Article:
Huang, J, Yang, P, Xiong, B et al. (5 more authors) (2022) Incorporating Neighboring Stimuli Data for Enhanced SSVEP-Based BCIs. IEEE Transactions on Instrumentation and Measurement, 71. 2521109. ISSN 0018-9456

https://doi.org/10.1109/TIM.2022.3219497

Incorporating Neighboring Stimuli Data for Enhanced SSVEP-Based BCIs

Jiayang Huang, Pengfei Yang, *Member, IEEE*, Bang Xiong, Quan Wang, Bo Wan, Ziling Ruan, Keyi Yang, and Zhi-Qiang Zhang, *Member, IEEE*

***Abstract*—Various spatial filters have been proposed to enhance the target identification performance of steady-state visual evoked potential (SSVEP)-based brain-computer interfaces (BCIs). The current methods only extract the target-related information from the corresponding stimulus to learn the spatial filter parameter. However, the SSVEP data from neighboring stimuli also contain frequency information of the target stimulus, which could be utilized to further improve the target identification performance. In this paper, we propose a new method incorporating SSVEPs from the neighboring stimuli to strengthen the target-related frequency information. First, The spatial filter is obtained by maximizing the summation of covariances of SSVEP data corresponding to the target and its neighboring stimuli. Then the correlation features between spatially filtered templates and test data are calculated for target detection. For the performance evaluation, we implemented the offline experiment using the 40-class benchmark dataset from 35 subjects and the 12-target self-collected dataset from 11 subjects. Compared with the state-of-art spatial filtering methods, the proposed method showed superiority in classification accuracy and information transfer rate (ITR). The comparison results demonstrate the effectiveness of the proposed spatial filter for target identification in SSVEP-based BCIs.**

***Index Terms*—Brain-computer interfaces (BCIs), neighboring stimuli, spatial filter, steady-state visual evoked potential (SSVEP).**

I. Introduction

Brain-computer interfaces (BCIs) have emerged as the direct channel for people to interact with the outside world through brain activities [1]. Due to high signal-to-noise ratios (SNRs) and fast information transfer rates (ITRs) [2], the steady-state visual evoked potential (SSVEP)-based BCI [3] has been one of the most popular electroencephalography (EEG)-based BCIs [4]. Specifically, SSVEP-based BCIs have been used in varieties of applications, such as prosthesis control [5], spelling [6], and robot control [7].

This work was supported in part by the Shaanxi Key Technology R&D Program under Grant No. 2021ZDLGY07-01, in part by the National Natural Science Foundation of China under Grants No.61972302 and No.61962019. (Corresponding author: Pengfei Yang; Zhi-Qiang Zhang.)

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 under Application No. 20200017.

Jiayang Huang, Pengfei Yang, Bang Xiong, Quan Wang, Bo Wan, Ziling Ruan, and Keyi Yang are with School of Computer Science and Technology, Xidian University, Xi'an, 710071, China (e-mail: jyhuang1@stu.xidian.edu.cn; pfyang@xidian.edu.cn; bxiong@stu.xidian.edu.cn; qwang@xidian.edu.cn; wanbo@xidian.edu.cn; 21031211775@stu.xidian.edu.cn; 21031211507@stu.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, Leeds LS2 9JT, U.K(e-mail: z.zhang3@leeds.ac.uk).

Target identification is an essential step for constructing SSVEP-based BCIs, which is to translate SSVEP signals into commands [8]. So far, researchers have proposed various spatial filters for target identification in SSVEP-based BCIs [9]. Canonical correlation analysis (CCA) [10], the most widely used spatial filter, was used to find a pair of weights for the maximal correlation coefficient between EEG signals and sine-cosine reference signals. To further boost the SSVEP detection performance, expansions of CCA developed from two aspects. On the one hand, CCA was combined with signal processing techniques, such as filter bank CCA (FBCCA) [11], binary subband CCA (BsCCA) [12], and multivariate variational mode decomposition CCA (MVMD-CCA) [13], to extract SSVEP harmonic frequency information for target recognition. On the other hand, individual calibration data was incorporated into CCA, such as L1-regularized multi-way CCA (L1-MCCA) [14], multi-set CCA (MsetCCA) [15], individual template CCA (IT-CCA) [16], and extended CCA (eCCA) [17], to improve the SSVEP frequency detection performance. Besides CCA-based expansions, Nakanishi et al. introduced task-related component analysis (TRCA) into SSVEP-based BCIs [18] by maximizing the inter-trial covariance of individual data. Zhang et al. proposed correlated component analysis (CORCA) [19] to construct template signals by maximizing inter-subject covariance. Kiran Kumar et al. proposed the sum of squared correlation (SSCOR) method [20] to maximize inter-session individual data. In the aforementioned algorithms, the spatial filters were obtained with SSVEP data corresponding to the target stimulus, but the information from other stimuli was understudied. It is reported that SSVEPs from different stimuli share a common spatial pattern [21] [22], so incorporating data from other stimuli to train the spatial filter would make further improvements in target detection [18]. Moreover, according to the competitive interactions of SSVEPs [23] [24], the target SSVEPs contain frequency information from neighboring stimuli at close distance, and vice versa. Therefore, SSVEPs across neighboring-location stimuli share a common spatial pattern and contain common frequency information. In this way, utilizing the training data from not only the target stimulus but also the nearby-location stimuli to train a spatial filter would be an effective way to improve target frequency detection performance.

To further enhance the target detection performance of SSVEP-based BCIs, SSVEP data from other stimuli was also utilized in some previous studies. For instance, Cohen et al. proposed rhythmic entrainment source separation (RESS)

[25] which directly suppresses the adjacent frequencies from the target SSVEP. Zhang et al. proposed a multi-objective optimization-based high-pass spatial filtering method (termed as MOO) [26] to reduce the influence of target-irrelevant information by minimizing the correlation between the target and other stimuli SSVEPs. Both methods suppress the adjacent frequency information to improve the target detection performance, which would lose target-related frequency information from other stimuli. Nakanishi et al. proposed ensemble TRCA (eTRCA) [18] which integrates the spatial filters from all targets to make further improvements, but the ensemble technique would yield redundancy and greatly increase the computation costs. Wong et al. proposed multi-stimulus eCCA (ms-eCCA) and multi-stimulus eTRCA (ms-eTRCA) [27] which concatenates EEG data from the target and its neighboring stimuli for spatial filter training. Here, the neighboring stimuli denote the stimuli with nearby flickering frequencies. The multi-stimulus methods utilized the information from nearby-frequency stimuli to extract target-related components, but they neglected the neighboring-location stimuli which do not flash at nearby frequencies.

To address the drawbacks above, a new method incorporating SSVEP data from the neighboring-location stimuli is proposed to enhance the target identification performance of SSVEP-based BCIs. The main contributions of this study include: 1) a novel spatial filtering method is introduced, which incorporates SSVEP data from neighboring-location stimuli to strengthen the target-related frequency information; 2) the spatial filter can effectively improve the target detection performance and save computational costs; 3) a more efficient training process, which works with small training data while achieving high classification performance; 4) a self-collected SSVEP dataset containing 12 visual stimulations to verify the effectiveness of this proposed algorithm. The proposed method is also verified by the 40-class benchmark dataset [28] with the offline experiment. Extensive comparisons were conducted with MVMD-CCA, TRCA, SSCOR, RESS, MOO, and ms-TRCA on classification performance. The experimental results illustrate the superiority of the proposed spatial filter in the aspects of classification accuracy, ITR, sensitivity, and specificity, which demonstrates the efficiency of using SSVEP data from the target and its neighboring stimuli for improving the target identification performance of SSVEP-based BCIs. The rest of the article is arranged as follows: Section II introduces the materials and methods. In section III, the experimental results with discussions are reported. Finally, the conclusion is presented in the last section.

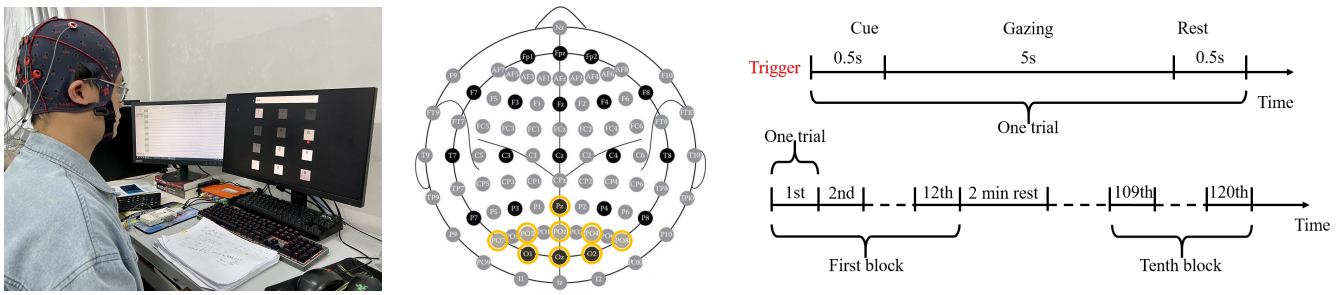
II. Materials and Methods

A. Data Description

The public benchmark dataset [28] and a self-collected dataset (termed Dataset I) were used in this study. The benchmark dataset consists of 64-channel EEG data collected from 35 healthy subjects stimulated by 40 stimuli, which is modulated at 8–15.8 Hz frequencies with an interval of 0.2 Hz and 0-1.5π phases with an interval of 0.5π using joint frequency and phase modulation (JFPM) [29]. The collected data were downsampled to 250 Hz. The dataset contains 6-block data, and each block is comprised of 0.5-s gaze shifting cue, 5-s stimulation, and 0.5-s rest. The EEG data from 9 channels (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, and O2) were used for performance evaluation in this study. In addition, the details of the self-collected dataset are described as follows.

1) Subjects: Eleven healthy subjects who participated in this study (four females and seven males, aged from 23 to 27) have normal or corrected-to-normal vision and no brain-related diseases. During the experiment, each subject was asked to sit in a comfortable position in a dark and silent room at a distance of 60 cm in front of the LCD screen (Fig. 1(a)). All subjects were informed of the experimental process and protocols and signed the informed consent before the experiment. The experiment is approved by the Research Ethics Committee of Xidian University under Application No. 20200017.

2) Stimulation Interface Design: In this study, a 4×3 matrix of visual stimuli was designed, which is coded by the JFPM method. The interface was displayed on a 23.6-inch LCD screen which has a 1920-×1080-pixel resolution and a 60-Hz refresh rate. The frequency ranged from 9.25 Hz to 14.75 Hz with an interval of 0.5 Hz, and the phase range was from 0π to 1.5π with an interval of 0.5π [18] [28]. The stimulation program was coded with MATLAB using the Psychophysics Toolbox Version 3 [30].



(a) (b) (c)

Fig. 1. The EEG recording environment is shown in (a). The electrodes used for EEG collection are circled in yellow (b). The experimental paradigm design is presented in (c).

3) EEG Collection: The EEG collection equipment is g.USBamp-Research with 256-Hz sampling rate. According to the 10-20 standard system, 9 Ag/AgCl electrodes (Pz, PO7, PO3, POz, PO4, PO8, O1, Oz, and O2) were selected from the parietal and occipital regions (Fig. 1(b)) which is the reflection zone of SSVEP signals. The ground and reference electrodes were placed at FPz and right earlobe respectively. During the EEG collection, electrode impedances were kept below 10 kΩ and an event trigger produced by the stimulation program was sent through the parallel port to the amplifier and computer simultaneously. The event trigger was to mark the time when all stimuli began to flash simultaneously. For each subject, 6-block EEG data were collected in the BCI experiments. The same as the benchmark dataset, each block contains 0.5-s cue, 5-s stimulation, and 0.5-s rest. During the stimulation, subjects were asked to avoid eye blinks. To prevent visual fatigue, there was a two-minute rest between two successive blocks (Fig. 1(c)).

B. Data Preprocessing

Firstly, the 9-channel EEG data epochs were extracted as [0.64 s 0.64+d s] according to the 0.5-s cue and the 0.14-s visual latency, where d is the size of the time window for target recognition. And then, all the extracted data epochs were filtered by a 6-order Butterworth filter with the 7-90 Hz band. A notch filter at 50 Hz is utilized to eliminate the power-line noise. After the preparation was completed, all the data processing and target detection were then performed.

C. Target Identification Method

In this section, the spatial filter using SSVEP data from the target and its neighboring stimuli is introduced and used as the target identification method. We will elaborate on neighbor definition, spatial filter training, and testing below.

1) Neighbors Definition: As shown in the figure 2, the neighbors of the n-th stimulus are defined as the horizontally and vertically adjacent stimuli. Due to the location of the target stimulus, the number of its neighbors m varies between 2 and 4. To simply the description, we denote the neighbors of n-th stimulus as the l_n^1-th, l_n^2-th, ..., l_n^m-th stimuli.

2) Training stage: The individual calibration data for the n-th stimulus are denoted as $\boldsymbol{\mathcal{X}}_n \in \mathbb{R}^{N_c \times N_s \times K}$, where N_c, N_s, and K represent the number of channels, the data length of the data epoch, and the number of training trials respectively.

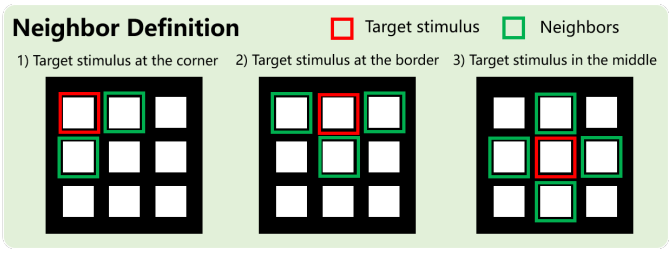


Fig. 2. The neighbors of the target stimulus are defined as its horizontally and vertically adjacent stimuli, and the number of its neighbors varies between 2 and 4.

For any trial h ($h = 1, 2, \cdots, K$), the calibration data for the n-th stimulus is denoted as $\boldsymbol{\mathcal{X}}_n^h \in \mathbb{R}^{N_c \times N_s}$, so that

$$\boldsymbol{\mathcal{X}}_n = \left[\boldsymbol{\mathcal{X}}_n^1, \boldsymbol{\mathcal{X}}_n^2, \cdots, \boldsymbol{\mathcal{X}}_n^K\right]. \tag{1}$$

The collection of the neighbors of the n-th stimulus is defined as

$$\boldsymbol{Y}_n = \left\{\boldsymbol{\mathcal{X}}_{l_n^1}, \boldsymbol{\mathcal{X}}_{l_n^2}, \cdots, \boldsymbol{\mathcal{X}}_{l_n^m}\right\}, \tag{2}$$

where $\boldsymbol{\mathcal{X}}_{l_n^i}$ is the training data from l_n^i-th stimulus, and $i = 1, 2, \cdots, m$. In this method, both $\boldsymbol{\mathcal{X}}_n$ and $\boldsymbol{Y}_n$ are used as the training data to obtain the spatial filter for the n-th stimulus.

With the training data, the filter bank technique [11] is first applied to extract information from the harmonic frequencies. The subbands of the filter bank are implemented by zero-phase Chebyshev type I infinite impulse response (IIR) filters. The b-th subband is at the frequency range of [b×8 Hz, 88 Hz]. For each subband, there adds 2 Hz bandwidth to both sides of the passband. For each $\boldsymbol{\mathcal{X}}_j^h$, it will be decomposed into N_b sub-band components $\boldsymbol{\mathcal{X}}_j^{h,1}, \boldsymbol{\mathcal{X}}_j^{h,2}, \cdots, \boldsymbol{\mathcal{X}}_j^{h,N_b}$, where $j = 1, 2, \cdots, N_f$, N_f is the number of stimuli and N_b is the total number of subbands.

For each subband b ($b = 1, 2, \cdots, N_b$) , the spatial filter for the n-th stimulus is obtained by maximizing the summation of the auto-covariances of the templates. The summation of the auto-covariances of $\overline{\boldsymbol{\mathcal{X}}}_n^b$ and $\overline{\boldsymbol{\mathcal{X}}}_{l_n^i}^b$ is denoted as:

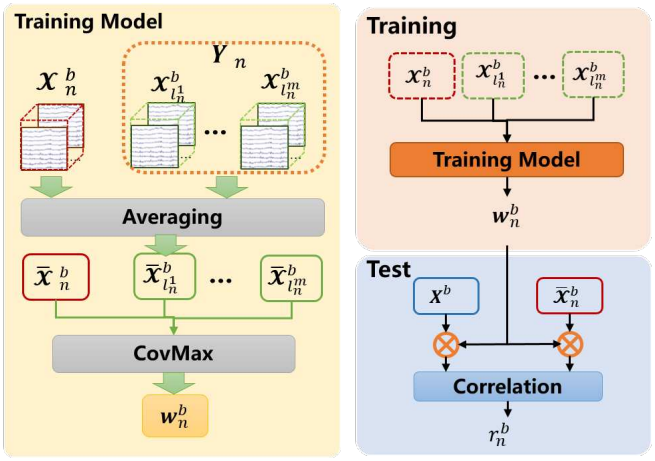
$$\boldsymbol{S} = \operatorname{cov}\left(\overline{\boldsymbol{\mathcal{X}}}_n^b\right) + \sum_{i=1}^{m} \operatorname{cov}\left(\overline{\boldsymbol{\mathcal{X}}}_{l_n^i}^b\right) \tag{3}$$

where

$$\overline{\boldsymbol{\mathcal{X}}}_n^b = \frac{1}{K}\sum_{h=1}^{K} \boldsymbol{\mathcal{X}}_n^{h,b} \tag{4}$$

and

$$\overline{\boldsymbol{\mathcal{X}}}_{l_n^i}^b = \frac{1}{K}\sum_{h=1}^{K} \boldsymbol{\mathcal{X}}_{l_n^i}^{h,b}. \tag{5}$$



(a) (b)

Fig. 3. The proposed training model of each sub-band is shown in (a). The flowchart of the proposed target identification method is illustrated in (b).

To obtain a finite solution, the sum of covariances of all-trial data from n-th stimulus is constrained as:

$$\boldsymbol{w}^\top \left(\sum_{h=1}^{K} \mathrm{cov}\left(\boldsymbol{\mathcal{X}}_n^{h,b} \right) \right) \boldsymbol{w} = 1. \tag{6}$$

The constraint also applies to the neighboring-stimulus data, which means for any i,

$$\boldsymbol{w}^\top \left(\sum_{h=1}^{K} \mathrm{cov}\left(\boldsymbol{\mathcal{X}}_{l_n^i}^{h,b} \right) \right) \boldsymbol{w} = 1. \tag{7}$$

The constrained optimization problem is defined as:

$$\hat{\boldsymbol{w}}_n^b = \underset{\boldsymbol{w}}{\mathrm{argmax}} \frac{\boldsymbol{w}^\top \boldsymbol{S} \boldsymbol{w}}{\boldsymbol{w}^\top \boldsymbol{Q} \boldsymbol{w}}, \tag{8}$$

where

$$\boldsymbol{Q} = \sum_{h=1}^{K} \mathrm{cov}\left(\boldsymbol{\mathcal{X}}_n^{h,b} \right) + \sum_{i=1}^{m} \sum_{h=1}^{K} \mathrm{cov}\left(\boldsymbol{\mathcal{X}}_{l_n^i}^{h,b} \right). \tag{9}$$

Via solving the equation (8) with generalized eigendecomposition of $\boldsymbol{Q}^{-1}\boldsymbol{S}$, the spatial filter $\hat{\boldsymbol{w}}_n^b$ is determined as the eigenvector corresponding to the largest eigenvalue. The constrained optimization process will repeat N_b times to ensure each subband gets its spatial filter.

3) Test stage: With the spatial filter $\boldsymbol{w}_n^b$ trained by the training model, the target frequency of the single-trial test data $\boldsymbol{X} \in \mathbb{R}^{N_c \times N_s}$ can be identified, as shown in Fig. 3(b). Firstly, with filter bank analysis, the single-trial test data $\boldsymbol{X} \in \mathbb{R}^{N_c \times N_s}$ is decomposed into N_b subbands with $\boldsymbol{X}^b$ as the data sub-band b. Then the test data $\boldsymbol{X}^b$ and the n-th stimulus template $\overline{\boldsymbol{\mathcal{X}}}_n^b$ are respectively spatially filtered with the optimal solution $\boldsymbol{w}_n^b$. For any sub-band b, the correlation coefficient between the spatially filtered test data and template is calculated as:

$$\gamma_n^b = \rho\left(\left(\boldsymbol{X}^b \right)^\top \boldsymbol{w}_n^b, \left(\overline{\boldsymbol{\mathcal{X}}}_n^b \right)^\top \boldsymbol{w}_n^b \right), \tag{10}$$

where $\rho(s_1, s_2)$ is the Pearson's correlation coefficient between s_1 and s_2 [31]. By integrating the γ_n^b from all subbands, the final correlation feature γ_n is calculated as:

$$\gamma_n = \sum_{b=1}^{N_b} c(b) \cdot \left(\gamma_n^b \right)^2, \tag{11}$$

where the weights for the subband component $c(b) = b^{-1.25} + 0.25$ is to maximize the classification performance [11]. The target frequency $\hat{f}$ with the largest correlation coefficient is defined as:

$$\hat{f} = \underset{n}{\mathrm{argmax}}\, \gamma_n, n = 1, 2, \cdots, N_f. \tag{12}$$

D. Performance Evaluation

In this study, the classification accuracy, ITR, sensitivity, and specificity estimates were computed to evaluate the target identification performance of the proposed method. The estimates were calculated by using leave-one-out cross-validation. For both datasets with 6 blocks, 5-block data were used as the training dataset and 1-block data was used as the test data (i.e., K = 5). In the offline experiment, each-block data was used as the test data, so the whole procedure was repeated 6 times for each subject. Therefore, for each subject, the overall result was obtained by averaging across blocks.

The classification accuracy is defined as the percentage of the correct predictions out of all predictions. ITR is the amount of information transferred per minute, defined as:

$$\mathrm{ITR} = \frac{60}{T} \times \left[\log_2 N_f + P \times \log_2 P + (1-P) \times \log_2 \left(\frac{1-P}{N_f - 1} \right) \right], \tag{13}$$

where T is the selection time for each target, including gazing time and 0.5-s gaze-shifting time, N_f is the number of stimuli, and P represents the classification accuracy. Besides, the sensitivity and specificity of the classification model were also calculated to further evaluate the target detection performance of the proposed method. Sensitivity refers to a true positive rate, which is defined as:

$$\mathrm{Sensitivity} = \frac{TP}{TP + FN}. \tag{14}$$

Specificity refers to the true negative rate, defined as:

$$\mathrm{Specificity} = \frac{TN}{TN + FP}. \tag{15}$$

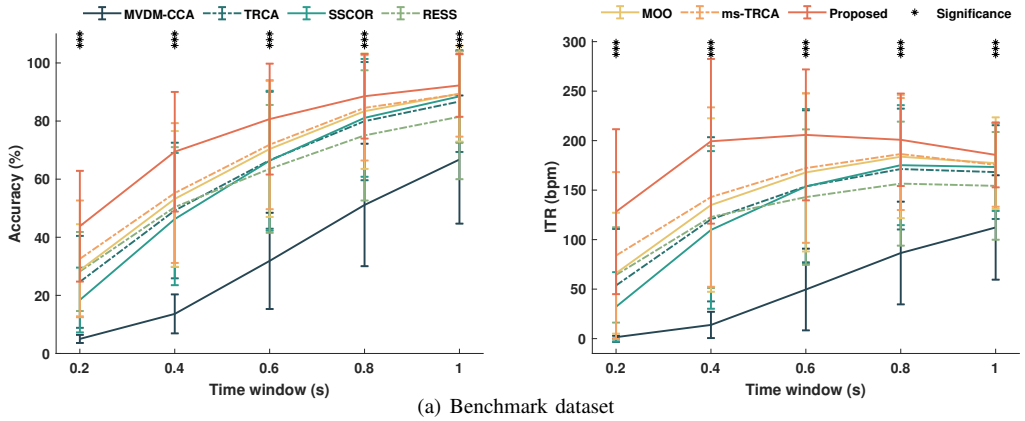
The terms TP, TN, FP, and FN represent the number of true positives, true negatives, false positives, and false negatives respectively.

III. Experimental results with Discussions

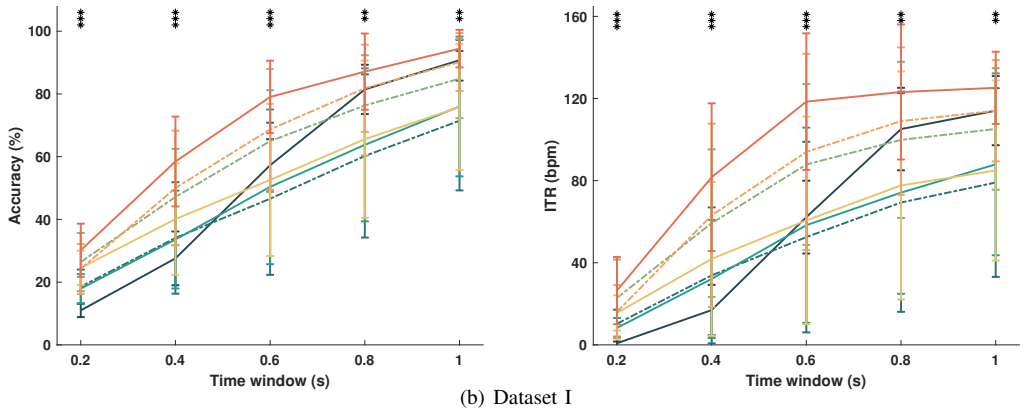
In this section, we first evaluated the performance of the proposed method on the 40-target benchmark dataset [28] and the 12-target self-collected SSVEP dataset. Extensive comparisons of target detection performance evaluated by the classification accuracy, ITR, sensitivity, and specificity were implemented between the proposed method and state-of-the-art SSVEP target recognition methods. And then, the influences of different parameters such as the filter bank analysis, the number of training blocks, and the number of neighbors were also reported. Based on the experimental results, the discussions were presented in the last subsection.

A. Target Recognition Performance

The target recognition performance was compared between the proposed method and state-of-art methods, such as MVMD-CCA, TRCA, SSCOR, RESS, MOO, and ms-TRCA. From the results shown in Fig. 4, the accuracies of all methods gradually increase with the time windows. Generally, the proposed method reached the highest accuracies and ITRs among all methods regardless of time windows. To be specific, in Fig. 4(a), the highest ITRs of the proposed method are obtained with a 0.6-s time window reaching 205.79±66.13 bpm. With 12-target Dataset I (Fig. 4(b)), the proposed method reached the highest ITR of 123.16±32.87 bpm with a 0.8-s time window. To intuitively reveal the significant difference between our proposed method and other compared methods, pairwise analyses were implemented using paired t-tests on these two datasets. Table I listed the multiple comparison



(a) Benchmark dataset



(b) Dataset I

Fig. 4. Averaged accuracies and ITRs across subjects of the proposed method and six compared methods as MVMD-CCA, TRCA, SSCOR, RESS, MOO, and ms-TRCA using different time windows on the benchmark dataset (a) and Dataset I (b). The time window increases from 0.2 s to 1.0 s with a step of 0.2 s. The error bar represents the standard deviation. The asterisks indicate significant differences between the seven methods obtained by paired t-tests ($*p<0.05$, $**p<0.01$, $***p<0.001$).

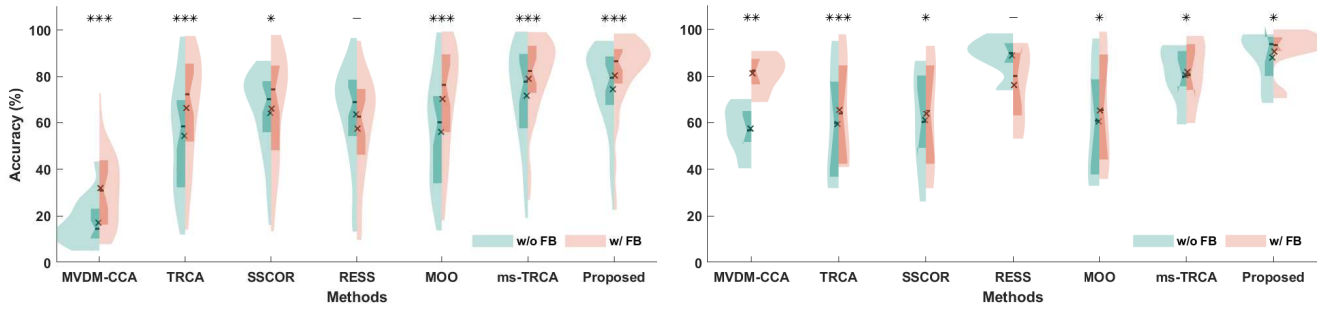
TABLE I
THE SIGNIFICANT DIFFERENCE IN CLASSIFICATION ACCURACY BETWEEN THE PROPOSED AND OTHER STATE-OF-ART METHODS OBTAINED BY PAIRED T-TESTS

Methods	**Time windows**									
	Benchmark Dataset					**Dataset I**				
	0.2 s	***0.4 s***	***0.6 s***	***0.8 s***	***1.0 s***	***0.2 s***	***0.4 s***	***0.6 s***	***0.8 s***	***1.0 s***
MVMD-CCA vs. Proposed method*	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0001	0.0476	0.1819
TRCA vs. Proposed method	<0.0001	<0.0001	0.0074	0.0464	0.1127	0.0011	0.0020	0.0007	0.0055	0.0035
SSCOR vs. Proposed method	<0.0001	<0.0001	0.0075	0.0824	0.2498	0.0004	0.0009	0.0023	0.0100	0.0154
RESS vs. Proposed method	<0.0001	<0.0001	<0.0001	0.0039	0.0100	0.3335	0.0893	0.0295	0.0497	0.0335
MOO vs. Proposed method	0.0006	0.0029	0.0475	0.2199	0.4005	0.0795	0.0147	0.0039	0.0184	0.0078
ms-TRCA vs. Proposed method	0.0190	0.0098	0.0497	0.3126	0.3345	0.0960	0.1415	0.0417	0.0522	0.2109

* The proposed method serves as the control group.

results in terms of SSVEP target recognition accuracy with different time windows. It is shown that with the time window smaller than 0.6 s, the proposed outperformed the other competing methods by a significant margin especially with the benchmark dataset, verifying the effectiveness of using information from neighboring stimuli. With a time window larger than 0.8 s, the differences in accuracies were no longer statistically significant, since the accuracies for different methods are gradually close to 100%. In conclusion, by using SSVEPs corresponding to the target and its neighboring stimuli, the proposed spatial filter can effectively improve the SSVEP stimulus recognition performance.

Apart from classification accuracy and ITR, the results of the sensitivity and specificity are also calculated as the evaluation of the target recognition performance. Table II reports the averaged sensitivities and specificities across subjects with two different datasets. From the results listed in the table, it can be seen that the sensitivity results are consistent with the accuracy and ITR results. The proposed method reached the highest sensitivity among all seven competing methods, which indicates that the proposed method shows the greatest capability of detecting positive samples. While for the specificities, all methods reached the results close to 100% and had slight differences from each other, indicating that these methods can correctly classify negative samples. Both sensitivity and specificity results further demonstrate that



(a) Benchmark dataset (b) Dataset I

Fig. 5. Classification accuracy distributions of the four competing methods without filter bank technique (w/o FB) and with filter bank analysis (w/ FB) on two datasets. Here the time windows were set to 0.6 s and 0.8 s respectively for the two datasets. The top and bottom sides of the violin represent the maximum and minimum without an outlier. The width of the violin refers to the probability density estimate. For the box in the violin, the top and bottom of the box represent the 25th and 75th percentiles respectively. The middle line and the x-mark indicate the median value and the mean value respectively. The asterisks indicate significant differences between methods w/o FB and w/ FB obtained by paired t-tests (∗p<0.05, ∗∗p<0.01, ∗∗∗p<0.001).

TABLE II
THE AVERAGED SENSITIVITY AND SPECIFICITY RESULTS ACROSS SUBJECTS OF THE PROPOSED AND STATE-OF-ART METHODS

Datasets	Methods	Sensitivity (%)	Specificity (%)
Benchmark dataset	MVDM-CCA	33.62±16.56	98.21±0.42
	TRCA	66.70±23.49	99.15±0.60
	SSCOR	81.86±19.47	99.53±0.49
	RESS	55.76±21.19	98.86±0.54
	MOO	70.23±22.49	98.56±0.57
	ms-TRCA	76.56±21.05	99.26±0.52
	Proposed method	**82.04±19.60**	**99.54±0.50**
Dataset I	MVDM-CCA	72.27±10.68	97.47±0.97
	TRCA	59.09±20.54	96.28±1.86
	SSCOR	70.83±21.84	97.35±1.99
	RESS	78.88±13.58	98.08±1.23
	MOO	69.18±15.58	97.28±1.07
	ms-TRCA	80.24±19.68	98.53±1.25
	Proposed method	**87.12±12.14**	**98.83±1.10**

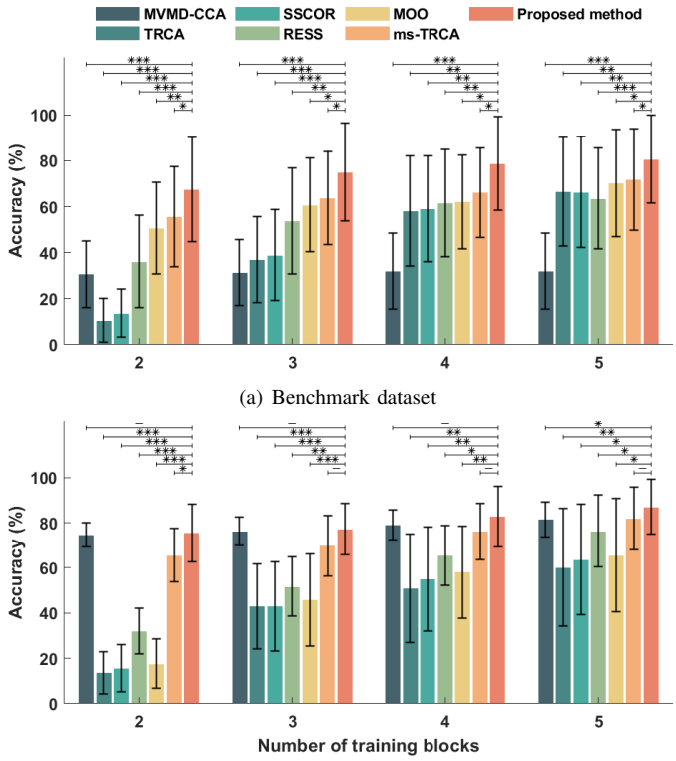
The data lengths were set to 0.6 s and 0.8 s respectively for the two datasets, where the highest ITRs were achieved.

the proposed method can effectively boost the target detection performance.

B. The Influence of Parameters on Performance

In order to further evaluate the performance of the proposed algorithm, we explored the impact of the filter bank analysis, the number of training blocks, and the number of neighbors on the SSVEP target recognition accuracy.

1) The filter bank analysis: The filter bank method is a widely-used signal processing technique to decompose signals into multiple sub-bands for extracting harmonic frequency components. To explore the impact of the filter bank analysis on target recognition, we conducted comparisons between various methods with and without filter bank analysis. Here, the numbers of subbands were set as 5 and 3 for two datasets respectively, where the highest accuracies were achieved. Fig. 5 depicts the classification accuracy distributions of the seven methods without filter bank technique (w/o FB) and with filter bank analysis (w/ FB) on two datasets via the half violin plot. From the figure, it is shown that the classification accuracies of the proposed method are higher than those of the other six methods with or without filter bank analysis. To be



(a) Benchmark dataset

(b) Dataset I

Fig. 6. Averaged classification accuracies across subjects of the benchmark dataset at 0.6 s data length (a) and Dataset I at 0.8 s data length (b) with different numbers of training blocks varying from 2 to 5. The vertical error bars represent standard deviations. The asterisks indicate significant differences between the seven methods obtained by paired t-tests (∗p<0.05, ∗∗p<0.01, ∗∗∗p<0.001).

specific, the proposed method reached the highest mean value and median value among the competing methods under two conditions. And except for the RESS method, the classification performance has been improved by employing the filter bank analysis due to the information from the harmonic frequencies.

2) The number of training blocks: In the proposed method, the spatial filters were learned from individual training data. The impact of training blocks on target identification performance can also be investigated. Figure 6 presents the averaged

classification accuracies across subjects obtained with different numbers of training blocks (K) on two datasets. As the figures illustrated, the averaged target recognition accuracies of TRCA, SSCOR, RESS, and MOO methods gradually increased with the number of training blocks, while those of ms-TRCA and the proposed method increased slightly with the number of training blocks with 5% accuracy. Besides, there shows no obvious difference in the MVDM-CCA classification accuracy as the number of training blocks increases, since the MVDM-CCA is a training-free method. From Fig. 6(a), the proposed method and ms-TRCA can still achieve satisfactory performance with only 2-block training data, which implies that employing data from neighboring stimuli for spatial filter training would reduce the sensitivities to the number of training blocks. The result in Fig. 6(b) shows consistency with the benchmark dataset, that is, with more than 2-block training data, the proposed method can get satisfactory classification results. Overall, these results indicate that even with limited training blocks, the proposed method reached the highest classification accuracy among all methods, showing an improvement in target recognition performance.

3) The number of neighbors: By employing SSVEPs from both target and its neighboring stimuli, the proposed method obtained enhanced performance on target identification. It remains unclear how the number of neighbors influences the target recognition performance, which needs further investigation. The results of the classification accuracies with the different number of neighbors on the benchmark dataset are set out in Fig. 7. The comparison was conducted between ms-TRCA and our proposed method at 0.6-s data length with the number of neighbors (m) increasing from 0 to 8. In the proposed method, the neighbors are defined as the nearby-location stimuli. Therefore, when $m = 4$, it represents the horizontal and vertical neighbors. When $m = 8$, it indicates all neighbors surrounding the target stimulus, including horizontal, vertical, and diagonal neighbors. It is also worth pointing out here that $m = 4$ or 8 only applies to those stimuli in the middle, and the stimuli on the border or corner will have a smaller number of neighbors for both cases. From the graph, we can see that as the number of neighbors increases from 0 to 4, the accuracies

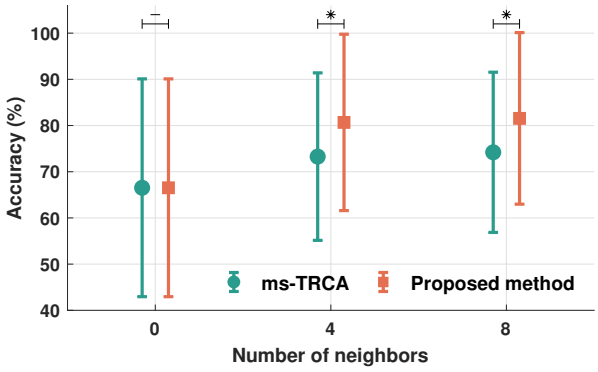


Fig. 7. The averaged target recognition accuracies across all subjects with different numbers of neighbors varying from 0 to 8 on the benchmark dataset. The vertical error bars represent standard deviations. The asterisks indicate significant differences between the two methods obtained by paired t-tests (∗p<0.05, ∗∗p<0.01, ∗∗∗p<0.001).

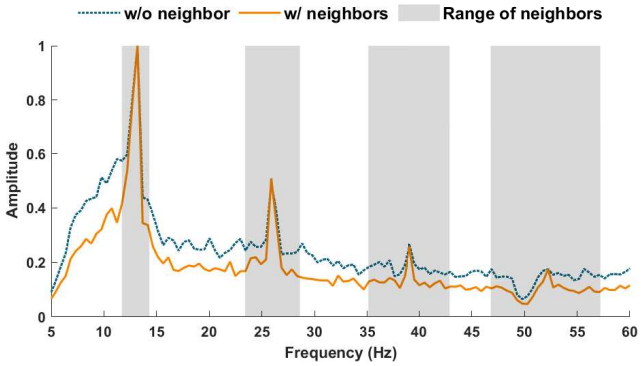


Fig. 8. The frequency spectrums correspond to the 2.0-s SSVEPs of 13 Hz stimulus with (w/) and without(w/o) the neighboring stimuli respectively. The frequency amplitudes were computed with Fourier transform and then normalized to 1. And the grey area represents the frequency range of the neighboring stimuli.

of the two methods show increments of more than 6%. While as the number of neighbors increases from 4 to 8, the accuracy of the proposed method remains stable with a slight difference, while ms-TRCA showed a slight increment with 2%. Nevertheless, the proposed method showed better performance than ms-TRCA regardless of the number of neighbors. Therefore, the incorporation of neighboring stimuli data does contribute to improving target detection performance, and the horizontal and vertical neighbors are adequate for improving the target identification performance.

C. Discussions

1) The frequency spectrum analysis: In this study, we constructed a novel spatial filter incorporating SSVEPs from the neighboring stimuli for target identification. From the experimental results, it is demonstrated that learning a spatial filter by maximizing the summation of covariances of SSVEP data corresponding to the target and its neighboring stimuli is an effective solution for enhancing the performance of SSVEP-based BCIs. The spatially filtered template signals would include more accurate frequency information of the SSVEP response. To take an example of 13 Hz stimulus, the frequency spectrums of 2.0-s SSVEP signals with(w/) and without(w/o) the neighboring stimuli are provided in Fig. 8. From the plot, it can be seen that without incorporating the neighboring stimuli, target SSVEPs include not only the major target frequency but also the frequency information from the neighboring stimuli. By contrast, the SSVEPs spatially filtered by the proposed method mainly contain the target frequency information. It demonstrates that the proposed method effectively strengthens the target-related frequency information, leading to higher classification accuracies.

2) The discriminability of correlation coefficients: With more accurate target-relevant frequency information extracted from the EEG signals, more discriminative correlation coefficients would be obtained for target identification. To intuitively present the contribution of the accurate frequency information, we choose 13 Hz as the exemplar stimulus frequencies from the benchmark dataset. Figure 9 illustrates the correlation

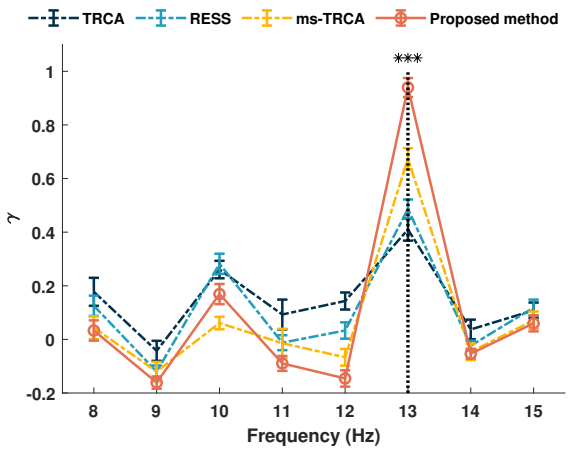
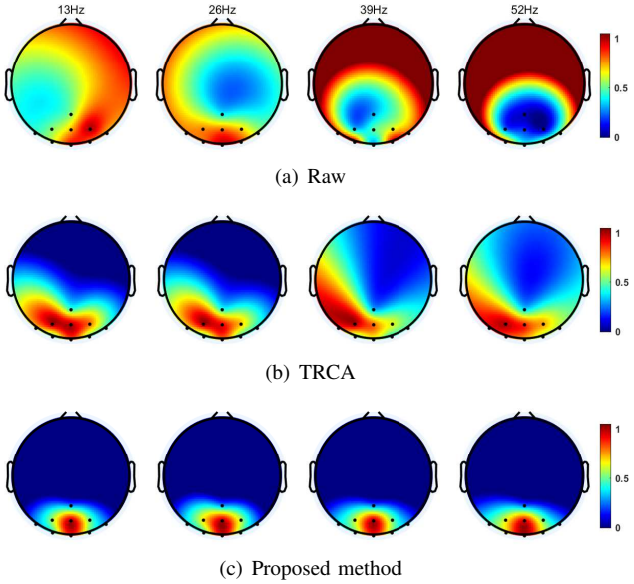


Fig. 9. Averaged correlation coefficients γ for 0.2-s SSVEPs at 13 Hz from the benchmark dataset across subjects. The error bars represent standard deviations. The dotted line represents the target stimulus frequency. The asterisks indicate significant differences between the four methods obtained by paired t-tests ($***p<0.001$).



(a) Raw

(b) TRCA

(c) Proposed method

Fig. 10. Averaged activation pattern across subjects of 13-Hz SSVEPs at the fundamental, second, third, and fourth harmonic frequencies (i.e., 13 Hz, 26 Hz, 39 Hz, and 52 Hz) on the benchmark dataset. The activation patterns were computed with 2 training blocks and then normalized with 1.0-s EEG data.

coefficients γ between 0.2-s test SSVEPs and the 8-15 Hz template signals of the proposed spatial filters and the compared methods TRCA, RESS, and ms-TRCA. As can be seen from the graph, with TRCA and RESS spatial filters, there shows a slight difference in correlation coefficients between certain non-target frequencies and the target frequency, increasing the false rate. Compared to TRCA and RESS, the features obtained by ms-TRCA presented a slightly stronger contrast between the target and non-target frequencies but were inferior to the proposed method. With the proposed method, the correlations γ reached the highest at the target frequency and the lowest at the neighboring frequency, compared with the other three methods. Therefore, with the more discriminative features, the proposed method can effectively improve the target recognition performance.

3) Activation pattern of the proposed method: In addition to the correlation coefficients, the activation pattern is also a profile of the feature space, which can directly present the performance of the spatial filtering methods. Here, the stimulus of 13 Hz was taken as an example. Fig. 10 illustrates the averaged activation pattern of the 13 Hz and its harmonic frequencies on the benchmark dataset for the three conditions, without spatial filter (Fig. 10(a)), with TRCA spatial filter (Fig. 10(b)), and with the proposed spatial filter (Fig. 10(c)). The activation pattern was estimated by averaging the amplitude spectrum of 1.0-s SSVEPs from 2 training blocks and all 35 subjects. It is shown in the figure that with a spatial filter, the activation patterns are more densely in the reflection zone at the fundamental and harmonic frequencies. Moreover, compared with TRCA, the activation pattern of the proposed method is more symmetrically distributed in the occipital cortex by using neighboring stimuli information. These results indicate that the proposed spatial filter can effectively suppress the noise and strengthen the frequency information of the SSVEP response, thus boosting the SSVEP frequency detection performance with limited training blocks.

4) Computational speed of the proposed method: The above experiments were implemented on a Lenovo PC with the Intel(R) Xeon(R) Silver 4116 CPU @ 2.10GHz, 32 GB RAM, and 64-bit Windows 10 OS using Matlab 2022a. In the spatial filter training stage, the computational time of the proposed method for the benchmark dataset and Dataset I were 5.53 s and 1.73 s respectively. For comparison, the training of the ms-TRCA spatial filter costs 6.69 s and 2.45 s computational time respectively. After the spatial filter training, the averaged costs of each time window for performing target recognition of both methods on the two datasets were 0.07 s and 0.03 s, respectively. Even though the two methods have the same target identification speed, our proposed method saves some training time compared to the state-of-art ms-TRCA method. Therefore, we can say that the proposed method can improve the target detection performance with less calibration time.

5) Future directions: In this study, a new method incorporating SSVEP data from the neighboring stimuli is developed for the JFPM-coded SSVEP-based BCIs, which shows its efficiency through offline experiments. There is abundant room for further progress in constructing SSVEP-based BCIs. Apart from the classical JFPM-coded target, the proposed method has the potential to be utilized in the spatially-coded visual BCIs [32] for target identification, which encoded one target with multiple stimuli by the spatial information [33] [34]. In this way, further research should be undertaken to investigate the feasibility of the proposed method in spatially-coded BCIs. Furthermore, the proposed method is a proof-of-concept that verified the improvement in offline experiments. For practical applications, the proposed method can incorporate the dynamic stopping strategy [35] or sliding window strategy [36] to meet the requirements of robust control in real-world scenarios.

IV. Conclusion

This study proposed a novel method incorporating the neighboring-location stimuli data to train a spatial filter for boosting the target detection performance of SSVEP-based BCIs. In this method, the spatial filter was obtained by maximizing the summation of covariances of SSVEPs from the target and its neighboring stimuli. Through extensive comparisons with state-of-art spatial filtering methods, the classification performance of the proposed method was verified offline on two different SSVEP datasets. The experimental results illustrated that the proposed method outperformed the other spatial filters in classification accuracies and ITRs with less training time, which demonstrated the effectiveness of the proposed approach for enhancing the SSVEP frequency identification performance.

Acknowledgment

All authors would like to thank the participants for data acquisition 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, 2000.

[2] M. Cheng *et al.*, "Design and implementation of a brain-computer interface with high transfer rates," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 10, pp. 1181–1186, 2002.

[3] L. Angrisani *et al.*, "A Single-Channel SSVEP-Based Instrument With Off-the-Shelf Components for Trainingless Brain-Computer Interfaces," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 10, pp. 3616–3625, 2019.

[4] X. Yu *et al.*, "A New Framework for Automatic Detection of Motor and Mental Imagery EEG Signals for Robust BCI Systems," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.

[5] G. R. Muller-Putz and G. Pfurtscheller, "Control of an Electrical Prosthesis With an SSVEP-Based BCI," *IEEE Trans. Biomed. Eng.*, no. 1, pp. 361–364, 2008.

[6] E. Yin *et al.*, "A Dynamically Optimized SSVEP Brain–Computer Interface (BCI) Speller," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1447–1456, 2015.

[7] P. Arpaia *et al.*, "Wearable Brain–Computer Interface Instrumentation for Robot-Based Rehabilitation by Augmented Reality," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6362–6371, 2020.

[8] Y. Zhang *et al.*, "Data Analytics in Steady-State Visual Evoked Potential-Based Brain–Computer Interface: A Review," *IEEE Sensors J.*, vol. 21, no. 2, pp. 1124–1138, 2021.

[9] C. M. Wong *et al.*, "Spatial filtering in SSVEP-based BCIs: Unified framework and new improvements," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 11, pp. 3057–3072, 2020.

[10] Z. Lin *et al.*, "Frequency Recognition Based on Canonical Correlation Analysis for SSVEP-Based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2610–2614, 2006.

[11] X. Chen *et al.*, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain-computer interface," *J. Neural Eng.*, vol. 12, no. 4, p. 046008, 2015.

[12] M. R. Islam *et al.*, "Unsupervised frequency-recognition method of SSVEPs using a filter bank implementation of binary subband CCA," *J. Neural Eng.*, vol. 14, no. 2, p. 026007, 2017.

[13] K. Wang *et al.*, "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, 2022.

[14] Y. Zhang *et al.*, "L1-Regularized Multiway Canonical Correlation Analysis for SSVEP-Based BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 6, pp. 887–896, 2013.

[15] Y. Zhang *et al.*, "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis," *Int. J. Neural Syst.*, vol. 24, no. 04, p. 1450013, 2014.

[16] G. Bin *et al.*, "A high-speed BCI based on code modulation VEP," *J. Neural Eng.*, vol. 8, no. 2, p. 025015, 2011.

[17] Y. Wang *et al.*, "Enhancing detection of steady-state visual evoked potentials using individual training data," in *Proc. 36th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2014, pp. 3037–3040.

[18] M. Nakanishi *et al.*, "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, 2018.

[19] 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, 2018.

[20] G. R. Kiran Kumar and M. Ramasubba Reddy, "Designing a Sum of Squared Correlations Framework for Enhancing SSVEP-Based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 10, pp. 2044–2050, 2019.

[21] R. Srinivasan, F. A. Bibi, and P. L. Nunez, "Steady-state visual evoked potentials: distributed local sources and wave-like dynamics are sensitive to flicker frequency," *Brain Topogr.*, vol. 18, no. 3, pp. 167–187, 2006.

[22] J. M. Ales and A. M. Norcia, "Assessing direction-specific adaptation using the steady-state visual evoked potential: Results from EEG source imaging," *J. Vis.*, vol. 9, no. 7, pp. 8–8, 2009.

[23] M. A. Pastor *et al.*, "Human cerebral activation during steady-state visual-evoked responses," *J. Neurosci.*, vol. 23, no. 37, pp. 11 621–11 627, 2003.

[24] S. Fuchs *et al.*, "Attentional bias of competitive interactions in neuronal networks of early visual processing in the human brain," *NeuroImage*, vol. 41, no. 3, pp. 1086–1101, 2008.

[25] M. X. Cohen and R. Gulbinaite, "Rhythmic entrainment source separation: Optimizing analyses of neural responses to rhythmic sensory stimulation," *NeuroImage*, vol. 147, pp. 43–56, 2017.

[26] Y. Zhang *et al.*, "Multi-Objective Optimisation-based High-pass Spatial Filtering for SSVEP-based Brain-Computer Interfaces," *IEEE Trans. Instrum. Meas.*, 2022.

[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, p. 016026, 2020.

[28] Y. Wang *et al.*, "A Benchmark Dataset for SSVEP-Based Brain–Computer Interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1746–1752, 2017.

[29] X. Chen *et al.*, "Hybrid frequency and phase coding for a high-speed SSVEP-based BCI speller," in *Proc. 36th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2014, pp. 3993–3996.

[30] D. H. Brainard and S. Vision, "The psychophysics toolbox," *Spat Vision*, vol. 10, no. 4, pp. 433–436, 1997.

[31] J. Benesty *et al.*, "Pearson correlation coefficient," in *Noise reduction in speech processing*. Springer, 2009, pp. 1–4.

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

[33] A. Maye, D. Zhang, and A. K. Engel, "Utilizing Retinotopic Mapping for a Multi-Target SSVEP BCI With a Single Flicker Frequency," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 7, pp. 1026–1036, 2017.

[34] N. Zhang *et al.*, "Retinotopic and topographic analyses with gaze restriction for steady-state visual evoked potentials," *Scientific Reports*, vol. 9, no. 1, 2019.

[35] J. Jiang *et al.*, "Incorporation of dynamic stopping strategy into the high-speed SSVEP-based BCIs," *J. Neural Eng.*, vol. 15, no. 4, p. 046025, 2018.

[36] P. Gaur *et al.*, "A Sliding Window Common Spatial Pattern for Enhancing Motor Imagery Classification in EEG-BCI," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021.