

Phase-Locked Time-Shift Data Augmentation Method for SSVEP Brain-Computer Interfaces

Ximing Mai, Jikun Ai[✉], *Member, IEEE*, Yuxuan Wei[✉], Xiangyang Zhu[✉],
and Jianjun Meng[✉], *Senior Member, IEEE*

Abstract—Steady-state visual evoked potential (SSVEP) based brain-computer interfaces (BCIs) have achieved an information transfer rate (ITR) of over 300 bits/min, but abundant training data is required. The performance of SSVEP algorithms deteriorates greatly under limited data, and the existing time-shift data augmentation method fails to improve it because the phase-locked requirement between training samples is violated. To address this issue, this study proposes a novel augmentation method, namely phase-locked time-shift (PLTS), for SSVEP-BCI. The similarity between epochs at different time moments was evaluated, and a unique time-shift step was calculated for each class to augment additional data epochs in each trial. The results showed that the PLTS significantly improved the classification performance of SSVEP algorithms on the BETA SSVEP datasets. Moreover, under the condition of one calibration block, by slightly prolonging the calibration duration (from 48 s to 51.5 s), the ITR increased from 40.88 ± 4.54 bits/min to 122.61 ± 7.05 bits/min with the PLTS. This study provides a new perspective on augmenting data epochs for training-based SSVEP-BCI, promotes the classification accuracy and ITR under limited training data, and thus facilitates the real-life applications of SSVEP-based brain spellers.

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

I. INTRODUCTION

BRAIN-COMPUTER interfaces (BCIs) directly decode users' intentions from electrophysiological signals to control the external devices without the neuromuscular pathway [1]. Among the noninvasive modalities, the steady-state visual evoked potential (SSVEP)-based BCI has attracted widespread attention due to its advantages in high signal-to-noise rate (SNR) and little user training time [2]. In recent

years, SSVEP-BCI has been successfully developed in brain-actuated wheelchairs [3], [4], robotic arms [5], [6], spellers [7], [8], and rehabilitations [9], [10].

Target detection methods play an important role in enhancing the performance of SSVEP-BCI [11]. In the past few decades, the detection methods have evolved from detecting with single channel, e.g., power spectrum density analysis (PSDA) [12], into detecting with multiple channels, e.g., canonical correlation analysis (CCA) [13], minimum energy combination (MEC) [14], multivariate synchronization index (MSI) [15], improved versions of CCA [16], [17], and filter-bank canonical correlation analysis (FBCCA) [7]. Moreover, training-based detection methods, especially spatial filtering-based algorithms, have been developed recently. The spatial filtering methods have gained wide attention and have proved to be effective in the field of BCI [18], [19]. In SSVEP-BCI, the spatial filters are calculated using the individual training data to reduce the non-relevant brain activities and noises, and extract the relevant SSVEP activity to improve the signal-to-noise ratio (SNR) from the given multi-channel EEG signal. The extended CCA (eCCA) proposed by Chen et al. [20] calculated the spatial filters using CCA between temporal templates, artificial reference signal, and the test data. The ensemble task-related component analysis (eTRCA) proposed by Nakanishi et al. [8] derived the spatial filters by solving the inter-trial covariance maximization problem, and it greatly promoted the classification performance. The extended versions of eTRCA [21], [22] were proposed thereafter to further improve the performance of eTRCA. The task-discriminant component analysis (TDCA) proposed by Liu et al. [23] sought to find a uniform spatial filter of all classes with a discriminative model, and it significantly outperformed eTRCA in public datasets and self-collected dataset. Despite the fact that spatial filtering based detection methods have achieved high performance in SSVEP-BCI, they require a large amount of individual training data. For example, the eCCA and eTRCA require a calibration session of 12 blocks with each block containing 40 trials before applying the SSVEP-BCI system for online detection. The collection of sufficient training data is time-consuming and might cause a visual fatigue, resulting in user's low acceptance of the BCI system.

Manuscript received 24 May 2023; revised 16 September 2023; accepted 18 September 2023. Date of publication 10 October 2023; date of current version 24 October 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 52175023 and Grant 91948302. (Corresponding authors: Jianjun Meng; Xiangyang Zhu.)

The authors are with the State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: mexyzhu@sjtu.edu.cn; mengjianjunxs008@sjtu.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2023.3323351

Therefore, researchers have explored the feasibility of detecting targets under limited SSVEP training data. For example, transfer learning-based methods [24], [25] can use the training data from existing subjects in the same experiment or training data collected from other experiments using different acquisition devices to train the detection model for the new subject directly. Generative adversarial networks (GAN) [26] can generate multiple synthetic data given the existing data from other subjects or experiments and little data from the new subject. By doing so, the new subject can be free from the long training time. However, these methods still need to collect a large amount of data in advance, and fine-tuning is required for numerous parameters in the model. Some researchers utilize the existing but unused data for detection model training, including using data from other classes which had the neighbor frequencies [27] or locations on the screen [28] to the target class to train the spatial filters or using the unlabeled test data for obtaining new spatial filters for detection [29]. Another attempt is to perform data transformation to the data epochs to obtain multiple new augmented epochs. Luo et al. [30] proposed source aliasing matrix estimation (SAME) to augment new data by combining the derived SSVEP source component from the training data and the estimated Gaussian noise. Li et al. [31] augmented new data by performing sample-based transformations, including performance-measure-based (PMB) time wrap, frequency noise addition, and frequency masking onto the data epochs.

Time-shift augmentation on the experimental data is another frequently used method in BCI field to extract more training samples for the model training [32], [33]. In the offline calibration data of SSVEP-BCI, each trial duration is set to 3 s or 5 s, and data epochs of $[d, d + T_w]$ s in each trial is extracted for training the spatial filtering algorithms. d indicates the SSVEP latency of 0.13 s or 0.14 s [34], [35], and T_w indicates the selected window length for online decoding, which is set to around 0.6 s and is much shorter than the trial duration. However, low classification performance was obtained in previous SSVEP studies [36], [37] when time-shift augmentation method was directly used to augment more data epochs for training the spatial filtering algorithm. The possible reason is that it violates the time- and phase-locked requirement between different data epochs [30], [36], [37]. Since subjects steadily gaze at the cued stimulation target for the entire trial, multiple epochs with similar SSVEP characteristics could be extracted within one trial. Thus, the feasibility of applying the time-shift augmentation method to spatial filtering-based algorithms needs to be explored further.

This study proposes a novel time-shift data augmentation method to augment more data epochs for training-based SSVEP algorithms. In detail, the similarity between the original data epoch $[d, d + T_w]$ s and data epochs at different time moment $[a_t, a_t + T_w]$ s within one trial was evaluated. Based on the similarity results, the phase-locked time-shift (PLTS) data augmentation method was proposed, in which a unique time-shift step for each SSVEP class was set to augment additional epochs within one trial. Then, the proposed PLTS augmentation method was verified on the state-of-the-art

(SOTA) spatial filtering algorithms, i.e., eTRCA and TDCA, for the Benchmark and the BETA dataset. Finally, PLTS was implemented in a situation of short calibration time, the single calibration block condition, and it showed significant improvement, which promoted the real-life applications of SSVEP-BCI.

The paper is organized as follows. The training-based SSVEP-BCI, the data augmentation process, the used public datasets, and evaluation metrics are described in Section II. The results are presented in Section III, with discussion and conclusion followed in Sections IV and V, respectively.

II. METHODS

A. Training-Based Method in SSVEP-BCI

As shown in Fig. 1, in SSVEP-BCIs, N_f visual stimuli are presented to subjects, and the experiment consists of N_b blocks, each including N_f trials corresponding to N_f stimuli. Each trial starts with a visual cue indicating the target stimulus, which lasts for 0.5s. Subjects are instructed to shift their gaze to the target within the cue duration and gaze at it. After the cue, all stimuli start to flicker for T_n s. Then, the screen is blank for 0.5s before the next trial begins. Epoch of $[d, d + T_w]$ s in each trial is extracted, in which the time moment 0 indicates the stimulus onset, d indicates the SSVEP response latency [34], and T_w indicates the time window length for training and decoding.

The SSVEP detection model can be trained using the individual calibration data $\mathbf{X}_n^{(j)} \in R^{N_{ch} \times N_s}$ ($n \in [1, N_f]$, $j \in [1, N_b]$, $n, j \in N^*$). n indicates the stimulus class index, and N_f is the number of stimulus classes in total; j indicates the index of the training sample, and N_b is the number of blocks / training samples for each class; N_{ch} indicates the number of channels for each epoch, and N_s indicates the data length extracted in each trial (which is equal to $T_w \times f_s$, f_s indicates the sampling rate). First, the spatial filter $\mathbf{W}_n \in R^{N_{ch} \times 1}$ is calculated (the detailed calculation process of spatial filter in eTRCA and TDCA could be found in [8] and [23]):

$$\mathbf{W}_n = f(\mathbf{X}_n^{(1)}, \mathbf{X}_n^{(2)}, \dots, \mathbf{X}_n^{(N_b)}), \quad n = 1, 2, \dots, N_f \quad (1)$$

Then, the temporal template for each class is obtained by averaging across different training samples:

$$\bar{\mathbf{X}}_n = \frac{1}{N_b} \sum_{j=1}^{N_b} \mathbf{X}_n^{(j)}, \quad n = 1, 2, \dots, N_f \quad (2)$$

Finally, given a test data $\mathbf{X}_t \in R^{N_{ch} \times N_s}$, the correlation coefficient between the projected temporal template and the projected test data is calculated:

$$\rho_n = \text{corr}(\bar{\mathbf{X}}_n^T \mathbf{W}_n, \mathbf{X}_t^T \mathbf{W}_n) \quad (3)$$

and the target class can be identified by choosing the following equation:

$$\tau = \underset{n}{\text{argmax}} \rho_n, \quad n = 1, 2, \dots, N_f \quad (4)$$

The filter bank technique is usually implemented along with the training-based SSVEP detection methods, in which

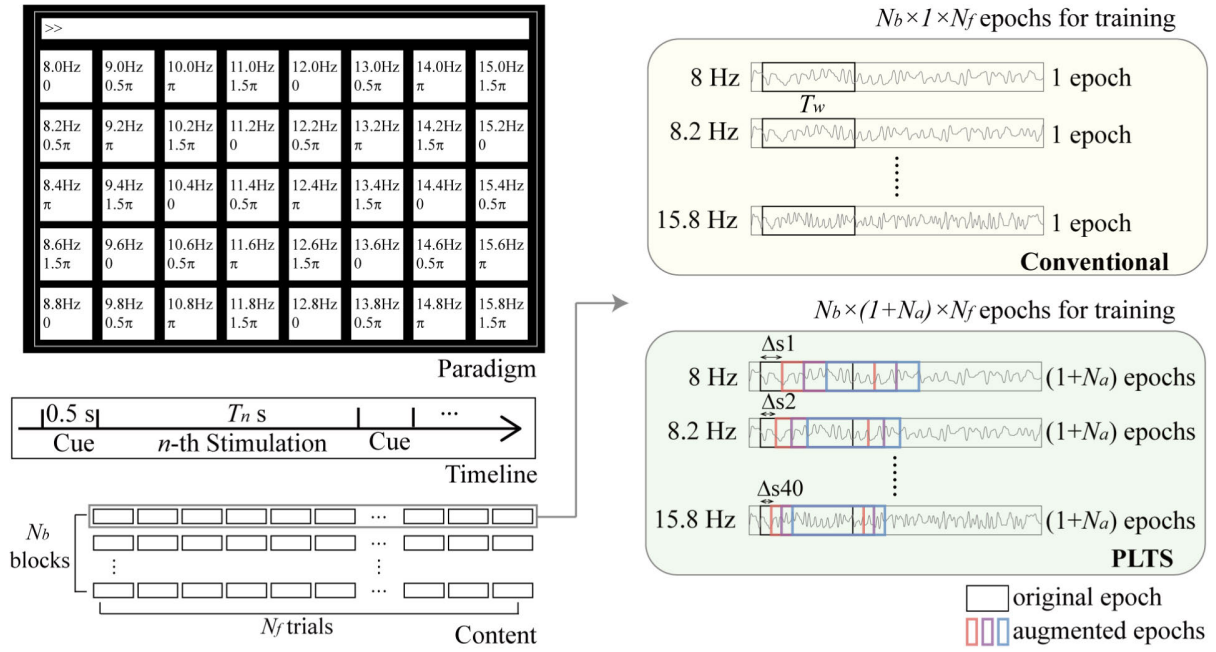


Fig. 1. Methodology overview. In SSVEP-BCI, the calibration session includes N_b blocks, each containing N_f trials corresponding to all N_f visual stimuli with different frequencies and phases. In each trial, subjects are required to gaze at the cued stimulus for T_n s. Conventionally, data epoch of $[d, d + T_w]$ s in each trial is extracted for model training, where d indicates the SSVEP response latency and T_w indicates the data length for training and decoding. In the proposed PLTS augmentation method, a unique time-shift step is assigned for each SSVEP class and therefore $(1 + N_a)$ epochs can be extracted in each trial, providing more data for training the model.

data are decomposed into N_m sub-band components to utilize the harmonic information. In detail, the m -th zero-phase Chebyshev Type I infinite impulse response (IIR) filters with the lower and upper cut-off frequencies of $[8m, 90]$ Hz is applied to both the training data and the test data. After calculating the above-mentioned spatial filters, temporal templates, and the correlation coefficients in each sub-band component, a weighted sum of squares of the combined correlation coefficients is calculated and the class with the maximum weighted correlation is selected as the target class:

$$\rho_i = \sum_{m=1}^{N_m} w_m \left(\rho_i^{(m)} \right)^2 \quad (5)$$

w_m is defined according to [7]:

$$w_m = m^{-1.25} + 0.25, \quad m = 1, 2, \dots, N_m. \quad (6)$$

B. Data Augmentation

1) **Finding Epochs With High Similarity:** To fully utilize the experiment data, apart from the original epoch $\mathbf{x}_o[d, d + T_w]$ s, we hypothesize that data in $[d + T_w, T_n]$ s in each trial can also be extracted for SSVEP detection model training since subjects steadily gaze at the flickering stimulus, and the SSVEP is evoked for the entire trial. In detail, the augmented epoch \mathbf{x}_a extracted from $[a_t, a_t + T_w]$ s should have similar SSVEP characteristics with \mathbf{x}_o , and can be served as a new training sample for the detection model. a_t indicates the start time moment for extracting the augmented epoch, and $(a_t + T_w)$ indicates the end time moment for extracting it. The optimal

augmented epoch \mathbf{x}_a can be determined by finding the time moment a with the maximum similarity (where $a = a_t \times f_s$):

$$a = \operatorname{argmax} \operatorname{Similarity}(\mathbf{x}_o, \mathbf{x}_a) \quad (7)$$

The similarity is measured by the correlation between \mathbf{x}_o and \mathbf{x}_a :

$$\operatorname{Similarity}(\mathbf{x}_o, \mathbf{x}_a) = \frac{\sum_{i=1}^{N_s} (x_o[i] - \bar{x}_o)(x_a[i] - \bar{x}_a)}{N-1} \sqrt{\frac{\sum_{i=1}^{N_s} (x_o[i] - \bar{x}_o)^2}{N_s-1} \frac{\sum_{i=1}^{N_s} (x_a[i] - \bar{x}_a)^2}{N_s-1}} \quad (8)$$

The grid search method can be used, in which the time moment a is set to $d \times f_s + 1, d \times f_s + 2, \dots, (T_n - T_w) \times f_s$ to calculate the similarity between \mathbf{x}_o and all possible \mathbf{x}_a , and the time moment with the maximum similarity is selected to be the optimal a .

2) **PLTS Augmentation Method:** In SSVEP-BCI, the augmented epoch \mathbf{x}_a shows high similarity with the original epoch \mathbf{x}_o periodically with the change of time moment a , and this period is equal to the reciprocal of the stimulation frequency of the trial (see Section III-A for detail). Therefore, the phase-locked time-shift (PLTS) SSVEP augmentation method is proposed in this study: given a trial data of n -th class (the stimulation frequency is f_n Hz), in addition to the original epoch, time-shift augmentation can be performed with a unique step Δs_n :

$$\Delta s_n = \frac{f_s}{f_n} \quad (9)$$

and augmented epochs $[a_t, a_t + T_w]$ s can be severed as additional training samples. The time moment a is:

$$a = a_t \times f_s = p \times \Delta s_n, \quad p = 1, 2, \dots, N_a \quad (10)$$

In practice, total N_a augmented epochs are extracted within one trial (the value of N_a is determined in Section III-B). Note that the PLTS SSVEP augmentation method is different from the conventional time-shift augmentation method because the step Δs of time shift in PLTS is different for each class, whereas Δs in the conventional method is set to a fixed value, e.g., time samples of 50ms or 100ms, for all class.

By implementing the PLTS augmentation method, the number of training samples for each class in N_b blocks can be enlarged from N_b to $N_b \times (1 + N_a)$. The pseudo code for PLTS augmentation and model training is described as follows, and the process of PLTS is summarized in Fig. 1.

Algorithm 1 PLTS Data Augmentation

Input: the original four-dimension data matrix $X_o \in \mathbb{R}^{N \times N_{ch} \times N_b \times N_f}$ (N is the total data sample number; N_{ch} is the channel number), the n_a -th augmented epoch extracted from n_b -th block $X \in \mathbb{R}^{N_s \times N_{ch}}$, the stimulation frequency for n -th class f_n , the time-shift step for n -th class Δs_n , the sampling rate f_s , the data sample number for each extracted epoch N_s , the number of training block N_b , the number of augmented epochs N_a

Output: the augmented data matrix $X_a \in \mathbb{R}^{N_s \times N_{ch} \times (N_b N_a) \times N_f}$

```

1  for  $n \leftarrow 1$  to  $N_f$  do
2    Compute the unique step  $\Delta s_n = \frac{f_s}{f_n}$ 
3    for  $n_b \leftarrow 1$  to  $N_b$  do
4      for  $n_a \leftarrow 1$  to  $N_a$  do
5        Extract augmented epoch  $X$  from  $X_o$ 
6        Append  $X$  into  $X_a$ 
7  Train the detection model using  $X_a$ 

```

C. Evaluation

1) *Data Used in This Study:* The similarities between the original epoch and the augmented epochs were evaluated by analyzing the SSVEP public dataset [38]. Ten healthy subjects participated in 15 blocks of a cued-spelling task on a 4×3 matrix of a virtual keypad, each block contained 12 trials. The frequencies of the flickering stimuli ranged from 9.25Hz to 14.75Hz with an interval of 0.5Hz, and the phase was set to 0 initially and incremented with 0.5π for different stimuli. The stimulation duration of each trial was 4 s, and the sampling rate was 256 Hz. The epoch data of channel Oz was used to calculate the similarity.

The value of N_a in the proposed PLTS data augmentation method was determined using the Benchmark dataset [34], and the performance of PLTS was verified on the BETA dataset [35]. For the Benchmark dataset, 35 subjects participated in six blocks of a cued-spelling task on a 5×8 matrix of a virtual keyboard. For the BETA dataset, 70 subjects participated in four blocks of a cued-spelling task on a QWERTY virtual keyboard. The stimulus frequencies for Benchmark and

BETA dataset ranged from 8.0Hz to 15.8Hz with an interval of 0.2Hz, and the phase was set to 0 initially and incremented with 0.5π for different stimuli. The stimulation duration was 5s in the Benchmark dataset, and 2s or 3s in the BETA dataset. The sampling rate was 250Hz. Signals from nine classical EEG channels (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, and O2) were chosen for detection. The SSVEP response latency d was set to 140ms in the Benchmark and 130ms in the BETA. Time window length T_w was set to 0.3s, 0.4s, ..., 1s.

2) *Performance Comparison:* In this study, the SSVEP detection methods (eTRCA [8] and TDCA [23]) with and without the proposed PLTS data augmentation method were compared to evaluate the performance of PLTS. Moreover, the SAME [30] data augmentation method used in SSVEP BCI was also included for comparison.

3) *Evaluation Metrics:* The correlation coefficient was used to evaluate the similarity between the original epoch and augmented epochs. The classification accuracy and information transfer rate (ITR) were evaluated using k -fold cross-validation ($k = 6$ for the Benchmark dataset, and $k = 4$ for the BETA dataset). The ITR in units of bits per min (bpm) is calculated by [39]:

$$ITR = \left(\log_2 M + P \log_2 P + (1 - P) \log_2 \frac{1 - P}{M - 1} \right) \times \frac{60}{T}$$

where M indicates the number of classes, P indicates the classification accuracy, and T in a unit of second indicates the target selection time, which is the sum of gaze time T_w and the 0.5s gaze shifting time.

4) *Statistical Analysis:* When applicable, results were expressed as mean \pm SEM (standard error of the mean) unless otherwise stated. The error bars shown in the figures indicated the SEM. One-way and two-way repeated measures analysis of variance (ANOVA) was applied to test the difference in classification accuracy and ITR of eTRCA and TDCA in different numbers of training blocks N_b and different augmentation conditions. The Greenhouse-Geisser correction was used if the data didn't conform to the sphericity assumption by Mauchly's test of sphericity. All post hoc pairwise comparisons were Bonferroni corrected. The alpha level of statistical significance was set at 0.05.

III. RESULTS

A. Similarity Between Original AND Augmented Epochs

To find the optimal augmented epoch, the similarity between original epoch x_0 and augmented epoch x_a at the time moment a of $[d \times f_s + 1, d \times f_s + 256]$ time samples was computed. Fig. 2(a) shows examples of similarity changes with the time moment a at 9.75Hz, 10.75Hz, and 14.75Hz respectively. The blue dotted curve indicates the calculated similarity for each subject, and the black curve indicates the average similarity across all subjects. It can be seen that the maximum similarity occurs periodically at all three frequency conditions. The time moments of the first three maximum similarity was 26, 52, 79 for 9.75Hz, 24, 48, 72 for 10.75Hz, and 18, 34, 52 for 14.75Hz, respectively. The period, i.e., the interval between the time moments, decreases with the increase of the stimulation frequency, and moreover, it is equal to the period of the

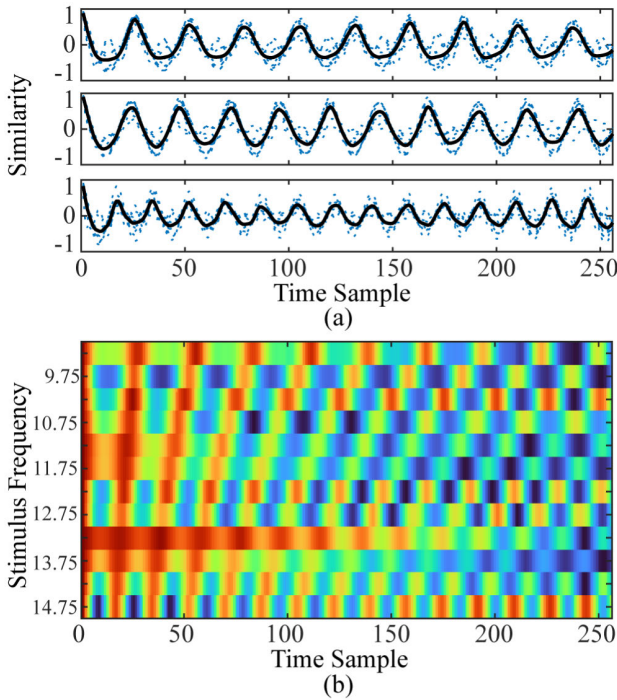


Fig. 2. Similarity between the original epoch and the augmented epochs. The time moment of the augmented epochs ranges from $d \times f_s + 1$ to $d \times f_s + f_s$. (a) Example similarity calculated using class data of 9.75 Hz, 10.75 Hz, and 14.75 Hz, respectively. The blue dotted curve indicates the calculated similarity for each subject, and the black curve indicates the average similarity across all subjects. (b) Similarity of all classes from 9.25 Hz to 14.75 Hz. Color indicates the similarity at each time moment of each class.

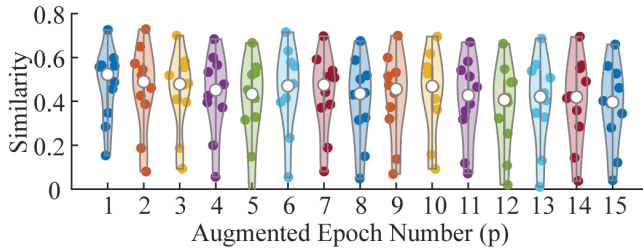


Fig. 3. The change of periodically occurred maximum similarity. The x-axis indicates the ordinal number p of the augmented epochs, the y-axis indicates the similarity between the original epoch and the p th augmented epoch.

stimulation frequency $\frac{f_s}{f_n}$. Furthermore, the similarity for all stimulation frequencies is illustrated in Fig. 2(b), in which the phenomenon of periodic maximum similarity exists at all frequency conditions, and the period is also decreased with the increase of the stimulation frequency. Therefore, given a trial data of stimulation frequency f_n , optimal augmented epochs can be obtained by extracting epochs at time moment $a = p \times \frac{f_s}{f_n}$, $p = 1, 2, \dots$, which show maximum similarity to the original epoch. p indicates the ordinal number of augmented epochs.

To evaluate the changes of periodically occurred maximum similarity, the first 15 similarity between the original epoch and p -th augmented epoch for each stimulation frequency and each subject was calculated and averaged, as shown in Fig. 3. The similarity shows a trend of decrease with the increase of

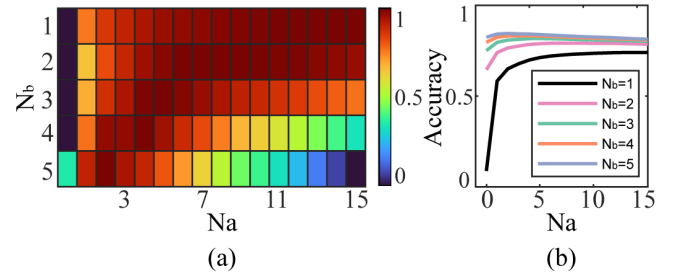


Fig. 4. The classification accuracy of eTRCA under different training block number N_b and different augmentation number N_a for the Benchmark dataset presented in (a) heatmap figure and (b) curve figure. The accuracy value in the heatmap was normalized to $[0, 1]$ for each N_b to better illustrate the accuracy change with different N_a .

ordinal number p , but it remains above 0.4 for all augmented epochs.

B. PLTS Parameter Optimization

To explore the classification accuracy improvement when using PLTS augmentation method with different augment number N_a , a grid search method was conducted to evaluate eTRCA with PLTS for the Benchmark dataset. As shown in Fig. 4, with the increase of augment number N_a , the classification accuracy first increased, then remained stable or slightly decreased. Two-way repeated measures ANOVA revealed that two within-subject factors N_b and N_a and their interaction were statistically significant for the Benchmark dataset ($N_b : F(1.10, 37.40) = 333.68, p < 0.001$, $N_a : F(1.40, 47.43) = 149.32, p < 0.001$, $N_b \times N_a : F(1.66, 56.34) = 204.61, p < 0.001$). The N_a in which the accuracy reached maximum or became stable was selected as the optimal parameter to reduce the computation cost in training. Therefore, N_a was set to 6, 5, 5, 3, and 3 for 1, 2, 3, 4, and 5 training blocks, respectively.

C. Classification Performance With PLTS

To testify the efficacy of PLTS augmentation method, the eTRCA and TDCA performance without any data augmentation method (Original), with PLTS (w/PLTS), and with SAME (w/SAME) were compared, as shown in Table I. The results showed that PLTS promoted the classification accuracy of eTRCA by 49.60%, 9.80%, and 5.00% for 1, 2, and 3 training blocks respectively, and it promoted the classification accuracy of TDCA by 33.67%, 2.86%, and 0.57%. When compared to the SAME, the PLTS promoted the classification accuracy of eTRCA by 12.78%, -0.43% , and -1.04% for 1, 2, and 3 training blocks respectively, and it promoted the classification accuracy of TDCA by 4.95%, 0.19%, and -0.92% . Two-way repeated measures ANOVA showed that augment methods, training blocks, and their interaction were statistically significant for both eTRCA and TDCA ($p < 0.001$). One-way repeated measures ANOVA and post hoc test revealed that the classification accuracy of PLTS was significantly higher than that without any data augmentation for both eTRCA and TDCA for all training blocks conditions ($p < 0.001$); compared with the SAME, classification accuracy of PLTS for both eTRCA and TDCA was significantly higher for one

TABLE I
CLASSIFICATION PERFORMANCE COMPARISON AMONG ALGORITHMS WITH DIFFERENT AUGMENTATION METHODS INCLUDING ORIGINAL, SAME, AND PLTS

Augmentation methods	Classification accuracy with different augmentation methods and different number of training blocks					
	eTRCA			TDCA		
	$N_b = 1$	$N_b = 2$	$N_b = 3$	$N_b = 1$	$N_b = 2$	$N_b = 3$
Original	$9.15 \pm 0.92\%$	$56.13 \pm 3.04\%$	$65.22 \pm 2.91\%$	$27.33 \pm 2.18\%$	$64.96 \pm 2.83\%$	$71.35 \pm 2.63\%$
w/SAME	$45.97 \pm 2.90\%$	$66.36 \pm 2.77\%$	$71.26 \pm 2.63\%$	$56.04 \pm 2.79\%$	$67.63 \pm 2.68\%$	$72.83 \pm 2.55\%$
w/PLTS	$58.75 \pm 2.66\%$	$65.93 \pm 2.63\%$	$70.22 \pm 2.51\%$	$60.99 \pm 2.59\%$	$67.82 \pm 2.53\%$	$71.92 \pm 2.44\%$
P1 vs. P2	Pairwise comparison between the proposed PLTS and original / w/SAME					
w/PLTS vs. Original	>>>	>>>	>>>	>>>	>>>	>>>
w/PLTS vs. SAME	>>>	-	<<	>>>	-	<<

Note: > (or <) indicates P1 (w/PLTS) is significantly higher (or lower) than P2 (original / w/SAME) ($p < 0.05$). >> (or <<) indicates P1 is significantly higher (or lower) than P2 ($p < 0.01$). >>> (or <<<) indicates P1 is significantly higher (or lower) than P2 ($p < 0.001$). - indicates P1 is not significantly higher or lower than P2.

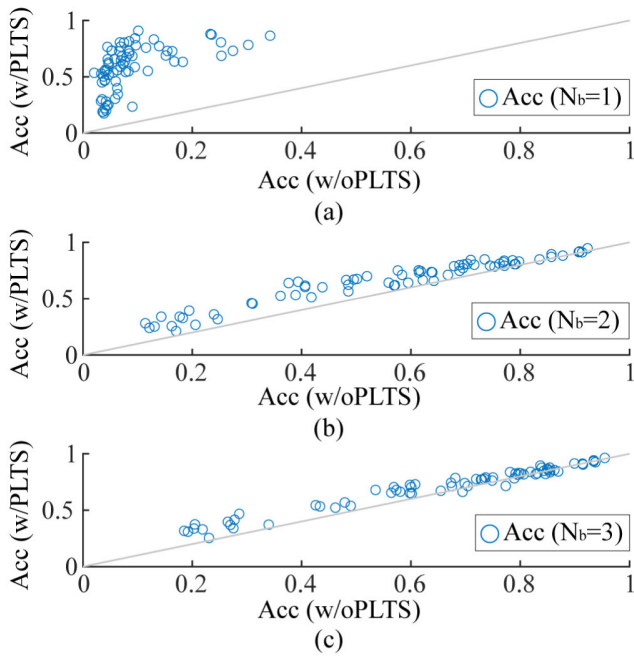


Fig. 5. The classification accuracy distribution of all subjects for the BETA dataset when (a) $N_b = 1$, (b) $N_b = 1$, and (c) $N_b = 1$, respectively. Each blue circle indicated the accuracy for each subject. The x value indicates the accuracy obtained by the subject without the PLTS, the y value indicates the accuracy obtained with the PLTS.

training block ($p < 0.001$), comparable for two training blocks ($p = 0.456$, $p = 1.000$) and significantly lower for three training blocks ($p = 0.005$, $p = 0.009$) with a difference of around 1%.

Fig. 5 demonstrates the detail of eTRCA performance in the BETA dataset with and without PLTS. Each blue circle indicated the accuracy for each subject. It can be seen that when one block of data was available for training, all blue circles scattered at the upper left area, indicating that the performance of all subjects was greatly promoted by the PLTS. With the increase of the number of training blocks, circles were closer to the central line ($y = x$), indicating that the PLTS had lower accuracy improvement.

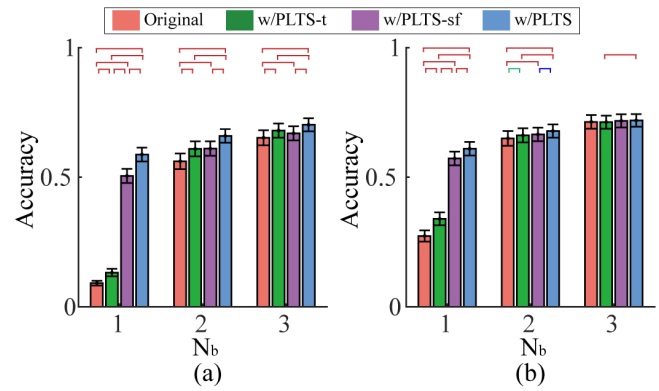


Fig. 6. The classification accuracy under different augmentation conditions, for (a) eTRCA and (b) TDCA, respectively. Conditions include the algorithm without PLTS (Original), with updated temporal templates (w/PLTS-t), with updated spatial filters (w/PLTS-sf), and with both the updated temporal templates and updated spatial filters (w/PLTS). The green brackets indicate $p < 0.05$, the blue brackets indicate $p < 0.01$, and the red brackets indicate $p < 0.001$.

D. Spatial Filters and Temporal Templates Improvement by PLTS

To further explore the contribution of PLTS to SSVEP detection methods, four augmentation conditions were considered, including classification without PLTS augmentation (w/oPLTS), classification with updated spatial filters after utilizing PLTS (w/PLTS-sf), classification with updated temporal templates after utilizing PLTS (w/PLTS-t), classification with both updated spatial filters \bar{W}_n and temporal templates $\bar{\chi}_n$ after utilizing PLTS (w/PLTS). As shown in Fig. 6, performance without PLTS had the lowest performance, and that with PLTS achieved the highest performance for both eTRCA and TDCA, and all numbers of training blocks N_b . Two-way repeated measures ANOVA revealed that two within-subject factors N_b and augmentation conditions and their interaction were statistically significant in eTRCA and TDCA ($p < 0.001$). One-way repeated measures ANOVA and the post hoc test showed that the performance with PLTS was significantly higher than all other three conditions in both eTRCA and TDCA for all training block conditions, except for $N_b = 3$ in TDCA. The performance of w/PLTS-sf was significantly

higher than that of w/PLTS-t in both eTRCA and TDCA for $N_b = 1$.

E. Performance Under Single Calibration Block

It is necessary for training-based SSVEP-BCI to diminish the calibration time before online application since a long calibration time causes visual fatigue and users' low acceptance of the BCI system. In this study, four simulated conditions of calibration session before the online decoding were compared, including (1) only one calibration block with a trial duration of T_o ($N_b = 1$); (2) only one calibration block with the expanded trial duration of $(T_o + 1 \times \frac{1}{f_n})$, in which PLTS with one augmented epoch was implemented ($N_b = 1, N_a = 1$); (3) only one calibration block with the expanded trial duration of $(T_o + N_a \times \frac{1}{f_n})$, in which PLTS with a total of N_a augmented epochs was implemented ($N_b = 1, N_a = 6$); (4) two calibration blocks ($N_b = 2$). Supposing that the gaze shift time for each trial was 0.5 s, and the rest time between training blocks was 60 s, then the calibration time D (unit: second) for the four conditions were: (1) $N_f \times (T_o + 0.5)$; (2) $N_f \times (T_o + 0.5) + \sum_{n=1}^{N_f} \frac{1}{f_n}$; (3) $N_f \times (T_o + 0.5) + 6 \times \sum_{n=1}^{N_f} \frac{1}{f_n}$; (4) $2 \times N_f \times (T_o + 0.5) + 60$, respectively.

Fig. 7 illustrates classification accuracy and ITR under different time window length T_o for online decoding and different calibration conditions. For $T_o = 0.7$ s, the condition ($N_b = 1$) required a short calibration time $D_1 = 48$ s, but it had extremely low ITR for eTRCA and TDCA, which was impractical for online applications. The condition ($N_b = 1, N_a = 1$) required a slightly longer calibration time $D_2 = 48 + 3.50 = 51.5$ s, but achieved much higher ITR compared to the condition ($N_b = 1$), i.e., 95.56 ± 7.34 vs 6.47 ± 1.30 bits/min for eTRCA, and 122.61 ± 7.05 vs 40.88 ± 4.54 bits/min for TDCA. The condition ($N_b = 1, N_a = 6$) required a calibration time $D_3 = 48 + 21.0 = 69$ s, and had comparable ITR, and lower ITR compared to the condition ($N_b = 2$) for eTRCA and TDCA, respectively (note that the calibration time $D_4 = 156$ s for the condition $N_b = 2$). One-way repeated measures ANOVA was used to evaluate the difference among the maximum ITR achieved by different calibration conditions, and the results revealed that the calibration condition was statistically significant in eTRCA ($F(1.28, 88.40) = 375.65, p < 0.001$) and TDCA ($F(1.23, 84.93) = 589.22, p < 0.001$). These results indicated that the problem of low SSVEP performance with one training block can be solved by slightly prolonging the calibration time by 3.5 s and implementing the PLTS augmentation method.

IV. DISCUSSION

This study proposes the PLTS augmentation method for SSVEP-BCI under limited training data. The state-of-the-art algorithms eTRCA and TDCA showed improved performance when combined with the PLTS. Especially when the training block was one, the classification accuracy improvement of eTRCA was 49.60%, and TDCA's was 33.67%. Moreover, the PLTS significantly outperformed the SAME by 12.78% and 4.95% for eTRCA and TDCA, respectively. The results

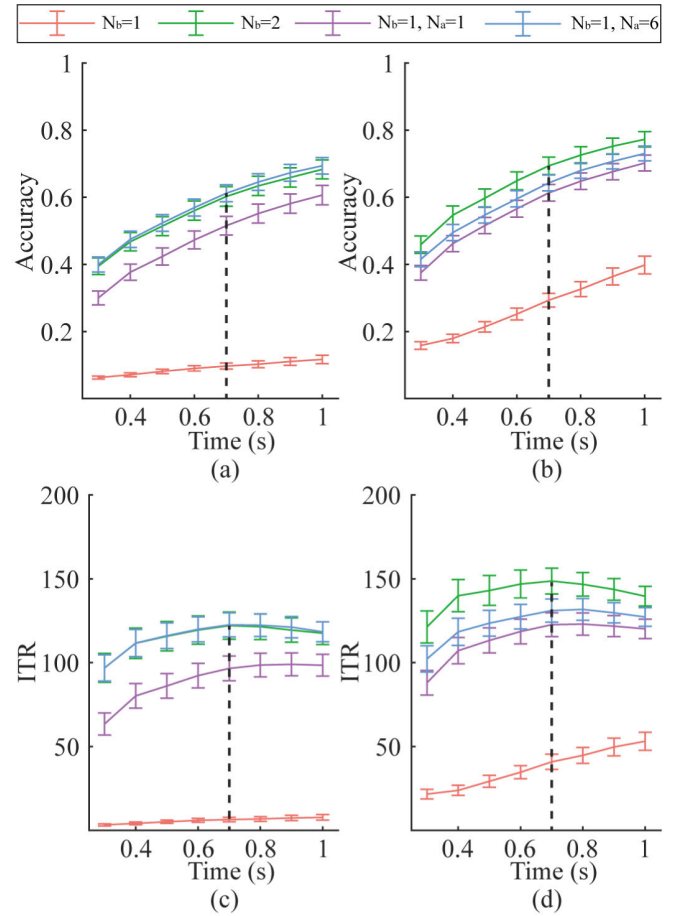


Fig. 7. The classification performance achieved by different calibration conditions. Sub-figures (a) and (b) indicated the accuracy of eTRCA and TDCA, and sub-figures (c) and (d) indicated the ITR of eTRCA and TDCA, respectively. The orange line indicates the condition of only one calibration block with the trial duration of optimal window length T_o , which is selected by the offline analysis. The purple line indicates the condition of only one calibration block with a prolonged trial duration of $T_o + \frac{1}{f_n}$ for each class of f_n Hz. The blue line indicates the condition of only one calibration block with a prolonged trial duration of $T_o + 6 \times \frac{1}{f_n}$. The green line indicates the condition of two calibration block with the trial duration of T_o . Note that PLTS is implemented in the conditions with the prolonged trial duration, whereas the conditions with the trial duration of T_o do not allow for PLTS implementation due to the lack of extra data in each trial.

validated the efficacy of the proposed PLTS data augmentation method.

A. New Time-Shift Augmentation Strategy – PLTS

Data augmentation method could improve the classification performance given the limited amount of training data [40]. Particularly, the time-shift data augmentation method with a fixed step of 50 ms or 100 ms is usually implemented in the BCI field to extract additional training samples in limited training data [32], [33]. However, this augmentation method resulted in low performance when combined with the spatial filtering-based SSVEP algorithms [36], [37]. One of the reasons is that, in spatial filtering-based SSVEP algorithms, all the training samples in one class are temporally averaged to eliminate the noise contained in EEG signals and to obtain the pure SSVEP component, i.e., the temporal templates referred

to in this study. To do so, the initial phases between training samples must maintain the phase-locked requirement, or the SSVEP component would be canceled out. For example, the sum of two signals with the same period but a phase difference of π is 0. The same initial phase of stimulation in each trial and the precise timing in the SSVEP-BCI experiment guarantee the phase-locked requirement [34], [35], but the augmenting process causes different phases between the original training samples and the augmented training samples, resulting in low classification performance.

In this study, PLTS augmentation method augmented new epochs based on the period of the stimulation frequency f_n Hz, so that the phase difference between the SSVEP component contained in the original epochs and those contained in the augmented epochs were integers of 2π , which was equivalent to 0, so that the phase-locked requirement was maintained and the additional augmented training samples enhanced the classification performance. The main difference between the PLTS augmentation method and the conventional augmentation method is that, by utilizing the stimulation frequency information of each class, the time-shift step in PLTS is unique for each class, which is $\frac{f_s}{f_n}$ data samples rather than a fixed step of 50 ms or 100 ms, so that the SSVEP component can be obtained after the temporal averaging and the target can be decoded accurately.

B. Single Calibration Block in SSVEP-BCI

The training-based SSVEP detection methods show advanced classification performance compared to the training-free methods [35], but they need sufficient calibration data to train the detection model. Previous studies reported low performance obtained by the training-based methods when only one training block was available [8], [23]. Since a relatively long stimulation duration of 3 s or 5 s was set in the offline dataset, instead of extracting only one epoch as the conventional studies did, this study proposed the PLTS augmentation method to extract multiple epochs within one trial. Therefore, the eTRCA and TDCA with the PLTS achieved significantly higher classification accuracy, with the improvement ranging from 34.1% to 62.5%.

When conducting a calibration session before the online SSVEP-BCI, the stimulation duration in each trial is set to the optimal time window T_o determined in the offline analysis, so the PLTS can't be implemented directly since there is no additional data in each trial to perform the time shift. Nevertheless, with a slightly prolonged trial duration ($T_o + \frac{1}{f_n}$) and PLTS implementation, the classification performance can be greatly improved. In this study, the ITR was 5.42-40.88 bits/min when only one calibration block was used for training, and the ITR was enhanced to 95.56-163.57 bits/min by PLTS when still one calibration block but with a slightly prolonged duration was used for training. Note that given the stimulation frequency ranging from 8 to 15.8 Hz, the extra stimulation time $\frac{1}{f_n}$ in each trial was 0.064-0.125 s, which is nearly imperceptible since the average human reaction time is 0.2 s [41]. The total extra time was 3.5s, which was small compared to the one-block duration

of 48 s (less than 8 %). To sum up, by slightly prolonging the stimulation duration in each trial and implementing the PLTS, the online SSVEP-BCI system with a high ITR can be achieved under one calibration block.

C. Explanation of the Periodic Similarity

In this study, the similarity between the original epoch x_o and augmented epochs x_a at different time moment in each trial was measured by the Pearson correlation coefficient, and results showed that the maximum correlation occurred periodically, and the period was equal to $\frac{f_s}{f_n}$, where f_n is the stimulation frequency of the given trial. Given two signals, their Pearson correlation coefficient can be high only if they have a similar period and phase. Since the evoked SSVEP in each trial was a periodic signal, x_o and x_a at different time moments, all had the same period; their correlation reached the maximum when the time moment difference between x_o and x_a was integer multiples of period $k \times \frac{f_s}{f_n}$, in which they had the same phase (phase-locked to each other). Since the non-stationary characteristics and noise contained in the measured SSVEP signals, the maximum correlation was less than 1 and varied with time.

Consider the case of zero mean and unit variance in equation (8), it can be simplified into:

$$\begin{aligned} \text{Similarity}(x_o, x_a) &= \frac{\sum_{i=1}^{N_s} (x_o[i]) (x_a[i])}{N-1} \\ &= \frac{1}{N-1} \sum_{i=1}^{N_s} (x_o[i]) (x_o[i+a]) \end{aligned} \quad (11)$$

which was equivalent to the definition of the autocorrelation function (ACF) $R_{x,x}(n)$:

$$R_{x,x}(n) = \sum_{m=-\infty}^{+\infty} x(m) x(m+n). \quad (12)$$

The ACF response of a signal with the period T_n is also a periodic signal, and the maximum correlation occurs periodically at $k \times T_n$, which coincided with our calculated similarity in Fig. 2. Therefore, these results validated the rationality of extracting epochs at time moments $k \times \frac{f_s}{f_n}$ for data augmentation.

D. PLTS Implementation

PLTS is an augmentation data method for SSVEP-BCI and can be effectively combined with the spatial filtering based training methods. With the PLTS, additional time is required to train the detection model since the number of epochs increases from $N_f \times N_b \times 1$ to $N_f \times N_b \times (1 + N_a)$. Still, the training time is acceptable since the original training time of eTRCA and TDCA is relatively short. When a PC with a 2.90 GHz CPU is equipped, and time window is set to 0.7 s and the training block number is 5, the training time is 0.21 s for eTRCA and 1.16 s for TDCA, and the training time with PLTS of $N_a = 3$ is 0.82 s for eTRCA and 5.89 s for TDCA. Besides, it is easy to implement PLTS into the training-based SSVEP-BCI because only one additional step of the time-shift process

to extract augmented epochs in each trial is required, and the spatial filtering calculation, temporal template calculation, and the detection are all the same with the conventional algorithm implementation.

E. Limitations and Future Works

Despite the efficacy of the proposed PLTS was validated in this study, there is still room for further improvement. First, as shown in Fig. 2(b), the similarity level between the augmented and original epochs was different for classes of different stimulus frequencies. This indicates that the optimal augmentation number N_a might be different for each class, and finding the optimal N_a for each individual class could further improve the performance of the PLTS. Second, when the training block N_b was one, the performance improvement brought by the PLTS augmented epochs was mainly achieved by estimating valid spatial filters, rather than valid temporal templates, as Fig. 6 indicated. Similar phenomenon was also found in the SAME study [30]. The possibility of acquiring a valid temporal template using augmented epochs will be investigated in the future. Third, the performance enhancement of the PLTS was less superior when $N_b > 1$ compared to the condition of $N_b = 1$. Meanwhile, the SAME was comparable to the PLTS when $N_b = 2$, and it outperformed the PLTS by 1% when $N_b = 3$. Therefore, the proposed PLTS and the existing data augmentation methods, e.g., the SAME, might be combined using feature fusion technique [42], which may further promote the classification performance, and will also be explored in our future study.

V. CONCLUSION

This study proposes a new time-shift data augmentation method, PLTS, for SSVEP-BCI. By setting a unique time-shift step for each SSVEP class, the new augmented epochs retain similar SSVEP characteristics to the original epochs, and, therefore, can be used for training the spatial filtering-based detection model. The results on two public datasets illustrate that PLTS can significantly improve the performance of the SOTA algorithms, eTRCA and TDCA. Furthermore, the PLTS can solve the insufficient performance problem under a single calibration block, which promotes the ITR from 34.90 ± 6.25 bits/min to 163.57 ± 6.25 bits/min. Therefore, PLTS advances the SSVEP-BCI performance under limited training data and facilitates real-life applications in high-speed brain spellers.

ACKNOWLEDGMENT

The authors would like to thank all the members in their laboratory for their suggestion on this research.

REFERENCES

- [1] B. He, H. Yuan, J. Meng, and S. Gao, "Brain-computer interfaces," in *Neural Engineering*, B. He, Ed. Cham, Switzerland: Springer, 2020, pp. 131–183, doi: 10.1007/978-3-030-43395-6_4.
- [2] F.-B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives," *Prog. Neurobiol.*, vol. 90, no. 4, pp. 418–438, Apr. 2010, doi: 10.1016/j.pneurobio.2009.11.005.
- [3] X. Deng, Z. L. Yu, C. Lin, Z. Gu, and Y. Li, "A Bayesian shared control approach for wheelchair robot with brain machine interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 1, pp. 328–338, Jan. 2020, doi: 10.1109/TNSRE.2019.2958076.
- [4] V. Sakkalis, M. Krana, C. Farmaki, C. Bourazanis, D. Gaitatzis, and M. Padiaditis, "Augmented reality driven steady-state visual evoked potentials for wheelchair navigation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 2960–2969, 2022, doi: 10.1109/TNSRE.2022.3215695.
- [5] X. Chen, X. Huang, Y. Wang, and X. Gao, "Combination of augmented reality based brain-computer interface and computer vision for high-level control of a robotic arm," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 12, pp. 3140–3147, Dec. 2020, doi: 10.1109/TNSRE.2020.3038209.
- [6] L. Chen et al., "Adaptive asynchronous control system of robotic arm based on augmented reality-assisted brain-computer interface," *J. Neural Eng.*, vol. 18, no. 6, Dec. 2021, Art. no. 066005, doi: 10.1088/1741-2552/ac3044.
- [7] 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, doi: 10.1088/1741-2560/12/4/046008.
- [8] 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, doi: 10.1109/TBME.2017.2694818.
- [9] N. Guo et al., "SSVEP-based brain computer interface controlled soft robotic glove for post-stroke hand function rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1737–1744, 2022, doi: 10.1109/TNSRE.2022.3185262.
- [10] H. Lim, S. Kim, and J. Ku, "Distraction classification during target tracking tasks involving target and cursor flickering using EEGNet," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1113–1119, 2022, doi: 10.1109/TNSRE.2022.3168829.
- [11] 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, doi: 10.1109/JSEN.2020.3017491.
- [12] M. Cheng, X. Gao, S. Gao, and D. Xu, "Design and implementation of a brain-computer interface with high transfer rates," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 10, pp. 1181–1186, Oct. 2002, doi: 10.1109/TBME.2002.803536.
- [13] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method," *J. Neural Eng.*, vol. 6, no. 4, Aug. 2009, Art. no. 046002, doi: 10.1088/1741-2560/6/4/046002.
- [14] O. Friman, I. Volosyak, and A. Graser, "Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 742–750, Apr. 2007, doi: 10.1109/TBME.2006.889160.
- [15] Y. Zhang, P. Xu, K. Cheng, and D. Yao, "Multivariate synchronization index for frequency recognition of SSVEP-based brain-computer interface," *J. Neurosci. Methods*, vol. 221, pp. 32–40, Jan. 2014, doi: 10.1016/j.jneumeth.2013.07.018.
- [16] J. Pan, X. Gao, F. Duan, Z. Yan, and S. Gao, "Enhancing the classification accuracy of steady-state visual evoked potential-based brain-computer interfaces using phase constrained canonical correlation analysis," *J. Neural Eng.*, vol. 8, no. 3, Jun. 2011, Art. no. 036027, doi: 10.1088/1741-2560/8/3/036027.
- [17] C. Liu, M. Duan, Z. Duan, A. Liu, Z. Lu, and H. Wang, "An SSVEP-based BCI with LEDs visual stimuli using dynamic window CCA algorithm," *Biomed. Signal Process. Control*, vol. 76, Jul. 2022, Art. no. 103727, doi: 10.1016/j.bspc.2022.103727.
- [18] K. Keng Ang, Z. Yang Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Hong Kong, Jun. 2008, pp. 2390–2397, doi: 10.1109/IJCNN.2008.4634130.
- [19] J. Jiang, C. Wang, J. Wu, W. Qin, M. Xu, and E. Yin, "Temporal combination pattern optimization based on feature selection method for motor imagery BCIs," *Frontiers Hum. Neurosci.*, vol. 14, p. 231, Jun. 2020, doi: 10.3389/fnhum.2020.00231.
- [20] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain-computer interface," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 44, pp. 6058–6067, Nov. 2015, doi: 10.1073/pnas.1508080112.

- [21] J. Huang, P. Yang, B. Xiong, B. Wan, K. Su, and Z.-Q. 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, doi: [10.1109/TNSRE.2022.3162029](https://doi.org/10.1109/TNSRE.2022.3162029).
- [22] 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, doi: [10.1088/1741-2552/abfdfa](https://doi.org/10.1088/1741-2552/abfdfa).
- [23] 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, doi: [10.1109/TNSRE.2021.3114340](https://doi.org/10.1109/TNSRE.2021.3114340).
- [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, doi: [10.1109/TBME.2021.3105331](https://doi.org/10.1109/TBME.2021.3105331).
- [25] Y. Zhang, S. Q. Xie, C. Shi, J. Li, and Z.-Q. 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, doi: [10.1109/TNSRE.2023.3250953](https://doi.org/10.1109/TNSRE.2023.3250953).
- [26] 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," *Exp. Syst. Appl.*, vol. 203, Oct. 2022, Art. no. 117574, doi: [10.1016/j.eswa.2022.117574](https://doi.org/10.1016/j.eswa.2022.117574).
- [27] 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, doi: [10.1088/1741-2552/ab2373](https://doi.org/10.1088/1741-2552/ab2373).
- [28] J. Huang et al., "Incorporating neighboring stimuli data for enhanced SSVEP-based BCIs," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–9, 2022, doi: [10.1109/TIM.2022.3219497](https://doi.org/10.1109/TIM.2022.3219497).
- [29] L. Wang et al., "Enhancing performance of SSVEP-based BCI by unsupervised learning information from test trials," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Montreal, QC, Canada, Jul. 2020, pp. 3359–3362, doi: [10.1109/EMBC44109.2020.9176851](https://doi.org/10.1109/EMBC44109.2020.9176851).
- [30] 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](https://doi.org/10.1109/TBME.2022.3227036).
- [31] 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.
- [32] F. Fahimi, Z. Zhang, W. B. Goh, T.-S. Lee, K. K. Ang, and C. Guan, "Inter-subject transfer learning with an end-to-end deep convolutional neural network for EEG-based BCI," *J. Neural Eng.*, vol. 16, no. 2, Apr. 2019, Art. no. 026007, doi: [10.1088/1741-2552/aaf3f6](https://doi.org/10.1088/1741-2552/aaf3f6).
- [33] D. Chen et al., "Scalp EEG-based pain detection using convolutional neural network," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 274–285, 2022, doi: [10.1109/TNSRE.2022.3147673](https://doi.org/10.1109/TNSRE.2022.3147673).
- [34] 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, doi: [10.1109/TNSRE.2016.2627556](https://doi.org/10.1109/TNSRE.2016.2627556).
- [35] B. Liu, X. Huang, Y. Wang, X. Chen, and X. Gao, "BETA: A large benchmark database toward SSVEP-BCI application," *Frontiers Neurosci.*, vol. 14, p. 627, Jun. 2020, doi: [10.3389/fnins.2020.00627](https://doi.org/10.3389/fnins.2020.00627).
- [36] 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, doi: [10.1088/1741-2552/ab6a67](https://doi.org/10.1088/1741-2552/ab6a67).
- [37] 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, doi: [10.1088/1741-2552/aae5d8](https://doi.org/10.1088/1741-2552/aae5d8).
- [38] 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, doi: [10.1371/journal.pone.0140703](https://doi.org/10.1371/journal.pone.0140703).
- [39] 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.
- [40] Y. Pei et al., "Data augmentation: Using channel-level recombination to improve classification performance for motor imagery EEG," *Frontiers Human Neurosci.*, vol. 15, Mar. 2021, Art. no. 645952, doi: [10.3389/fnhum.2021.645952](https://doi.org/10.3389/fnhum.2021.645952).
- [41] I. J. Deary and G. Der, "Reaction time, age, and cognitive ability: Longitudinal findings from age 16 to 63 years in representative population samples," *Aging, Neuropsychol., Cognition*, vol. 12, no. 2, pp. 187–215, Jun. 2005, doi: [10.1080/13825580590969235](https://doi.org/10.1080/13825580590969235).
- [42] Y. Pei et al., "A tensor-based frequency features combination method for brain-computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 465–475, 2022, doi: [10.1109/TNSRE.2021.3125386](https://doi.org/10.1109/TNSRE.2021.3125386).