Latency Aligning Task-Related Component Analysis Using Wave Propagation for Enhancing SSVEP-Based BCIs

Jiayang Huang, Pengfei Yang, *Member, IEEE*, Bang Xiong, Bo Wan, Kejia Su, and Zhi-Qiang Zhang, *Member, IEEE*

Abstract—Due to the high robustness to artifacts, steady-state visual evoked potential (SSVEP) has been widely applied to construct high-speed brain-computer interfaces (BCIs). Thus far, many spatial filtering methods have been proposed to enhance the target identification performance for SSVEP-based BCIs, and task-related component analysis (TRCA) is among the most effective ones. In this paper, we further extend TRCA and propose a new method called Latency Aligning TRCA (LA-TRCA), which aligns visual latencies on channels to obtain accurate phase information from task-related signals. Based on the SSVEP wave propagation theory, SSVEP spreads from posterior occipital areas over the cortex with a fixed phase velocity. Via estimation of the phase velocity using phase shifts of channels, the visual latencies on different channels can be determined for inter-channel alignment. TRCA is then applied to aligned data epochs for target recognition. For the validation purpose, the classification performance comparison between the proposed LA-TRCA and TRCA-based expansions were performed on two different SSVEP datasets. The experimental results illustrated that the proposed LA-TRCA method outperformed the other TRCA-based expansions, which thus demonstrated the effectiveness of the proposed approach for enhancing the SSVEP detection performance.

Index Terms— Brain-computer interfaces (BCIs), latency aligning task-related component analysis (LA-TRCA), phase velocity, steady-state visual evoked potential (SSVEP), wave propagation.

I. INTRODUCTION

BRAIN-COMPUTER interfaces (BCIs) provide disabled people a direct communication interface to external devices via brain activities [1], [2]. EEG-based BCIs [3], such as steady-state visual evoked potential (SSVEP) [4], P300 potential [5], motor imagery (MI) [6], have widespread use as a noninvasive paradigm. SSVEP signals, evoked by a repetitive visual stimulation flickering with a specific frequency, have been widely explored in the past decades due to their high signal-to-noise ratios (SNRs) [7] and information transfer rates (ITRs) [8].

Thus far, many spatial filters have been developed for target identification [9] for SSVEP-based BCIs. The most widely used is canonical correlation analysis (CCA) [10], which seeks a pair of weights to maximize the correlation between SSVEP signals and sine-cosine reference signals. Some extended versions of the standard CCA were also proposed to further improve the performance of SSVEP detection via optimizing the predefined artificial reference. For instance, Zhang *et al.* proposed multiway canonical correlation analysis (Multiway CCA) [11] and L1-regularized multiway CCA (L1-MCCA) [12], which optimized reference signals by maximizing the correlation between multi-way (trial-way, channel-way) data. Zhang *et al.* also proposed multi-set CCA (MsetCCA) [13], which optimized reference signals by extracting common features shared by multiple sets of EEG data. Apart from artificial reference signals, Bin *et al.* proposed individual template CCA (IT-CCA) [14], in which averaged individual data across trials replaced sine-cosine signals as reference signals. Furthermore, Wang *et al.* proposed extended CCA (eCCA) [15], which incorporated reference signals with individual templates to construct spatial filters. Based on the task-related component analysis (TRCA) for functional neuroimaging [16], Nakanishi *et al.* introduced it in SSVEP-based BCIs [17], which extract the task-related component as a template by maximizing inter-trial covariance of individual training data. Similarly, Zhang *et al.* proposed correlated component analysis (CORCA) [18] to construct template signals by maximizing inter-subject covariance. Kiran Kumar and Ramasubba Reddy proposed the sum of squared correlation (SSCOR) method [19], which maximized the sum

of squared correlation of inter-session individual data to construct template signals. All the aforementioned methods implicitly assumed that the visual latency of all channels was at 0.14 s. Based on the volume conduction effects of EEG signals [20], the visual latency of different channels varies due to their location differences, and thus the template signals learned from the aforementioned methods would lose the phase information of SSVEP response.

In order to further improve the accuracy of SSVEP stimuli identification, visual latency or phase information was also considered in some previous methods. For instance, Pan *et al.* proposed phase constrained canonical correlation analysis (p-CCA) [21], in which the SSVEP response phase was estimated based on the latency estimates obtained by the linear fitting of SSVEP phase and stimulus frequency. Zhang *et al.* proposed a Spectrum and Phase Adaptive CCA (SPACCA) [22], in which a library of phase-shifting reference signals was constructed to accommodate the response time lag. Chabuda *et al.* proposed the phase synchronized comb filter [23] to extract the individual template via averaging multiple delayed versions of the signal. Tanka *et al.* proposed the cross-correlation task-related component analysis (xTRCA) method [24], in which the inter-trial latencies were estimated by maximizing trial-reproducibility through iterative optimization. However, according to the wave propagation theory [25], [26], the latency should be different for different channels when SSVEPs propagate from the primary visual cortex to the whole cortex with wave dynamics.

In this paper, a new method called Latency Aligning TRCA (LA-TRCA) is proposed to incorporate the TRCA method with accurate phase information. The visual latency of each channel corresponding to each stimulus is estimated based on the phase velocity of SSVEP wave propagation [27], [28]. The signals from different channels are aligned according to the latency estimates on different channels. The TRCA method is then applied to aligned SSVEP data for target recognition. An SSVEP dataset with 9-channel data stimulated by 12 flickers collected from 10 subjects and the benchmark dataset were used to evaluate the performance of the LA-TRCA method. The comparisons of classification performance between the proposed method and TRCA expansions, i.e., TRCA, CORCA, SSCOR, and xTRCA, were conducted. The comparison results indicated that the LA-TRCA method has an improvement over TRCA-based expansions with different time windows. The promising results demonstrate the efficiency of the LA-TRCA method for improving the target recognition performance of the SSVEP-based BCIs.

The remaining paper 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. METHODS AND MATERIALS

A. Data Acqusition and Preprocessing

This study designed an offline BCI experiment using an SSVEP BCI stimulation interface. 10 subjects who participated in this study (four females and six males, 23 to 27 years
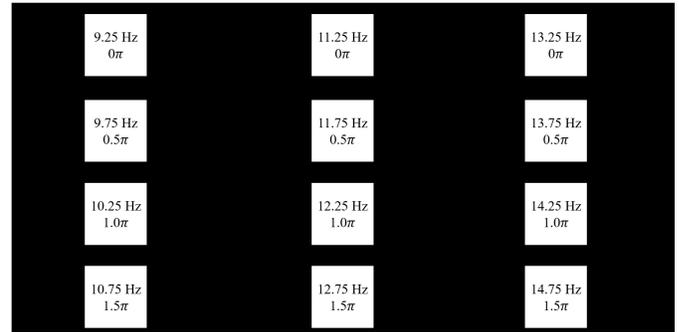

Fig. 1. The interface of the 12-target SSVEP BCI system with frequency and phase values of all stimuli.
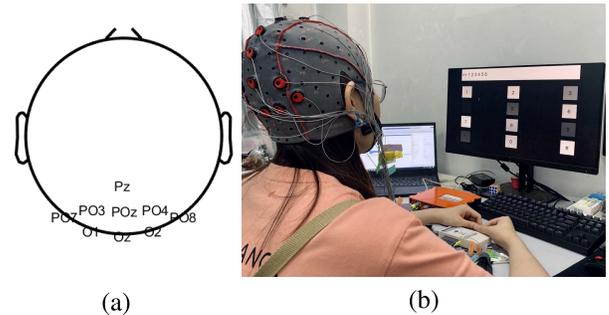

Fig. 2. The 9 electrodes configuration (a) and experimental environment (b) of EEG recording.

old, averaged 25 years old) have normal or corrected-to-normal vision and no brain-related diseases. Only two of them had experience with SSVEP BCIs, while the other eight participants were naive. Each subject was informed of the experimental process and protocols and signed the informed consent before the experiment.

This study designed a 4×3 matrix of visual stimuli, which is coded by joint frequency and phase modulation (JFPM) [29] method. The interface was displayed on a 23.6-inch LCD screen, which has a resolution of 1920- $\times$ 1080-pixel and a refresh rate of 60 Hz. Each flicker was displayed by a square block of 160- $\times$ 160-pixel. The distances between two adjacent stimuli were 100 pixels and 500 pixels vertically and horizontally. 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 (Fig. 1) [30], [31]. This study developed the stimulus program with MATLAB using the Psychophysics Toolbox Version 3 [32].

The acquisition equipment used for EEG recording 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 to collect SSVEP signals (Fig. 2(a)) based on the related studies [17], [30]. The FPz electrode was the ground channel. The reference channel was placed on the right ear of the subject. Electrode impedances were kept below 10 kΩ during data collection. An event trigger indicates the beginning of data collection. It is produced by the stimulus program, which is sent through the parallel port
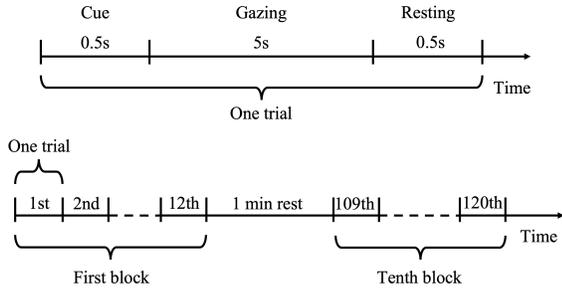
Fig. 3. Experimental procedure for EEG signal acquisition.

to the amplifier and computer simultaneously. According to the event trigger, EEG signals were collected starting from the trigger, and the stimulation began simultaneously. In this way, the EEG signals would be synchronized with the visual stimulation. 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. 2(b)).

For each subject, 10-block BCI experiments were comprised of the whole experiment. In one block experiment, subjects should gaze at one target for 5 seconds. 12 trials were completed corresponding to 12 flickers. To start each trial, a 0.5-s red dot cue was shown on the bottom of the target square. During the 0.5-s cue, each subject was supposed to shift their gaze to the target in time. Then all flickers started to flash simultaneously and lasted 5 seconds. During the stimulation, subjects were asked to avoid eye blinks. Finally, all stimuli stopped flashing for 0.5 s to be ready for the next trial. Each trial lasted 6 seconds. To prevent visual fatigue, there was a one-minute rest between two successive blocks. The 10-block experiments are repetitive experiments. The experiment time chart is shown in Fig. 3. Based on the highest harmonic frequency observed in the collected EEG [33], all data were filtered by a 6-order Butterworth filter from 8 Hz to 75 Hz. A notch filter at 50Hz was utilized to eliminate the power-line noise. After the data acquisition was completed, all the data processing and analysis were then performed.

Moreover, we also used the benchmark datatset [30] consisting of 64-channel EEG data collected from 35 healthy subjects (28 naive and 7 experienced subjects) stimulated by a 40-target BCI speller to verify the efficiency of our method. The stimulation frequencies were arranged from 8 to 15.8 Hz with an interval of 0.2 Hz. The phase ranged from 0π to 1.5π with the interval of 0.5π. The collected data were downsampled to 250 Hz. The EEG data from 9 channels (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, and O2) were used for the evaluation. All data were filtered from 7 Hz to 90 Hz with a 6-order Butterworth filter.

B. TRCA Method

In SSVEP-based BCIs, task-related components are extracted by maximizing the trial reproducibility via TRCA method [17]. Firstly, SSVEPs of the n-th stimulus $\mathcal{X}_n \in \mathbb{R}^{N_c \times N_s \times N_t}$, where N_c is the number of channels, N_s represents the data length of the data epoch in each trial, and N_t

indicates the number of trials for each stimulus, are decomposed into N_b sub-band components $\mathcal{X}_n^{(b)} \in \mathbb{R}^{N_c \times N_s \times N_t}$, $b = 1, 2, \ldots, N_b$ with filter bank analysis [34]. The subbands of filter bank are implemented by zero-phase Chebyshev type I infinite impulse response (IIR) filters. The b-th subband is at frequency range of $[b \times 8$ Hz, 88 Hz]. For each subband, there adds 2 Hz bandwidth to both sides of passband. From the individual calibration data $\mathcal{X}_n^{(b)}$, TRCA obtained spatial filters for the n-th stimulus $\mathbf{w}_n^{(b)} \in \mathbb{R}^{N_c}$. $\mathbf{w}_n^{(b)}$ is calculated as:

$$\mathbf{w}_n^{(b)} = \underset{\mathbf{w}}{\mathrm{argmax}} \frac{\mathbf{w}^\top S_n \mathbf{w}}{\mathbf{w}^\top Q_n \mathbf{w}}. \tag{1}$$

Q_n is defined as the sum of auto-covariance of the n-th stimulus:

$$Q_n = \sum_{\substack{i,j=1 \\ i \neq j}}^{N_t} \mathrm{Cov}\left(\mathcal{X}_n^i, \mathcal{X}_n^j\right). \tag{2}$$

The symmetric matrix S_n represents the sum of cross-covariance of the n-th stimulus, calculated as:

$$S_n = \sum_i^{N_t} \mathrm{Cov}\left(\mathcal{X}_n^i, \mathcal{X}_n^i\right), \tag{3}$$

where i and j represent the indexes of trials, $\mathcal{X}_n^i \in R^{N_c \times N_s}$ is the single trial data of n-th stimulus from training data. The correlation coefficient between the test data $X^{(b)} \in \mathbb{R}^{N_c \times N_s}$ of single trial and averaged training trials for n-th visual stimulus $\overline{\mathcal{X}}_n^{(b)} \in \mathbb{R}^{N_c \times N_s}$ is calculated as:

$$\gamma_n^{(b)} = \rho\left(\left(X^{(b)}\right)^\top \mathbf{w}_n^{(b)}, \left(\overline{\mathcal{X}}_n^{(b)}\right)^\top \mathbf{w}_n^{(b)}\right), \tag{4}$$

where $\rho(s_1, s_2)$ is to calculate the Pearson's correlation coefficient between two signals s_1 and s_2 [35]. By integrating all the correlation coefficients obtained from all subbands, the final features γ_n are calculated as:

$$\gamma_n = \sum_{b=1}^{N_b} c(b) \cdot \left(\gamma_n^{(b)}\right)^2, \tag{5}$$

where the weights for the sub-band component $c(b) = b^{-1.25} + 0.25$ is to maxmize the classification performance [34].

The target frequency $\hat{f}$ with the largest correlation coefficient is defined as:

$$\hat{f} = \underset{n}{\mathrm{argmax}}\ \gamma_n, \quad n = 1, 2, \ldots, N_f. \tag{6}$$

C. LA-TRCA Method

The flowchart of the proposed method is illustrated as Fig. 4. As shown in the figure, given the training data $\mathcal{X}_n \in \mathbb{R}^{N_c \times N_d \times N_t}$ of n-th stimulus, $\mathcal{X}_n$ was filtered by 9-15 Hz narrow-band filtering to estimate accurate instantaneous phase. $\overline{\mathcal{X}}_n = [\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{N_c}] \in \mathbb{R}^{N_c \times N_d}$ is obtained by averaging narrow-band filtered $\mathcal{X}_n$ across trials, where N_d is the data length in a trial of continuous EEG data which start at the stimulus onset corresponding to the event triggers. To begin
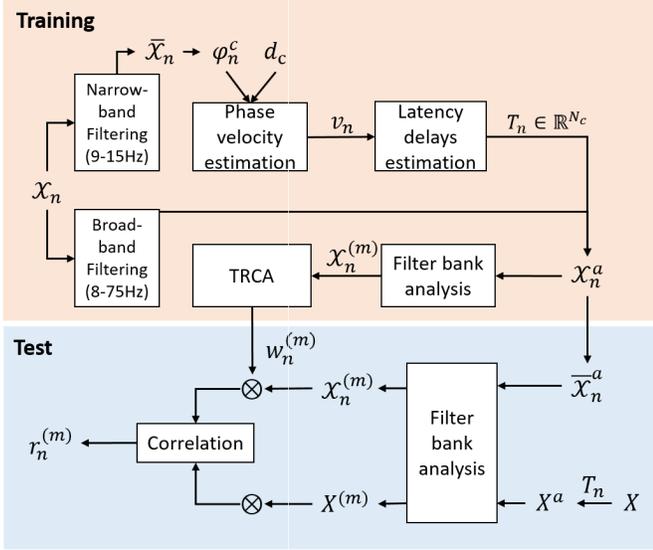
Fig. 4. The flowchart of LA-TRCA method.

Fig. 5. The topographic maps of power spectrum at the fundamental frequencies of all stimuli from one subject as example.

with, the response onsets $\boldsymbol{T}_n = [t_1, t_2, \ldots, t_{N_c}] \in \mathbb{R}^{N_c}$ on channels of n-th visual stimulus are determined for the alignment.

With $\overline{\boldsymbol{\mathcal{X}}}_n$ containing all-channel signals, the instantaneous phase φ_n^c of c-th channel are calculated with Hilbert transform [36]. The Hilbert transform $x_h(t)$ is the convolution of the signal $x(t)$ with the function $h(t) = \dfrac{1}{\pi t}$ as:

$$x_h(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau, \quad t = \frac{1}{fs}, \frac{2}{fs}, \ldots, \frac{N_d}{fs} \quad (7)$$

where τ is the integration variable, fs is the sampling rate, and $x(t)$ is the continuous data $\overline{\boldsymbol{\mathcal{X}}}_n$. Based on Euler's formula, the Hilbert transform is to extract a complex signal from the real part $x(t)$ as:

$$X(t) = x(t) + i \cdot x_h(t) = A_x(t) e^{i\varphi_x(t)}, \quad (8)$$

where $A_x(t)$ represents the instantaneous amplititude of signal $x(t)$. $\varphi_x(t)$ is the instantaneous phase of $x(t)$, which is calculated as:

$$\varphi_x(t) = \arctan \frac{x_h(t)}{x(t)}. \quad (9)$$

To define the phase velocity of SSVEP propagation, the source electrode should be determined. From the topographic maps shown in Fig. 5, the channel with the maximum SSVEP response is considered to be the one with minimum latency delay in this study. For all frequencies, the source channels are in the occipital sites. Channel POz is the first to respond for 11 out of 12 targets [37]. For consistency, henceforth, channel POz chooses to be the source electrode. Consistent with the examplar topology, the source electrodes of all subjects are selected as channel POz.

The phase difference $\Delta\varphi_c$ between c-th channel data $x_e(t)$ and source channel data $x_s(t)$ with phases $\varphi_e(t)$ and $\varphi_s(t)$ respectively, is computed as:

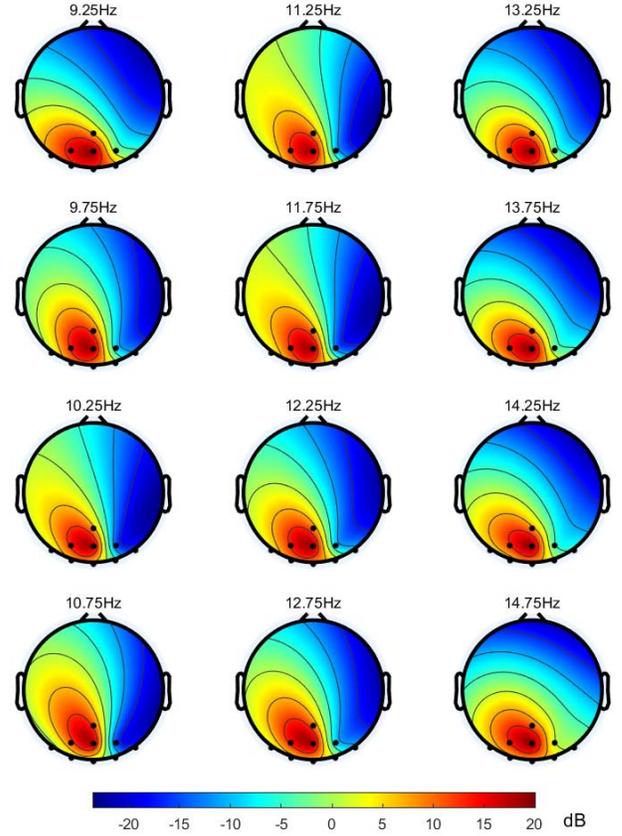$$\Delta\varphi_c = \arg\left(e^{j(\varphi_e(t) - \varphi_s(t))}\right), \quad (10)$$

where arg is the same operation as Eq. 9, thus getting the phase differences between all electrodes and source electrode $\boldsymbol{\phi}_n = [\Delta\varphi_1, \Delta\varphi_2, \ldots, \Delta\varphi_{N_c}] \in \mathbb{R}^{N_c}$. Given the electrode location (dx, dy, dz) in Cartesian coordinates, the distance d_c between c-th electrode (dx_e, dy_e, dz_e) and source electrode (dx_s, dy_s, dz_s) is estimated as:

$$\Delta d_c = \sqrt{(dx_e - dx_s)^2 + (dy_e - dy_s)^2 + (dz_e - dz_s)^2}, \quad (11)$$

then the distances between all electrodes and source electrode $\boldsymbol{d} = [\Delta d_1, \Delta d_2, \ldots, \Delta d_{N_c}] \in \mathbb{R}^{N_c}$ are obtained. Phase velocity v_n of n-th stimulus is the phase speed of a wave propagation and is calculated as:

$$v_n = \frac{1}{N_c} \sum_{c=1}^{N_c} 2\pi f (\Delta d_c / \Delta\varphi_c), \quad (12)$$

where f is the stimulus frequency.

The latency t_c^n on c-th channel at n-th stimulus then can be estimated with phase velocity v_n and the distance Δd_c as:

$$t_c^n = \frac{\Delta d_c}{v_n}, \quad (13)$$

the latencies of all channels $\boldsymbol{T}_n = [t_1^n, t_2^n, \ldots, t_{N_c}^n] \in \mathbb{R}^{N_c}$ is obtained.

With the latencies $\boldsymbol{T}_n$ for aligning data on multiple channels, the training data $\boldsymbol{\mathcal{X}}_n$ are extracted into $\boldsymbol{\mathcal{X}}_n^a \in \mathbb{R}^{N_c \times N_s \times N_t}$.
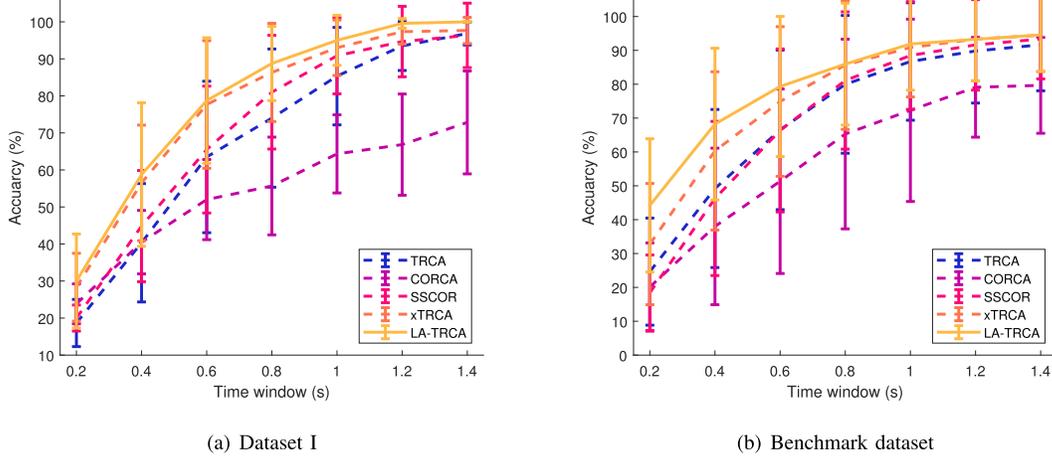
(a) Dataset I

(b) Benchmark dataset

Fig. 6. Averaged accuracy across subjects of various methods using different time windows on (a) self-collected dataset I and (b) benchmark dataset. The time window increase from 0.2 s to 1.4 s with a step of 0.2 s. The error bar represents the standard deviation.

With filter bank analysis, the aligned training data $\mathcal{X}_n^{\mathrm{a}}$ is decomposed into N_b sub-band components as $\mathcal{X}_n^{(b)}$, where $b = 1, 2, \ldots, N_b$. And the spatial filters $\boldsymbol{w}_n^{(b)} \in \mathbb{R}^{N_c}$ for the n-th stimulus are obtained through TRCA from $\mathcal{X}_n^{(b)}$. The calculation of $\boldsymbol{w}_n^{(b)}$ is implemented as Eq. 1:

$$\boldsymbol{w}_n^{(b)} = \underset{\boldsymbol{w}}{\arg\max} \frac{\boldsymbol{w}^\top \boldsymbol{S}_n^{\mathrm{a}} \boldsymbol{w}}{\boldsymbol{w}^\top \boldsymbol{Q}_n^{\mathrm{a}} \boldsymbol{w}}, \tag{14}$$

where $\boldsymbol{S}_n^{\mathrm{a}}$ and $\boldsymbol{Q}_n^{\mathrm{a}}$ are obtained as Eq. 2 and 3 with $\mathcal{X}_n^{\mathrm{a}}$.

The single-trial test data $\boldsymbol{X}^{\mathrm{a}} \in \mathbb{R}^{N_c \times N_s}$ is extracted from $\boldsymbol{X}$ with $\boldsymbol{T}_n$. Finally, with $\boldsymbol{w}_n^{(b)}$, the correlation coefficient between $\boldsymbol{X}^{\mathrm{a}}$ and averaged training data across trials for n-th visual stimulus $\overline{\mathcal{X}}_n^{\mathrm{a}} \in \mathbb{R}^{N_c \times N_s}$ is calculated as:

$$\gamma_n^{(b)} = \rho \left(\left(\boldsymbol{X}^{\mathrm{a}} \right)^\top \boldsymbol{w}_n^{(b)}, \left(\overline{\mathcal{X}}_n^{\mathrm{a}} \right)^\top \boldsymbol{w}_n^{(b)} \right). \tag{15}$$

The final features γ_n are obtained the same as Eq. 6. Finally, the target stimulus $\hat{f}$ with the largest correlation γ_n is obtained as Eq. 7.

D. Performance Evaluation

This study evaluated the target detection performance of the LA-TRCA method with the classification accuracy and ITR estimates. The accuracy of classification is defined as the rate of the number of correct predictions out of all predictions. ITR is the amount of communication information per minute, defined as:

$$\mathrm{ITR} = \frac{60}{T} \times \left[\log_2 N_f + P \times \log_2 P + (1 - P) \times \log_2 \left(\frac{1 - P}{N_f - 1} \right) \right], \tag{16}$$

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.

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section, we first evaluated the performance of the proposed LA-TRCA method on the self-collected SSVEP dataset and the 40-target benchmark dataset [30]. Extensive comparisons of classification accuracy and ITR were implemented between the proposed method and many state-of-the-art SSVEP target recognition methods. The influences of different parameters such as subject's experience, the number of training blocks, and the distances between adjacent stimuli were also reported. Based on the experimental results, the discussions of LA-TRCA were presented in the last subsection.

A. Target Detection Performance

The stimuli frequency classification performance comparison was conducted between the proposed LA-TRCA method and TRCA, CORCA, SSCOR, and xTRCA. The classification accuracy and ITRs were calculated by leave-one-out cross-validation to evaluate the recognition performance of these algorithms. Figure 6 illustrates the averaged classification accuracy across all subjects in the two datasets with different time windows. As we can see from the figure, the proposed LA-TRCA method can in general achieve the highest accuracy with different data lengths at all time windows with both datasets compared to the state-of-art methods. To reveal the significant difference in the classification accuracy between LA-TRCA and other methods, we further carried out a pairwise analysis between the proposed method and each comparison method on these two datasets. The multiple comparison results in terms of SSVEP recognition accuracy with different time windows are presented in Table I. It is obvious that with a time window smaller than 0.8 s, LA-TRCA outperformed TRCA, CORCA, and SSCOR by a significant margin, which proved the effectiveness of the latency-aligning operation for extracting phase information. With a time window exceeding 1.0 s, except CORCA, none of these differences were statistically significant, since accuracies for different methods gradually increased to 100%.

TABLE I

THE SIGNIFICANT DIFFERENCE ON CLASSIFICATION ACCURACY BETWEEN LA-TRCA AND OTHER STATE-OF-ART METHODS OBTAINED BY ONE-WAY REPEATED MEASURES ANOVAs

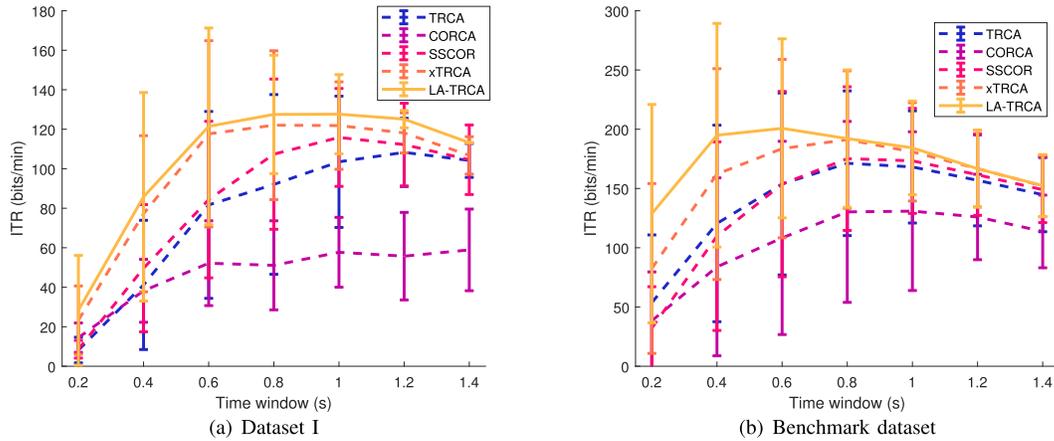| Methods | Time windows | | | | | | | | | | | | | |
| | Dataset I | | | | | | | Benchmark Dataset | | | | | | |
	0.2 s	0.4 s	0.6 s	0.8 s	1.0 s	1.2 s	1.4 s	0.2 s	0.4 s	0.6 s	0.8 s	1.0 s	1.2 s	1.4 s
TRCA vs. LA-TRCA*	0.0213	0.0324	0.0861	0.0410	0.0535	0.0108	0.0045	<0.0001	0.0009	0.0184	0.1966	0.1693	0.3015	0.3257
CORCA vs. LA-TRCA	0.1745	0.0140	0.0005	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
SSCOR vs. LA-TRCA	0.0274	0.0896	0.0991	0.1981	0.2981	0.1231	0.1996	<0.0001	0.0001	0.0183	0.2941	0.3399	0.5914	0.6574
xTRCA vs. LA-TRCA	0.7403	0.7782	0.8889	0.6512	0.5382	0.0637	0.0511	0.0133	0.1520	0.3899	0.9508	0.7573	0.9903	0.9816

Fig. 7. Averaged ITRs across subjects of various methods using different time windows on (a) self-collected dataset I and (b) benchmark dataset. The time window increased from 0.2 s to 1.4 s with a step of 0.2 s. The error bar represents the standard deviation.

In addition to the classification performance analysis, we further explored the ITRs across all subjects in the two datasets with different time windows, as shown in the Fig. 7. The tendency of ITRs is consistent with accuracies that the proposed LA-TRCA method can in general achieve the highest ITRs among different methods. With the collected 12-target dataset, LA-TRCA obtained the highest ITRs at 0.8-s time window as 127.63±29.96 bits/min. While for the benchmark dataset with 40 targets, LA-TRCA obtained the highest ITRs with 0.6-s data as 200.77±75.55 bits/min.

B. The Influence of Parameters on Performance

In order to further evaluate the performance of the proposed approach, we explored the impact of the subject's experience, the number of training blocks, and the distances between neighboring targets on the SSVEP recognition accuracies.

1) The Subject's Experience: The two datasets consist of data collected from naive and experienced subjects, and it is unclear how their experience will affect the SSVEP classification accuracy. Therefore, we performed a study to investigate the effect of subjects' experience on target detection performance. Fig. 8 illustrates the classification accuracy distributions of two kinds of subjects via violin plots. The subjects are divided into naive or experienced groups: 8 vs. 2 for the self-collected dataset and 28 vs. 7 for the benchmark dataset. As we can see from the figure, although the accuracy distributions of experienced subjects were slightly more concentrated to a higher median value, for both groups of subjects, all five methods obtained satisfactory results.

TABLE II

THE VERTICAL AND HORIZONTAL SPACINGS BETWEEN NEIGHBORING STIMULI OF DIFFERENT STIMULATION SETTINGS

Stimulation settings (rows×columns)	4×3*	3×4	2×6
Vertical spacing (pixels)	100	150	300
Horizontal spacing (pixels)	500	280	70

2) The Number of Training Blocks: In the LA-TRCA method, the template signals were constructed with individual training data, the target identification performance also depends on the number of training blocks. Figure 9 presents the averaged classification accuracies across subjects obtained with different numbers of training blocks (N_t). As shown in both figures, the averaged accuracies of all five methods gradually increased with the number of training blocks. And LA-TRCA consistently outperformed other TRCA-based methods on both datasets with sufficient training blocks ($N_t \geq 3$). As a result, the best target detection performance of LA-TRCA can be achieved by using 9-block data for training with the collected dataset, while 5 blocks with the benchmark dataset.

3) The Distance Between Neighboring Stimuli: In the stimulation interface, the different distances between adjacent targets may have different interference on target identification. We have conducted the 3 different stimulation interface settings as 4 × 3, 3 × 4, and 2 × 6 matrices. The details of the stimulation interface settings were listed in Table II. The identification accuracies of three stimulation interfaces at the time window of 1.5 s have been given in Fig. 10. As we can see from the figure, it is obvious that the 4 × 3 and 3 × 4 matrices

(a) Dataset I

(b) Benchmark dataset
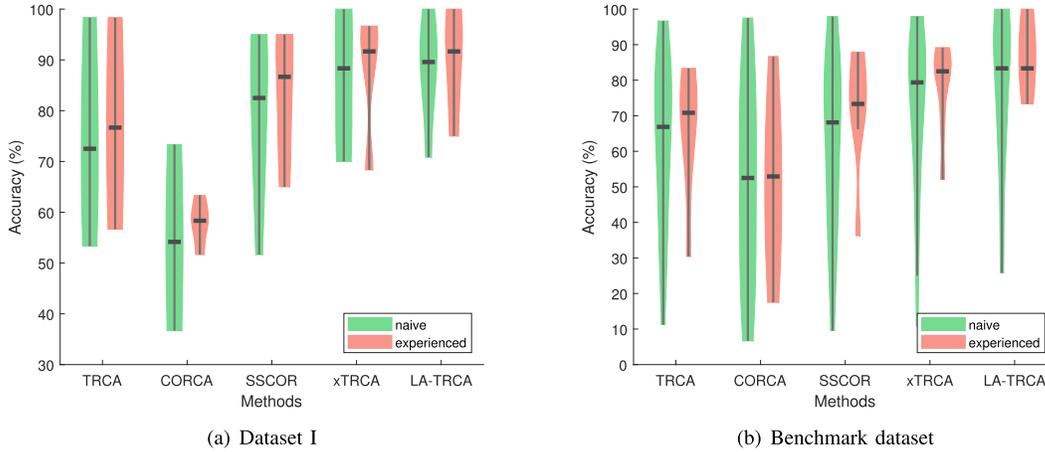
Fig. 8. Classification accuracy distributions of two groups of subjects with (a) the self-collected dataset I at 0.8 s and (b) the benchmark dataset at 0.6 s. The top and bottom sides of the violin represent the maximum and minimum. The vertical and horizontal lines represent the range of accuracies without outliners and the median value respectively. The width of the violin refers to the probability density estimate.

(a) Dataset I
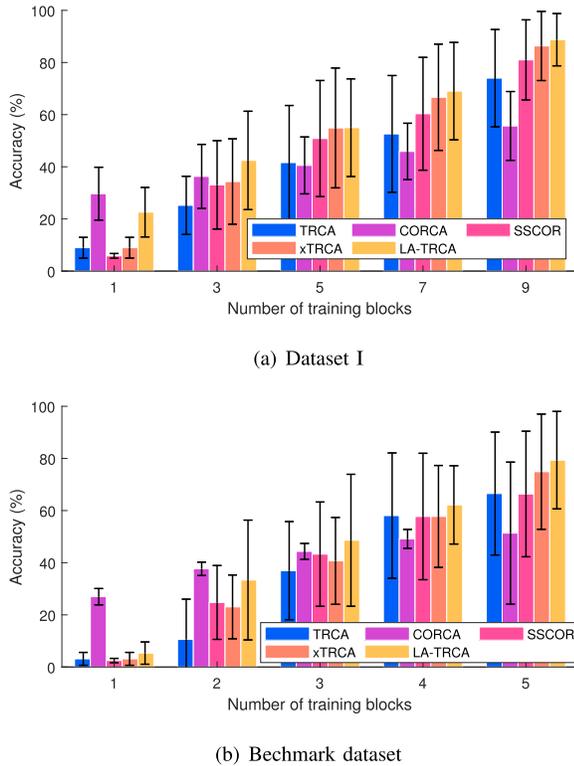
(b) Bechmark dataset

Fig. 9. Averaged classification accuracies across subjects of the self-collected dataset I (a) and the benchmark dataset (b) with different numbers of training blocks (N_t).Here the time windows were set to 0.8 s and 0.6 s respectively for the two datasets, where highest ITRs were achieved. The vertical error bars represent standard deviations.

obtained the similar result on accuracies with all methods, but the accuraies of the 2×6 matrix were significantly lower than the other two interface settings by approximately 20%. This is probably due to insufficient horizontal distance between neighboring stimuli, resulting weak SSVEP amplitude of the target stimulus [38]. Therefore, in our experiment settings, the minimal distance between adjacent stimulus flickers was set to at least 100 pixels to avoid this.
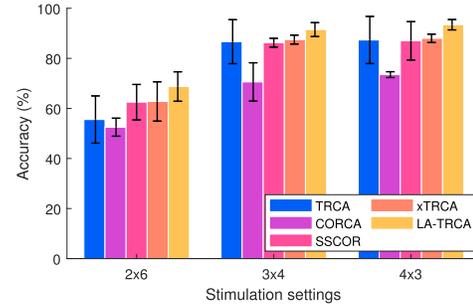
Fig. 10. The comparison of averaged classification accuracies using 1.5-s SSVEPs with different stimulation settings as 4×3, 3×4, and 2×6 matrix.

C. Discussions

1) Correlation Coefficient Discriminability: In this study, we aligned inter-channel latencies using the SSVEP wave propagation theory for accurate phase information. The experimental results have demonstrated that LA-TRCA could effectively extract task-related signals with accurate phase information. With accurate phase information, a more discriminative correlation coefficient would be obtained for target identification. To intuitively illustrate the contribution on phase information, we randomly choose 11.25 Hz as the exemplar stimuli frequency. Figure 11 shows the correlations coefficients between 0.5-s SSVEPs at 11.25 Hz and template signals corresponding to all stimuli. As the graph shows, with TRCA and xTRCA, certain non-target stimuli obtained similar correlation coefficients as the target frequency, increasing the false rate. By contrast, LA-TRCA with accurate phase information achieved a significant difference in feature values between target and non-target stimuli, which leads to correct target recognition.

2) Computational Cost of the LA-TRCA Method: The above experimental results were implemented using Matlab 2020b on a Lenovo PC with the Intel(R) Xeon(R) Silver 4116 CPU @ 2.10GHz, 32 GB RAM, and 64-bit Windows 10 OS.
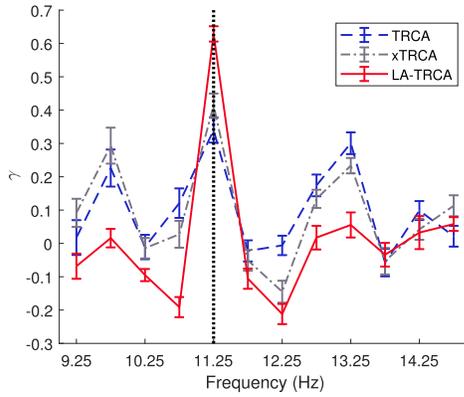
Fig. 11. Averaged feature value as correlation coefficients γ for 0.5-s SSVEPs at 11.25 Hz from the collected dataset across subjects. The error bars represent standard deviations. The dotted line represents the target stimulus frequency.

For the spatial filter training, it took up to 0.84 s for the self-collected dataset and 2.32 s for the benchmark dataset. Once the latency is determined and the spatial filter is trained, the averaged recognition time per time window for performing the proposed method on the two datasets were 0.07 s and 0.43 s, respectively. We would like to point out that although LA-TRCA has a more complicated training process than the TRCA method for latency aligning, and the training time is a bit long, it won't affect the computational speed of the target detection. We could always determine the latency and train the spatial filter offline before we move on to real-time SSVEP recognition.

3) Bottlenecks and Future Work: Despite promising results that have been achieved by the proposed LA-TRCA method, there is still room for improvements. First, the target detection performance can be further enhanced with time windows smaller than 0.5 s. The spatial filter can suppress certain noise by strengthening the SSVEP task-related components, but the SSVEP signal still includes a significant amount of spontaneous signals. In the future, different time filters [39] will be explored to improve the target detection performance for short time windows by further suppressing the spontaneous signals. In addition, our proposed method would require latency estimation and spatial filter training for each subject. It is time-consuming and also requires a large amount of training data. In the future, transfer learning approaches [40], [41] will be used to estimate the latency and spatial filter using less training data.

IV. CONCLUSION

In this study, a new method LA-TRCA was proposed to enhance the target detection performance of SSVEP-based BCIs. In LA-TRCA, accurate phase information can be obtained from task-related signals by inter-channel aligning and spatial filter training. The experimental results based on two different datasets support the improvement of LA-TRCA on target detection performance compared to state-of-art methods, which indicates the promising potentials for applications in real-life scenarios.

REFERENCES

[1] J. 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, Feb. 2000.

[2] Y. Yu, Y. Liu, E. Yin, J. Jiang, Z. Zhou, and D. Hu, "An asynchronous hybrid spelling approach based on EEG–EOG signals for Chinese character input," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 6, pp. 1292–1302, Jun. 2019.

[3] F. Lotte *et al.*, "A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update," *J. Neural Eng.*, vol. 15, no. 3, Jun. 2018, Art. no. 031005.

[4] E. Yin, Z. Zhou, J. Jiang, Y. Yu, and D. Hu, "A dynamically optimized SSVEP brain–computer interface (BCI) speller," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1447–1456, Jun. 2015.

[5] X. Xiao *et al.*, "Enhancement for P300-speller classification using multi-window discriminative canonical pattern matching," *J. Neural Eng.*, vol. 18, no. 4, Aug. 2021, Art. no. 046079.

[6] J. Jin, R. Xiao, I. Daly, Y. Miao, X. Wang, and A. Cichocki, "Internal feature selection method of CSP based on L1-norm and Dempster–Shafer theory," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 11, pp. 4814–4825, Nov. 2020.

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

[8] 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.

[9] 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, Aug. 2021.

[10] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2610–2614, Dec. 2006.

[11] Y. Zhang *et al.*, "Multiway canonical correlation analysis for frequency components recognition in SSVEP-based BCIs," in *Proc. Int. Conf. Neural Inf. Process.*, 2011, pp. 287–295.

[12] Y. Zhang, G. Zhou, J. Jin, M. Wang, X. Wang, and A. Cichocki, "L1-regularized multiway canonical correlation analysis for SSVEP-based BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 6, pp. 887–896, Nov. 2013.

[13] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis," *Int. J. Neural Syst.*, vol. 24, no. 4, pp. 1450013-1–1450013-14, 2014.

[14] G. Bin, X. Gao, Y. Wang, Y. Li, B. Hong, and S. Gao, "A high-speed BCI based on code modulation VEP," *J. Neural Eng.*, vol. 8, no. 2, Apr. 2011, Art. no. 025015.

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

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

[17] M. Nakanishi, Y. Wang, X. Chen, Y. Wang, X. Gao, and T.-P. Jung, "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 1, pp. 104–112, Jan. 2018.

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

[19] G. R. K. Kumar and M. R. 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, Oct. 2019.

[20] A. Omidvarnia, G. Azemi, B. Boashash, J. M. O'Toole, P. B. Colditz, and S. Vanhatalo, "Measuring time-varying information flow in scalp EEG signals: Orthogonalized partial directed coherence," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 3, pp. 680–693, Mar. 2014.

[21] 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.

[22] Z. Zhang, C. Wang, K. K. Ang, A. A. P. Wai, and C. G. Nanyang, "Spectrum and phase adaptive CCA for SSVEP-based brain computer interface," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 311–314.

[23] A. Chabuda, P. Durka, and J. Zygierewicz, "High frequency SSVEP-BCI with hardware stimuli control and phase-synchronized comb filter," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 344–352, Feb. 2018.

[24] H. Tanaka and M. Miyakoshi, "Cross-correlation task-related component analysis (xTRCA) for enhancing evoked and induced responses of event-related potentials," *NeuroImage*, vol. 197, pp. 177–190, Aug. 2019.

[25] F. Di Russo and D. Spinelli, "Electrophysiological evidence for an early attentional mechanism in visual processing in humans," *Vis. Res.*, vol. 39, no. 18, pp. 2975–2985, Sep. 1999.

[26] H. R. Wilson, R. Blake, and S.-H. Lee, "Dynamics of travelling waves in visual perception," *Nature*, vol. 412, no. 6850, pp. 907–910, Aug. 2001.

[27] G. R. Burkitt, R. B. Silberstein, P. J. Cadusch, and A. W. Wood, "Steady-state visual evoked potentials and travelling waves," *Clin. Neurophysiol.*, vol. 111, no. 2, pp. 246–258, Feb. 2000.

[28] P. L. Nunez and R. Srinivasan, "Neocortical dynamics due to axon propagation delays in cortico-cortical fibers: EEG traveling and standing waves with implications for top-down influences on local networks and white matter disease," *Brain Res.*, vol. 1542, pp. 138–166, Jan. 2014.

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

[30] 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, Nov. 2016.

[31] 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.

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

[33] G. R. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller, "Steady-state visual evoked potential (SSVEP)-based communication: Impact of harmonic frequency components," *J. Neural Eng.*, vol. 2, no. 4, p. 123, 2005.

[34] 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.

[35] J. Benesty *et al.*, "Pearson correlation coefficient," in *Noise Reduction in Speech Processing*. Berlin, Germany: Springer, 2009, pp. 1–4.

[36] M. Le Van Quyen *et al.*, "Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony," *J. Neurosci. Methods*, vol. 111, no. 2, pp. 83–98, Sep. 2001.

[37] N. Zhang *et al.*, "Retinotopic and topographic analyses with gaze restriction for steady-state visual evoked potentials," *Sci. Rep.*, vol. 9, no. 1, pp. 1–10, Dec. 2019.

[38] S. Fuchs, S. K. Andersen, T. Gruber, and M. M. Müller, "Attentional bias of competitive interactions in neuronal networks of early visual processing in the human brain," *NeuroImage*, vol. 41, no. 3, pp. 1086–1101, Jul. 2008.

[39] J. Jin, Z. Wang, R. Xu, C. Liu, X. Wang, and A. Cichocki, "Robust similarity measurement based on a novel time filter for SSVEPs detection," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Oct. 14, 2021, doi: 10.1109/TNNLS.2021.3118468.

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

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