PLTS) are both lower than those of INS-SF (w/ PLTS). In conclusion, the TLW-PLTS method can achieve higher SSVEP decoding performance in less time.

D. Offline Performance With Other Training-Based Methods

To further illustrate the contribution of the TLW-PLTS data augmentation method to the decoding performance, the classification accuracies of state-of-the-art training-based spatial filtering methods (i.e., TRCA, eTRCA, [14], and TDCA [18]) with and without the TLW-PLTS method are also compared. Figure 3 shows the classification accuracy distributions via a half-violin plot, using the single-trial training data from Dataset I. The number of augmented training trials is set to 5 for all methods; for TDCA, the number of subspaces and the number of delay points are set to 8 and 5 respectively. In the figure, the left violin represents the original spatial filtering method and the right violin represents the spatial filtering method after using the data augmentation method. From the figure, it can be seen that using only a single trial of calibration data for spatial filter training, the decoding accuracies of the TRCA, eTRCA, and TDCA methods under different time windows are all around 10%, indicating that the spatial filter and template signals computed by all three methods with a single training trial are not capable of decoding multi-target SSVEPs; whereas using the TLW-PLTS method to augment the training data, the three methods can obtain valid spatial filter and template signals, which can greatly improve the SSVEP decoding performance. These results demonstrate that the problem of low SSVEP decoding performance with a single trial can be effectively solved by using the TLW-PLTS method.

Furthermore, the decoding performance of state-of-the-art transfer learning methods with (w/) and without (w/o) the TLW-PLTS method are also compared using the least number of calibration trials from Dataset I. Here, three baseline transfer learning methods (i.e., cross-subject transfer method based on domain generalization (referred as DG-CST) [22], correlated component analysis (CORCA) [27], and inter- and intra-subject maximal correlation (IISMC) [28]) were conducted to verify the effectiveness of the proposed TLW-PLTS method. As shown in Fig. 4, the classification accuracy distributions are presented via violin plots. The time window is set as 1.0 s. The number of augmented training trials is set to 5 and the number of subbands of filter bank analysis is set to 5 for all three methods. For the settings of calibration trials, DG-CST requires at least 2-trial calibration data for transferred parameters calculation; CORCA and IISMC methods utilize single-trial calibration data. For the settings of the source domain, the numbers of source subjects in DG-CST and IISMC are both set as 10. For CORCA, the target subject is



(a) Dataset I

(b) Dataset II

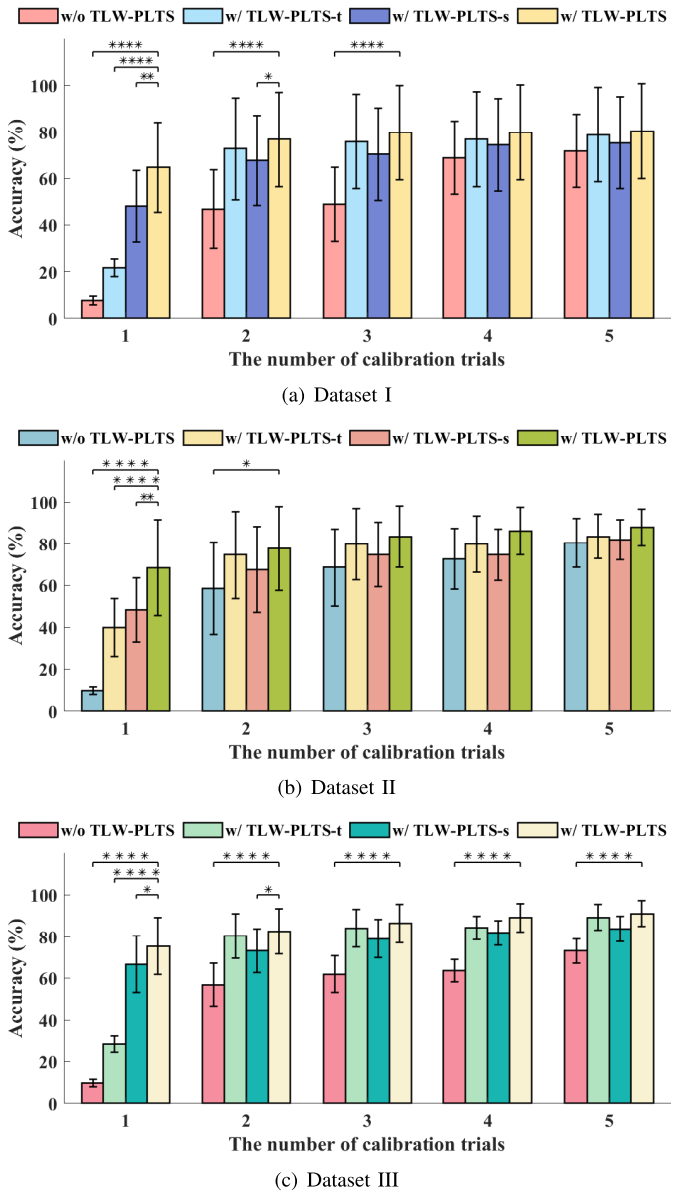
(c) Dataset III

Fig. 5. The averaged classification accuracies of INS-SF (w/ TLW-PLTS) using different numbers of augmented trials at 0.5-s and 1.0-s time windows. The colour bars represent the ranges of the classification accuracies.

sequentially selected one from all subjects, and the remaining subjects were treated as the source subjects. The figure also shows that the decoding performance of the training-based transfer learning method can be boosted significantly by using the proposed TLW-PLTS data augmentation method, due to the valid transferred spatial filter and templates.

E. The Effect of Parameters

1) The Number of Augmented Trials: The effect of the number of augmented trials on decoding performance is first investigated, which is the key parameter in this study. Figure 5 shows the averaged classification accuracies across subjects using different numbers of original trials (N_t) and augmented trials (N_k) on all datasets via heatmaps. In this case, the number of calibration trials N_t of the original training data for each dataset ranges from 1 to 5 with an interval of 1. The number of augmented trials for each dataset ranges from 1 to 15 with an interval of 1. Both 0.5-s and 1.0-s time windows are used to estimate the decoding performance. For Dataset I, the accuracy initially improves with an increase in augmented trials. But when $N_k > 6$, the accuracy starts to decline, especially in the 0.5-second window. Dataset II exhibits a similar trend, with accuracy increasing as N_k rises up to 7, followed by a slight decrease. For Dataset III, overall accuracies still show a decline in accuracy at levels of augmentation higher than 6. Therefore, the numbers of the augmented trials are set as 5, 7, and 5 respectively for the three datasets. Based on the findings in the PLTS method [36], the decline in accuracy results from decreasing similarity between augmented data and the original



(a) Dataset I

(b) Dataset II

(c) Dataset III

Fig. 6. Averaged classification accuracies across subjects under four different classification conditions, including without data augmentation (w/o TLW-PLTS), only with augmented template signals (w/ TLW-PLTS-t), only with augmented spatial filters (w/ TLW-PLTS-s), and with both augmented spatial filters and template signals (w/ TLW-PLTS).

data as N_k increases. The decoding model would become biased towards these less representative augmented epochs, leading to decreased performance on unseen test epochs and a subsequent drop in accuracy.

2) The Spatial Filter and Temporal Template: To investigate the enhancement of spatial filters and template signals on the SSVEP decoding performance, the classification accuracies were compared under four different classification conditions, including without data augmentation (w/o TLW-PLTS), only with augmented template signals (w/ TLW-PLTS-t), only with augmented spatial filters (w/ TLW-PLTS-s), and with both augmented spatial filters and template signals (w/ TLW-PLTS). From Fig. 6 it can be seen that w/o TLW-PLTS has the lowest accuracy and w/ TLW-PLTS has the highest accuracy for any

TABLE II
ONLINE SSVEP DECODING PERFORMANCE

Subject	No. of trials (Correct/Incorrect)	Accuracy (%)	ITR (bits/min)
S1	30(27/3)	90.00	151.09
S2	30(26/4)	86.67	139.49
S3	30(26/4)	86.67	139.49
S4	30(30/0)	100.00	195.54
S5	30(30/0)	100.00	195.54
S6	30(27/3)	90.00	151.09
S7	30(29/1)	96.67	177.76
S8	30(28/2)	93.33	163.67
S9	30(29/1)	96.67	177.76
S10	30(28/2)	93.33	163.67
S11	30(26/5)	86.67	139.49
S12	30(28/2)	93.33	163.67
Mean±STD	–	92.78±4.89	161.51±20.22

number of calibration trials. When $N_t = 1$, the accuracy of w/ TLW-PLTS-s computed for all datasets is higher than that of w/ TLW-PLTS-t, indicating that the augmented spatial filter contributes more to the decoding performance improvement than the template signal under a single calibration trial. When $N_t > 1$, the decoding accuracy of w/ TLW-PLTS-t obtained for all datasets is higher than that of w/ TLW-PLTS-s. Such results suggest that the augmented template signal makes a greater contribution to improving SSVEP decoding performance when multiple calibration trials are used. In conclusion, the experimental results show that the TLW-PLTS data enhancement method can not only effectively optimise the spatial filter, but also extract accurate template signals and improve the SSVEP decoding performance from both the spatial filter and template signals.

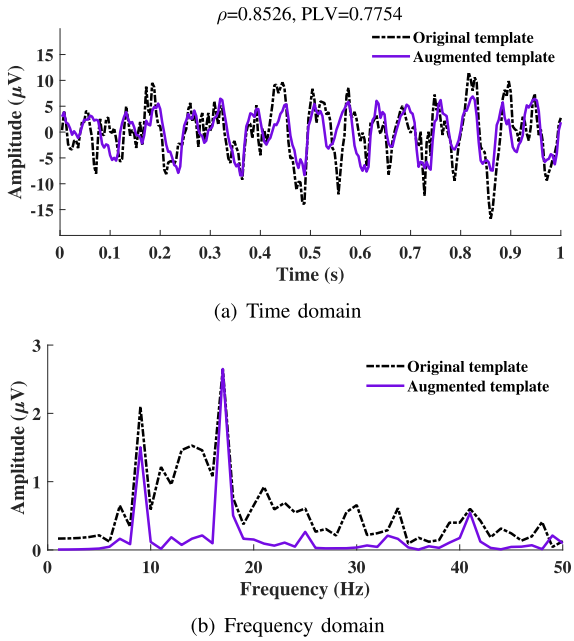
F. Online SSVEP Decoding Performance

The results of the online SSVEP decoding performance experiment are summarized in Table II and indicate a slightly higher performance compared to offline experiments. Individual accuracy varied from 86.67% to 100%, while the ITR ranged from 139.49 to 195.54 bits/min. On average, subjects achieved an accuracy of 92.78% with a standard deviation of 4.89%. The mean ITR was 161.51 bits/min, with a standard deviation of 20.22 bits/min. Notably, two subjects (S4 and S5) achieved a perfect score, with 100% accuracy across all 30 trials. The high performance observed across all subjects underscores the robustness and efficacy of the online SSVEP decoding method. These findings demonstrate not only the reliability of the offline methods but also the potential advantages over the online approach, highlighting its capability to consistently deliver high accuracy and efficient information transfer rates in real-life applications.

IV. DISCUSSIONS

A. Comparison With Transfer Learning Methods

In this paper, we propose a novel data augmentation method to facilitate SSVEP BCIs with less calibration effort. In SSVEP-based BCIs, both transfer learning and data augmentation methods have been introduced to facilitate the fast SSVEP decoding process with less calibration data. For the transfer learning methods, some methods, such as IISMC [28], use not only calibration data from the source subjects but also from the target subjects. In this way, the calibration time is not significantly reduced in terms of overall SSVEP-based BCIs. The other type of method only uses calibration data from source subjects, such as our previously published work DG-CST [22]. Such methods significantly reduced the calibration time for the overall SSVEP-based BCIs, but satisfactory decoding performance depends on the high similarity between the source and target data. Unlike transfer learning methods, data augmentation methods generate artificial SSVEP data with a small amount of calibration data. Although calibration data should be collected from each subject, single-trial calibration data would be sufficient to achieve satisfactory decoding performance using the proposed TLW-PLTS data augmentation method, as shown in Table I and Fig. 3. From the perspective of the overall BCI system, the calibration time is significantly reduced, while the high efficiency is also guaranteed, thus facilitating the fast calibration of the BCI system.



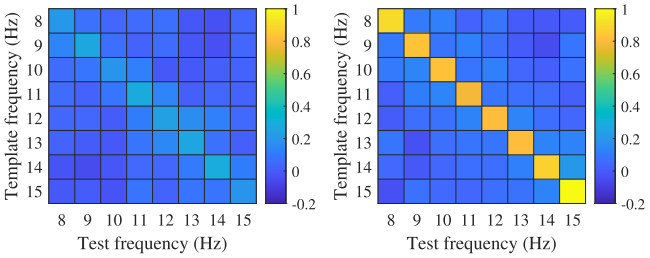
(a) Time domain

(b) Frequency domain

Fig. 7. Comparison between the original and augmented template signals in time (a) and frequency (b) domains.

B. The Feature Analysis of Augmented Template

In training-based spatial filtering methods, accurate SSVEP decoding mainly relies on a large number of SSVEP training signals. To reduce the training data acquisition and calibration effort, a TLW-PLTS data augmentation method is used to construct the augmented signals so that the augmented template can extract accurate SSVEP response features. To verify that the augmented template contains accurate SSVEP response features, using SSVEPs corresponding to 8.2 Hz, 0.5π from Dataset I as an example, Fig. 7 shows the features of the 1.0 s original SSVEP template and the augmented template from subject S3 on channel Oz in the time and frequency domains. As can be seen, the augmented template and the original template contain similar features in both the time and frequency



(a) w/o TLW-PLTS (b) w/ TLW-PLTS

Fig. 8. The correlation coefficients between the template signal and the test signals with (w/) and without (w/o) the TLW-PLTS method. Colour bars indicate the range of correlation coefficients.

domains. In terms of the time domain signals, the correlation between the two templates is 0.8526 and the phase-locked value (PLV) is 0.7754, showing high similarity between the original template and the augmented template. However, the time domain features of the augmented template are more stable compared to the original template. In the frequency domain, both templates contain frequency features of the fundamental and harmonic components, while the augmented template has stronger frequency domain response features due to the use of a temporal filter to suppress irrelevant noise in the data. These results show that the augmented training data constructed by the TLW-PLTS method contains accurate time-frequency SSVEP features, which can effectively optimise the spatial filter and template signals to improve SSVEP decoding performance.

C. The Correlation Feature Analysis

To visually present the contribution of the TLW-PLTS data augmentation method, 8-15 Hz in Dataset I is selected here as an example stimulus frequency to demonstrate the correlation coefficients between the template signal and the test signals with (w/) and without (w/o) the TLW-PLTS method, as shown in Fig. 8. In this case, the correlation coefficients were calculated using 1.0 s of data from a single trial in Dataset I as the original calibration data, and the final correlation eigenvalues were obtained by averaging the correlation coefficients from 35 subjects. The figure shows that when only single-trial calibration data are used as training data, the correlation coefficients of target and non-target stimulus frequencies are difficult to differentiate, resulting in a high detection error rate. When both single-trial calibration data and augmented data are used as training samples, the correlation coefficients corresponding to the target stimulus frequency are both close to 1 and the correlation coefficients corresponding to the non-target are close to 0 or even negative values. Therefore, such results can also show that using the TLW-PLTS method to augment the training data can effectively improve the decoding performance of SSVEP in single-trial training.

D. Limitations and Future Works

In this study, a new data augmentation method, TLW-PLTS, is developed for the fast-calibration SSVEP-based BCIs, which shows its efficiency through extensive comparison experiments. Further studies, which take the following issues into account, will need to be undertaken. The proposed TLW-PLTS data augmentation method extracts information from SSVEPs in both time and frequency domains. Therefore, in addition to the JFPM-coded target, the proposed method has the potential to be utilized to decode targets modulated by other methods [47], [48], [49]. Additionally, the proposed method achieved a higher ITR than SAME with more calibration time, as shown in Table I. In the future, the SAME method should be incorporated to obtain valid spatial filter and templates, which requires less calibration time. In addition, the TLW-PLTS method generates adequate trials with featured SSVEPs, which can be applied for transfer learning methods [21], [22], [23], [24], [25], [26] to further reduce calibration efforts.

V. Conclusion

This study proposes a new data augmentation method, TLW-PLTS, which uses limited calibration data to compute valid spatial filters and temporal templates for constructing fast-calibration SSVEP-based BCIs. The extensive comparison results with state-of-the-art spatial filtering methods showed the proposed TLW-PLTS method significantly enhanced the decoding performance with fewer calibration data. Therefore, the TLW-PLTS data augmentation method can effectively reduce calibration efforts, facilitating the practical applications of SSVEP-based BCIs.

Acknowledgment

The authors would like to thank the participants for data acquisition and online experiments in this study.

References

[1] J. R. Wolpaw et al., "Brain-computer interface technology: A review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 164–173, Jun. 2000.

[2] Y. Zhang, S. Q. Xie, H. Wang, and Z. Zhang, "Data analytics in steady-state visual evoked potential-based brain–computer interface: A review," *IEEE Sensors J.*, vol. 21, no. 2, pp. 1124–1138, Jan. 2021.

[3] F. Wang, Y. Wen, J. Bi, H. Li, and J. Sun, "A portable SSVEP-BCI system for rehabilitation exoskeleton in augmented reality environment," *Biomed. Signal Process. Control*, vol. 83, May 2023, Art. no. 104664.

[4] B. Xiong, J. Huang, B. Wan, C. Jiang, K. Su, and F. Wang, "A high-speed SSVEP-based speller using continuous spelling method," in *Communications in Computer and Information Science*. Cham, Switzerland: Springer, 2023, pp. 215–226.

[5] Y. Zhang, K. Qian, S. Q. Xie, C. Shi, J. Li, and Z.-Q. Zhang, "SSVEP-based brain–computer interface controlled robotic platform with velocity modulation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 3448–3458, 2023.

[6] Y. Zhang, Z. Li, S. Q. Xie, H. Wang, Z. Yu, and Z.-Q. Zhang, "Multi-objective optimization-based high-pass spatial filtering for SSVEP-based brain–computer interfaces," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–9, 2022.

[7] Y. Zhang, S. Q. Xie, H. Wang, C. Shi, and Z.-Q. Zhang, "Bayesian-based classification confidence estimation for enhancing SSVEP detection," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–12, 2023.

[8] 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, doi: 10.1109/TBME.2020. 2975552.

[9] 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.

[10] 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.

[11] K. Wang, D.-H. Zhai, Y. Xiong, L. Hu, and Y. Xia, "An MVMD-CCA recognition algorithm in SSVEP-based BCI and its application in robot control," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 5, pp. 2159–2167, May 2022.

[12] M. Norizadeh Cherloo, H. Kashefi Amiri, and M. R. Daliri, "Spatio-spectral CCA (SS-CCA): A novel approach for frequency recognition in SSVEP-based BCI," *J. Neurosci. Methods*, vol. 371, Apr. 2022, Art. no. 109499.

[13] Y. J. 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.*, Sep. 2014, pp. 3037–3040.

[14] 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.

[15] J. Huang, P. Yang, B. Xiong, B. Wan, K. Su, and Z. Zhang, "Latency aligning task-related component analysis using wave propagation for enhancing SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 851–859, 2022.

[16] J. Huang et al., "Incorporating neighboring stimuli data for enhanced SSVEP-based BCIs," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–9, 2022.

[17] 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.

[18] 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.

[19] 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. Cognit. Develop. Syst.*, vol. 14, no. 1, pp. 4–19, Mar. 2022.

[20] 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.

[21] Y. Zhang, S. Q. Xie, C. Shi, J. Li, and Z. Zhang, "Cross-subject transfer learning for boosting recognition performance in SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1574–1583, 2023.

[22] J. Huang et al., "Cross-subject transfer method based on domain generalization for facilitating calibration of SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 3307–3319, 2023.

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

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

[25] Z. Wang, C. M. Wong, A. Rosa, T. Qian, T.-P. Jung, and F. Wan, "Stimulus-stimulus transfer based on time-frequency-joint representation in SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 70, no. 2, pp. 603–615, Feb. 2023.

[26] B. Xiong, B. Wan, J. Huang, F. Li, X. Li, and P. Yang, "Cross-stimulus transfer method using common impulse response for fast calibration of SSVEP-based BCIs," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–14, 2024.

[27] Y. Zhang et al., "Correlated component analysis for enhancing the performance of SSVEP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 5, pp. 948–956, May 2018.

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

[29] C. Rommel, J. Paillard, T. Moreau, and A. Gramfort, "Data augmentation for learning predictive models on EEG: A systematic comparison," *J. Neural Eng.*, vol. 19, no. 6, Dec. 2022, Art. no. 066020.

[30] F. Fahimi, S. Dosen, K. K. Ang, N. Mrachacz-Kersting, and C. Guan, "Generative adversarial networks-based data augmentation for brain–computer interface," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 9, pp. 4039–4051, Sep. 2021.

[31] J. Kwon and C.-H. Im, "Novel signal-to-signal translation method based on StarGAN to generate artificial EEG for SSVEP-based brain–computer interfaces," *Expert Syst. Appl.*, vol. 203, Oct. 2022, Art. no. 117574.

[32] R. Li, L. Wang, P. N. Suganthan, and O. Sourina, "Sample-based data augmentation based on electroencephalogram intrinsic characteristics," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 10, pp. 4996–5003, Oct. 2022.

[33] R. Luo, M. Xu, X. Zhou, X. Xiao, T.-P. Jung, and D. Ming, "Data augmentation of SSVEPs using source aliasing matrix estimation for brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 70, no. 6, pp. 1775–1785, Jun. 2023, doi: 10.1109/TBME.2022.3227036.

[34] N. Waytowich et al., "Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials," *J. Neural Eng.*, vol. 15, no. 6, Dec. 2018, Art. no. 066031.

[35] A. Ravi, N. H. Beni, J. Manuel, and N. Jiang, "Comparing user-dependent and user-independent training of CNN for SSVEP BCI," *J. Neural Eng.*, vol. 17, no. 2, Apr. 2020, Art. no. 026028.

[36] X. Mai, J. Ai, Y. Wei, X. Zhu, and J. Meng, "Phase-locked time-shift data augmentation method for SSVEP brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 4096–4105, 2023.

[37] X. Xiao, L. Wang, M. Xu, K. Wang, T.-P. Jung, and D. Ming, "A data expansion technique based on training and testing sample to boost the detection of SSVEPs for brain-computer interfaces," *J. Neural Eng.*, vol. 20, no. 6, Dec. 2023, Art. no. 066017.

[38] H. Wang and W. Zheng, "Local temporal common spatial patterns for robust single-trial EEG classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 2, pp. 131–139, Apr. 2008.

[39] H. Wang, "Temporally local maximum signal fraction analysis for artifact removal from biomedical signals," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4919–4925, Sep. 2010.

[40] J. Jin, Z. Wang, R. Xu, C. Liu, X. Wang, and A. Cichocki, "Robust similarity measurement based on a novel time filter for SSVEPs detection," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 8, pp. 4096–4105, Aug. 2023.

[41] M. Norizadeh Cherloo, A. M. Mijani, L. Zhan, and M. R. Daliri, "A novel multiclass-based framework for P300 detection in BCI matrix speller: Temporal EEG patterns of non-target trials vary based on their position to previous target stimuli," *Eng. Appl. Artif. Intell.*, vol. 123, Aug. 2023, Art. no. 106381.

[42] 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.

[43] X. Chen, Y. Wang, M. Nakanishi, T.-P. Jung, and X. Gao, "Hybrid frequency and phase coding for a high-speed SSVEP-based BCI speller," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 3993–3996.

[44] 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.

[45] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *J. Amer. Stat. Assoc.*, vol. 74, no. 368, pp. 829–836, 1979.

[46] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in *Noise Reduction in Speech Processing*. Berlin, Germany: Springer, 2009, pp. 1–4, doi: 10.1007/978-3-642-00296-0_5.

[47] X. Chen et al., "Optimizing stimulus frequency ranges for building a high-rate high frequency SSVEP-BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1277–1286, 2023.

[48] G. Ming, W. Pei, X. Gao, and Y. Wang, "A high-performance SSVEP-based BCI using imperceptible flickers," *J. Neural Eng.*, vol. 20, no. 1, Feb. 2023, Art. no. 016042.

[49] R. Li et al., "A novel hybrid brain–computer interface combining the illusion-induced VEP and SSVEP," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 4760–4772, 2023.