Enhancing Detection of SSVEPs for a High-Speed Brain Speller Using Task-Related Component Analysis

Masaki Nakanishi, *Member, IEEE*, Yijun Wang, *Member, IEEE*, Xiaogang Chen, *Member, IEEE*, Yu-Te Wang, *Member, IEEE*, Xiaorong Gao, *Member, IEEE*, and Tzyy-Ping Jung, *Fellow, IEEE*

Abstract*—*Objective: **This study proposes and evaluates a novel data-driven spatial filtering approach for enhancing steady-state visual evoked potentials (SSVEPs) detection toward a high-speed brain-computer interface (BCI) speller. *Methods:* Task-related component analysis (TRCA), which can enhance reproducibility of SSVEPs across multiple trials, was employed to improve the signal-to-noise ratio (SNR) of SSVEP signals by removing background electroencephalographic (EEG) activities. An ensemble method was further developed to integrate TRCA filters corresponding to multiple stimulation frequencies. This study conducted a comparison of BCI performance between the proposed TRCA-based method and an extended canonical correlation analysis (CCA)-based method using a 40-class SSVEP dataset recorded from 12 subjects. An online BCI speller was further implemented using a cue-guided target selection task with 20 subjects and a free-spelling task with 10 of the subjects. *Results:* The offline comparison results indicate that the proposed TRCA-based approach can significantly improve the classification accuracy compared with the extended CCA-based method. Furthermore, the online BCI speller achieved averaged information transfer rates (ITRs) of 325.33 $\pm$ 38.17 bits/min with the cue-guided task and 198.67 $\pm$ 50.48 bits/min with the free-spelling task. *Conclusion:* This study validated the efficiency of the proposed TRCA-based method in implementing a high-speed SSVEP-based BCI. *Significance:* The high-speed SSVEP-based BCIs using the TRCA method have great potential for various applications in communication and control.**

***Index Terms*—Brain-computer interfaces (BCI), electroencephalography (EEG), steady-state visual evoked potentials (SSVEP), task-related component analysis (TRCA).**

Manuscript received November 21, 2016; revised February 25, 2017; accepted April 6, 2017. Date of publication April 19, 2016; date of current version December 20, 2017. This work was supported in part by Army Research Laboratory (W911NF-10-2-0022) and National Institute of Healths under Grant 1R21EY025056-01. The work of Y. Wang was also supported by the Recruitment Program for Young Professionals and National Natural Science Foundation of China under Grant 61671424. The work of X. Chen was supported by Young Talents Lift Project of Chinese Association of Science and Technology, and National Natural Science Foundation of China under Grant 61603416. *(Corresponding author: Yijun Wang.)*

M. Nakanishi, Y.-T. Wang, and T.-P. Jung are with the Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California.

Y. Wang is with the State Key Laboratory on Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China (e-mail: wangyj@semi.ac.cn).

X. Chen is with the Institute of Biomedical Engineering, Chinese Academy of Medical Sciences and Peking Union Medical College.

X. Gao is with the Department of Biomedical Engineering, Tsinghua University.

Digital Object Identifier 10.1109/TBME.2017.2694818

I. Introduction

BRAIN-COMPUTER interfaces (BCIs) based on steady-state visual evoked potentials (SSVEPs) have been attracting much attention due to high communication rates [1]–[3]. In a recent study, a 40-target SSVEP-based BCI achieved an information transfer rate (ITR) of 267 bits/min, to our knowledge, the highest ITR reported in BCI research to date [2]. To improve the performance of SSVEP-based BCIs, researchers are working on two main directions: (1) increasing the number of classes (i.e., visual stimuli), (2) improving target identification algorithms [4]. Although the number of frequencies that can be presented on a computer monitor is limited by its refresh rate, various stimulus presentation approaches inspired by multiple access (MA) technologies used in telecommunication have succeeded in rendering a large number of visual flickers on a computer monitor [2], [5]–[8]. For example, the frequency shift keying (FSK) method has been applied to increase the number of visual stimuli with different properties using the limited number of available frequencies [5]. A hybrid paradigm which combines SSVEP with other electroencephalogram (EEG) features (e.g., P300) is another way to increase the number of selections [6]. In several other studies, frequency-approximation approaches make it possible to render a large number of flickers with flexible frequencies on a computer monitor [7]–[9]. More recently, the efficacy of a joint frequency-phase modulation (JFPM) method has been demonstrated in our previous studies [2], [4], [10].

Target identification methods also play an important role in enhancing the performance of SSVEP-based BCIs [4]. The methods based on power spectrum density analysis (PSDA) such as discrete Fourier transform (DFT) were widely used to detect target frequency of SSVEPs from single-channel EEGs [11], [12]. In recent studies, temporal features of SSVEPs have been employed to develop more efficient methods than PSDA [13]. For example, canonical correlation analysis (CCA) between SSVEPs and sine-cosine reference signals has been used to improve the classification accuracy due to its ability to enhance the signal-to-noise ratio (SNR) of SSVEPs [13], [14]. Although the robustness of CCA-based method in detecting

SSVEPs has been proved, its performance is still affected by the interference from the spontaneous EEG activities. To reduce the misclassification rate caused by the spontaneous EEG signals, individual calibration data have been incorporated in CCA-based SSVEP detection [4], [15]–[18]. Zhang et al. proposed the multi-way CCA (MwayCCA) and the L1-reguralized MwayCCA (L1-MCCA) methods to optimize sine-cosine reference signals based on multiple standard CCA processes with the calibration data [15], [16]. The multi-set CCA (MsetCCA) method, in which the reference signals are optimized from common features in multiple calibration trials without any artificial signals (i.e., sine-cosine signals), showed better performance than the MwayCCA and the L1-MCCA [17]. In our recent studies, we proposed an extended CCA-based method to combine CCA-based spatial filtering and template matching to detect SSVEPs [4], [18]. In this method, after the CCA process, a separate procedure of correlation analysis (between testing data and individual templates) was used to enable the discrimination between different phases. A recent comparison study showed the combination method achieved the highest performance in these calibration data based methods [19].

Scalp EEG data recorded at a single electrode is generally considered as a simple sum of multiple time series from different cortical sources [20]. In other words, the cortical source signals can be reconstructed from a weighted linear summation of scalp EEG signals from multiple electrodes. Based on this idea, several spatial filtering methods have been applied to extract and investigate task-related source activities from scalp recordings [21]–[26]. For example, independent component analysis (ICA) is widely used for separating artifacts from EEG data. The ICA method is based only on statistical independence between source signals [21]–[24]. In several recent studies, Tanaka et al. developed a spatial filtering approach called task-related component analysis (TRCA), in which weight coefficients are optimized to maximize inter-trial covariance of time-locked near-infrared spectroscopy (NIRS) data [25], [26]. By maximizing the reproducibility of time-locked activities across trials, the TRCA method showed good performance in extracting task-related components. SSVEPs are known as time-locked photic driving responses to repetitive visual stimuli [27], [28]. A recent computer simulation has shown that the intra-subject variability of visual latencies was less than 2 ms [2]. Considering the stability of the latency in SSVEPs, the TRCA-based spatial filtering approach has great potential to improve the SNR of SSVEPs. To the best of our knowledge, the TRCA-based analysis has never been used in any EEG study including SSVEP-based BCIs.

This study proposed a TRCA-based target identification algorithm to improve the performance of an SSVEP-based BCI. The performance was evaluated using a 40-target SSVEP dataset recorded from 12 subjects [2]. In the performance evaluation, the extended CCA-based method [4] was used to compare the performance with the proposed method. After optimizing parameters (e.g., data length, the number of electrodes, the number of training trials, the number of sub-bands etc.), an online BCI experiment was conducted using the proposed method. The goal of this study is to demonstrate the efficacy of the TRCA-based method in fast detection of SSVEPs for a high-speed BCI speller.

Fig. 1. Stimulus design of the 40-target BCI system. Frequency and phase values specified for each target.

II. Methods and Materials

A. Experimental Environment

1) Subjects: This study designed an offline experiment and an online experiment separately. Twelve (5 females, mean age 25 years) and twenty (10 females, mean age 25 years) healthy adults with normal or corrected-to-normal vision participated in the offline and online experiments, respectively. The online experiment included a cue-guided target selection task and a free spelling task. All 20 subjects participated in the cue-guided task and 10 of them also participated in the free spelling task. Each participant was asked to read and sign an informed consent form approved by the Research Ethics Committee of Chinese Academy of Medical Sciences before the experiments.

2) Stimulus Design: This study used the sampled sinusoidal stimulation method [9] to present visual flickers coded by the JFPM method on a 23.6-inch liquid-crystal display (LCD) monitor. The resolution and refresh rate of the monitor were 1,920 × 1,080 pixels and 60 Hz, respectively. The user interface is a 5 × 8 matrix of visual stimuli containing 40 characters (26 English alphabet letters, 10 digits, and 4 other symbols). Specifically, 40 stimuli are tagged by linearly increasing frequencies and phases, of which the increments are both proportional to stimulus index (Fig. 1). The frequency range was selected from 8 Hz to 15.8 Hz with an interval of 0.2 Hz. The phase values started from 0 and the phase interval was 0.35 π. Each stimulus was rendered within a 140- × 140-pixel square. The characters were presented within a 32- × 32-pixel square at the center of each stimulus. The vertical and horizontal distances between two neighboring stimuli were 50 pixels. The stimulus program was developed under MATLAB (MathWorks, Inc.) using the Psychphysics Toolbox Version 3 [29].

3) EEG Recording: EEG data were recorded by nine electrodes over parietal and occipital areas (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, O2) using a Synamp2 system (Neuroscan, Inc.) at a sampling rate of 1,000 Hz. The reference electrode was placed at the vertex (Cz). Event triggers generated by the stimulus program were recorded on an event channel of the amplifier synchronized to the EEG data. In the online experiment, EEG data and trigger signals were recorded and analyzed by the online data analysis program in real time. The online data analysis program was developed under MATLAB.

B. Experimental Design

1) Offline BCI Experiment: In the offline experiment, the subjects performed a simulated online experiment to record data for offline analysis. The experiment consisted of 12 blocks. In each block, subjects were asked to gaze at one of the visual stimuli for 0.5 s, and completed 40 trials corresponding to all 40 stimuli. Each trial started with a visual cue (a red square) generated by the stimulus program indicating a target stimulus. The cue appeared for 0.5 s on the screen. Each subject was instructed to shift their gaze to the target as soon as possible within the cue duration. After that, all stimuli started to flicker simultaneously for 0.5 s on the monitor. To reduce eye movement artifacts, subjects were asked to avoid eye blinks during the stimulation period. To avoid visual fatigue, there was a rest for several minutes between two consecutive blocks.

Data epochs comprising nine-channel SSVEPs were first extracted according to the event triggers. All data epochs were down-sampled to 250 Hz and then band-pass filtered from 7 Hz to 90 Hz with an IIR filter. Zero-phase forward and reverse filtering was implemented using the filtfilt() function in MATLAB. Considering a latency delay in the visual pathway, the data epochs were extracted in [0.14 s 0.14 + d s], where the time 0 indicated stimulus onset and d indicated data length used in the offline analysis.

2) Online BCI Experiment: An online experiment was conducted to validate the feasibility of a high-speed BCI speller using the proposed method. The online experiment was divided into a training stage and a testing stage. The training stage consisted of 12 blocks, each including 40 trials. The data acquired in this stage were used as individual training data in the testing stage. The testing stage of the cue-guided selection task included 5 blocks, each also including 40 trials. Each trial lasted 0.8 s, including 0.3 s for visual stimulation and 0.5 s for gaze shifting. The cue for next target appeared right after the stimulus offset. Visual and auditory feedbacks were provided to the subjects in real time. A short beep was sounded after a target was correctly identified by the online analysis program. At the same time, the target character was typed in the text input field on the top of the screen. In the online experiment, the BCI system spelled at a rate of 75 characters/min (i.e., 0.8 s/character). In the testing stage of the free spelling task, the subjects were asked to input 'HIGH SPEED BCI' three times without visual cues (42 characters in total) [2]. 'Backspace' was used to remove an incorrect input. The gaze shifting time could vary between subjects according to their spelling accuracy. The ensemble TRCA-based method described below was used in the online experiment.

C. Target Identification Algorithm

This study employed the CCA and TRCA-based methods with individual calibration data to identify the target stimuli. Individual calibration data for the n-th stimulus and single-trial test data are denoted by a four-way tensor $\boldsymbol{\chi} = (\chi)_{njkh} \in \mathbb{R}^{N_f \times N_c \times N_s \times N_t}$ and a two-way tensor $\boldsymbol{X} \in \mathbb{R}^{N_c \times N_s}$, respectively. Here, n indicates the stimulus index, N_f is the number of stimuli, j indicates the channel index, N_c is the number of channels, k indicates the index of sample points, N_s is the number of sampling points in each trial, h indicates the index of training trials, and N_t is the number of training trials. The goal of target identification is to take an input X and assign it to one of N_f classes L_n. L_n corresponds to the stimulation frequency $f_n \in \{f_1, f_2, \cdots, f_{N_f}\}$. The filter bank analysis was also applied to decompose SSVEPs into sub-band components so that independent information embedded in the harmonic components can be extracted more efficiently [30]. According to the study [30], the lower and upper cut-off frequencies of the m-th sub-band were set to $m \times 8$ Hz and 90 Hz, respectively. After applying m-th zero-phase Chebyshev Type I Infinite impulse response (IIR) filters, the training data and the test data are denoted as $\boldsymbol{\chi}^{(m)} \in \mathbb{R}^{N_f \times N_c \times N_s \times N_t}$ and $\boldsymbol{X}^{(m)} \in \mathbb{R}^{N_c \times N_s}$, respectively. In the template matching-based framework of target identification, correlation-based feature values for the m-th sub-band and n-th stimuli can be calculated as $r_n^{(m)} = f(\boldsymbol{\chi}^{(m)}, \boldsymbol{X}^{(m)})$. A weighted sum of squares of the combined correlation values corresponding to all sub-band components was calculated as the feature for target identification:

$$\rho_n = \sum_{m=1}^{N_m} a(m) \cdot \left(r_n^{(m)}\right)^2 \tag{1}$$

where, N_m is the total number of sub-bands, and $a(m)$ was defined as $a(m) = m^{-1.25} + 0.25$ according to [30]. Target class L_τ can be identified by the following equation:

$$\tau = \arg \max_n \rho_n, \; n = 1, 2, \ldots, N_f \tag{2}$$

In SSVEP-based BCIs, it is of importance to find better feature extraction method $f(\cdot)$ to optimize the accuracy of target identification.

1) Extended CCA-Based Method: Our recent studies proposed an extended CCA-based method, which incorporated individual template signals obtained by averaging multiple training trials as $\bar{\chi}_{njk} = \frac{1}{N_t} \sum_{h=1}^{N_t} \chi_{njkh}$ [4], [9], [18], [19]. Correlation coefficients between projections of a test data $\boldsymbol{X}^{(m)}$ and an individual template $\bar{\boldsymbol{\chi}}_n^{(m)} \in \mathbb{R}^{N_c \times N_t}$ using CCA-based spatial filters are used as features for target identification. Specifically, the following three weight vectors are used as spatial filters to enhance the SNR of SSVEPs: (1) $\boldsymbol{W}_X^{(m)}(\boldsymbol{X}^{(m)}, \bar{\boldsymbol{\chi}}_n^{(m)})$ between the test set $\boldsymbol{X}^{(m)}$ and the individual template $\bar{\boldsymbol{\chi}}_n^{(m)}$, (2) $\boldsymbol{W}_X^{(m)}(\boldsymbol{X}^{(m)}, \boldsymbol{Y}_f)$ between the test set $\boldsymbol{X}^{(m)}$ and sine-cosine reference signals $\boldsymbol{Y}_f$, (3) $\boldsymbol{W}_{\bar{\chi}_n}^{(m)}(\bar{\boldsymbol{\chi}}_n^{(m)}, \boldsymbol{Y}_f)$ between the individual template $\bar{\boldsymbol{\chi}}_n^{(m)}$ and sine-cosine reference signals $\boldsymbol{Y}_f$. In addition, the similarity between $\boldsymbol{W}_X^{(m)}(\boldsymbol{X}^{(m)}, \bar{\boldsymbol{\chi}}_n^{(m)})$ and $\boldsymbol{W}_{\bar{\chi}_n}^{(m)}(\boldsymbol{X}^{(m)}, \bar{\boldsymbol{\chi}}_n^{(m)})$ is indirectly measured by calculating correlation coefficient between the projections of template signals using the spatial filters. A correlation vector $\hat{\boldsymbol{r}}_n^{(m)}$ is defined as follows eq.(3) shown at the bottom of the next page where $r(a, b)$ indicates the Pearson's correlation coefficient between two signals a and b. An ensemble classifier is used to combine the five features. In practice, the following weighted correlation coefficient $r_n^{(m)}$ is used as the final feature in target

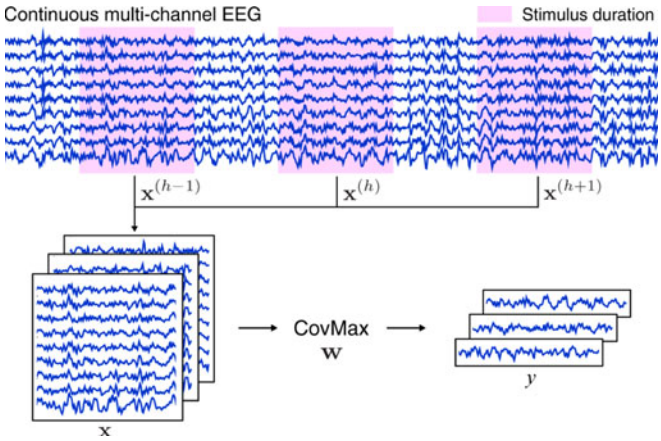

Fig. 2. Diagram of the task-related component analysis (TRCA) in SSVEP analysis.

identification:

$$r_n^{(m)} = \sum_{l=1}^{5} \text{sign}\left(r_{n,l}^{(m)}\right) \cdot \left(r_{n,l}^{(m)}\right)^2 \tag{4}$$

where $\text{sign}()$ is used to retain discriminative information from negative correlation coefficients between the test set and individual templates. Target can be identified by (2).

2) TRCA-Based Method: TRCA is the method that extracts task-related components efficiently by maximizing the reproducibility during the task period (Fig. 2) [25], [26]. Here, two source signals are assumed: 1) task-related signal $s(t) \in \mathbb{R}$; 2) task-unrelated signal $n(t) \in \mathbb{R}$. A linear generative model of observed multichannel EEG signal $\boldsymbol{x}(t) \in \mathbb{R}^{N_c}$ is assumed as:

$$x_j(t) = a_{1,j}s(t) + a_{2,j}n(t),\ j = 1, 2, \ldots, N_c \tag{5}$$

where, j is the index of channels, and $a_{1,j}$ and $a_{2,j}$ are mixing coefficients that project the source signals to the EEG signal. The problem is to recover the task-related component $s(t)$ from a linear sum of observed signals $\boldsymbol{x}(t)$ as:

$$\begin{aligned} y(t) &= \sum_{j=1}^{N_c} w_j x_j(t) \\ &= \sum_{j=1}^{N_c} \left(w_j a_{1,j} s(t) + w_j a_{2,j} n(t)\right). \end{aligned} \tag{6}$$

Ideally, the problem has a solution of $\sum_{j=1}^{N_c} w_j a_{1,j} = 1$ and $\sum_{j=1}^{N_c} w_j a_{2,j} = 0$, leading to the final solution $y(t) = s(t)$. The problem can be solved by inter-trial covariance maximization. The h-th trial of EEG signal and the estimated task-related component are described as $\boldsymbol{x}^{(h)}(t)$ and $y^{(h)}(t),\ h = 1, 2, \ldots, N_t$. The periods of $y^{(h)}(t)$ are fixed as $t \in [t_h,\ t_h + T]$. Here T is the duration of each task trial. The covariance between the h_1-th and the h_2-th trials of $y(t)$ is described as:

$$\begin{aligned} C_{h_1 h_2} &= \text{Cov}\left(y^{(h_1)}(t), y^{(h_2)}(t)\right) \\ &= \sum_{j_1, j_2 = 1}^{N_c} w_{j_1} w_{j_2} \text{Cov}\left(x_{j_1}^{(h_1)}(t), x_{j_2}^{(h_2)}(t)\right). \end{aligned} \tag{7}$$

All possible combinations of trials are summed as:

$$\begin{aligned} \sum_{\substack{h_1, h_2 = 1 \\ h_1 \neq h_2}}^{N_t} C_{h_1 h_2} &= \sum_{\substack{h_1, h_2 = 1 \\ h_1 \neq h_2}}^{N_t} \sum_{j_1, j_2 = 1}^{N_c} w_{j_1} w_{j_2} \\ &\quad \text{Cov}\left(x_{j_1}^{(h_1)}(t), x_{j_2}^{(h_2)}(t)\right) \\ &= \boldsymbol{w}^T \boldsymbol{S} \boldsymbol{w}. \end{aligned} \tag{8}$$

Here, the matrix $\boldsymbol{S} = (S_{j_1 j_2})_{1 \leq j_1, j_2 \leq N_c}$ is defined as:

$$S_{j_1 j_2} = \sum_{\substack{h_1, h_2 = 1 \\ h_1 \neq h_2}}^{N_t} \text{Cov}\left(x_{j_1}^{(h_1)}(t), x_{j_2}^{(h_2)}(t)\right). \tag{9}$$

To obtain a finite solution, the variance of $y(t)$ is constrained as:

$$\begin{aligned} \text{Var}(y(t)) &= \sum_{j_1, j_2 = 1}^{N_c} w_{j_1} w_{j_2} \text{Cov}(x_{j_1}(t), x_{j_2}(t)) \\ &= \boldsymbol{w}^T \boldsymbol{Q} \boldsymbol{w} \\ &= 1. \end{aligned} \tag{10}$$

The constrained optimization problem can be solved as:

$$\hat{\boldsymbol{w}} = \arg \max_{\boldsymbol{w}} \frac{\boldsymbol{w}^T \boldsymbol{S} \boldsymbol{w}}{\boldsymbol{w}^T \boldsymbol{Q} \boldsymbol{w}}. \tag{11}$$

The optimal coefficient vector is obtained as the eigenvector of the matrix $\boldsymbol{Q}^{-1}\boldsymbol{S}$. In general, N_c eigenvalues and eigenvectors are obtained from a $N_c \times N_c$ matrix. The eigenvalues $\boldsymbol{\lambda}$ of the

$$\hat{\boldsymbol{r}}_n^{(m)} = \begin{bmatrix} r_{n,1}^{(m)} \\ r_{n,2}^{(m)} \\ r_{n,3}^{(m)} \\ r_{n,4}^{(m)} \\ r_{n,5}^{(m)} \end{bmatrix} = \begin{bmatrix} \rho\left(\boldsymbol{X}^{(m)\mathrm{T}} \boldsymbol{W}_X^{(m)}\left(\boldsymbol{X}^{(m)}, \boldsymbol{Y}_f\right), \boldsymbol{Y}_f{}^{\mathrm{T}} \boldsymbol{W}_Y^{(m)}\left(\boldsymbol{X}^{(m)}, \boldsymbol{Y}_f\right)\right) \\ \rho\left(\boldsymbol{X}^{(m)\mathrm{T}} \boldsymbol{W}_X^{(m)}\left(\boldsymbol{X}^{(m)}, \bar{\boldsymbol{X}}_n^{(m)}\right), \bar{\boldsymbol{X}}_n^{(m)\mathrm{T}} \boldsymbol{W}_X^{(m)}\left(\boldsymbol{X}^{(m)}, \bar{\boldsymbol{X}}_n^{(m)}\right)\right) \\ \rho\left(\boldsymbol{X}^{(m)\mathrm{T}} \boldsymbol{W}_X^{(m)}\left(\boldsymbol{X}^{(m)}, \boldsymbol{Y}_f\right), \bar{\boldsymbol{X}}_n^{(m)\mathrm{T}} \boldsymbol{W}_X^{(m)}\left(\boldsymbol{X}^{(m)}, \boldsymbol{Y}_f\right)\right) \\ \rho\left(\boldsymbol{X}^{(m)\mathrm{T}} \boldsymbol{W}_{\bar{X}_n}^{(m)}\left(\bar{\boldsymbol{X}}_n^{(m)}, \boldsymbol{Y}_f\right), \bar{\boldsymbol{X}}_n^{(m)\mathrm{T}} \boldsymbol{W}_{\bar{X}_n}^{(m)}\left(\bar{\boldsymbol{X}}_n^{(m)}, \boldsymbol{Y}_f\right)\right) \\ \rho\left(\bar{\boldsymbol{X}}_n^{(m)\mathrm{T}} \boldsymbol{W}_X^{(m)}\left(\boldsymbol{X}^{(m)}, \bar{\boldsymbol{X}}_n^{(m)}\right), \bar{\boldsymbol{X}}_n^{(m)\mathrm{T}} \boldsymbol{W}_{\bar{X}_n}^{(m)}\left(\boldsymbol{X}^{(m)}, \bar{\boldsymbol{X}}_n^{(m)}\right)\right) \end{bmatrix} \tag{3}$$

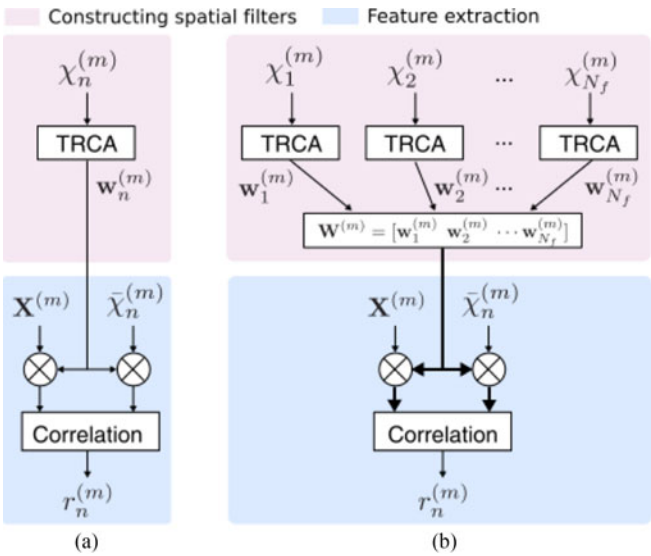

Fig. 3. Diagrams of the proposed methods. (a) The TRCA-based method and (b) The ensemble TRCA-based method. 1-D and 2-D correlation analyses were used in (a) and (b), respectively.

matrix $\boldsymbol{Q}^{-1}\boldsymbol{S}$, which can be obtained in a descending order, represents the value of the cost function for the corresponding eigenvector $\hat{\boldsymbol{w}}$. Therefore, the eigenvalues indicate the task consistency among multiple trials.

In SSVEP-based BCIs, TRCA can be used to design spatial filters for removing background EEG activities from scalp recordings. Fig. 3(a) shows the flowchart of the TRCA-based method. Spatial filters for the n-th stimulus $\boldsymbol{w}_n^{(m)} \in \mathbb{R}^{N_c}$. are obtained through TRCA from the individual calibration data $\boldsymbol{\chi}_n^{(m)}$ after applying the filter bank analysis. Here, $\boldsymbol{Q}$ in the (10) was calculated with a concatenated matrix of all the training trials $\boldsymbol{\chi}_n^{(m)}$. The correlation coefficient between single-trial test data $\boldsymbol{X}^{(m)} \in \mathbb{R}^{N_c \times N_s}$ and averaged training data across trials for n-th visual stimulus $\bar{\boldsymbol{\chi}}_n^{(m)} \in \mathbb{R}^{N_c \times N_s}$ is calculated as:

$$r_n^{(m)} = \rho\left(\left(\boldsymbol{X}^{(m)}\right)^T \boldsymbol{w}_n^{(m)}, \left(\bar{\boldsymbol{\chi}}_n^{(m)}\right)^T \boldsymbol{w}_n^{(m)}\right) \tag{12}$$

where, $\rho(a, b)$ indicates the Pearson's correlation analysis between two signals a and b.

3) Ensemble TRCA-Based Method: This study also proposes an ensemble TRCA-based target identification method shown in Fig. 3(b). Since there are N_f individual calibration data corresponding to all visual stimuli, N_f different spatial filters can be obtained. Ideally, they should be similar to each other because the mixing coefficients from SSVEP source signals to scalp recordings could be considered similar within the used frequency range [31], [32], which indicates the possibility of further improvements by integrating all spatial filters. An ensemble spatial filter $\boldsymbol{W}^{(m)} \in \mathbb{R}^{N_c \times N_f}$ is obtained as follows:

$$\boldsymbol{W}^{(m)} = \left[\boldsymbol{w}_1^{(m)}\ \boldsymbol{w}_2^{(m)}\ \cdots\ \boldsymbol{w}_{N_f}^{(m)}\right]. \tag{13}$$

Then the (12) can be modified as follows:

$$r_n^{(m)} = \rho\left(\left(\boldsymbol{X}^{(m)}\right)^T \boldsymbol{W}^{(m)}, \left(\bar{\boldsymbol{\chi}}_n^{(m)}\right)^T \boldsymbol{W}^{(m)}\right) \tag{14}$$

where, $\rho(\boldsymbol{a},\ \boldsymbol{b})$ indicates the two-dimensional correlation analysis between $\boldsymbol{a}$ and $\boldsymbol{b}$. After final features ρ_n are obtained by merging the correlation coefficients corresponding to all sub-band components, the intended target can be identified via (2).

III. Results

A. Target Identification Performance

Fig. 4(a) and (b) show the averaged classification accuracy and putative ITRs across all subjects with different data lengths from 100 ms to 500 ms with an interval of 100 ms. The number of channels (i.e., N_c), and the number of sub-bands (i.e., N_m) were set to 9, and 5, respectively, for all subjects. The accuracy and the ITR were estimated by a leave-one-out cross-validation, in which 11 blocks were used as training data and 1 block was used as test data (i.e., $N_t = 11$). The comparison of the three methods indicates that the proposed TRCA-based methods outperform the extended CCA-based method. The ensemble TRCA-based method achieved the highest performance regardless of data length. A two-way repeated measures ANOVA showed main effects of accuracy ($F(2, 22) = 124.90$, $p < 0.001$) and data length ($F(4, 44) = 127.56$, $p < 0.001$), and significant interaction of these two factors ($F(8, 88) = 21.78$, $p < 0.001$). This tendency of classification accuracy was consistent to that of the putative ITR. A two-way repeated measures ANOVA revealed significant main effects of the ITR ($F(2, 22) = 105.77$, $p < 0.001$) and data length ($F(4, 44) = 58.19$, $p < 0.001$), and significant interaction of them ($F(8,88) = 14.69$, $p < 0.001$). The highest ITRs were achieved with different data lengths for each method (Extended CCA: $d = 400$ ms, 262.66 $\pm$ 48.19 bits/min; TRCA: $d = 400$ ms, 278.54 $\pm$ 42.77 bits/min; Ensemble TRCA: $d = 300$ ms, 288.86 $\pm$ 62.57 bits/ min).

B. Parameter Optimization

Fig. 5(a), (b), and (c) show the classification accuracy with different parameters: the number of training trials (i.e., N_t), the number of electrodes (i.e., N_c), and the number of sub-bands (i.e., N_m). A grid-search was conducted to optimize the parameters, in which the accuracy was calculated with all combinations of the parameters using 300-ms long SSVEP data. As the result of the grid-search, the highest accuracy was obtained when the numbers of training trials, electrodes, and sub-bands were 11, 9, and 5, respectively. In the Fig. 5, therefore, except when each parameter was targeted for searching, the number of training data, electrodes, and sub-bands were fixed to 11, 9, and 5, respectively. Two-way repeated measures ANOVA showed significant main effects of algorithm ($F(2, 22) = 44.67$, $p < 0.001$) and the number of training trials ($F(9, 99) = 88.93$, $p < 0.001$). In the comparison of the number of electrodes, there were significant main effects of algorithm ($F(2, 22) = 34.64$, $p < 0.001$) and electrodes ($F(5, 55) = 10.93$, $p < 0.001$). On the other hand, there were significant main effects of

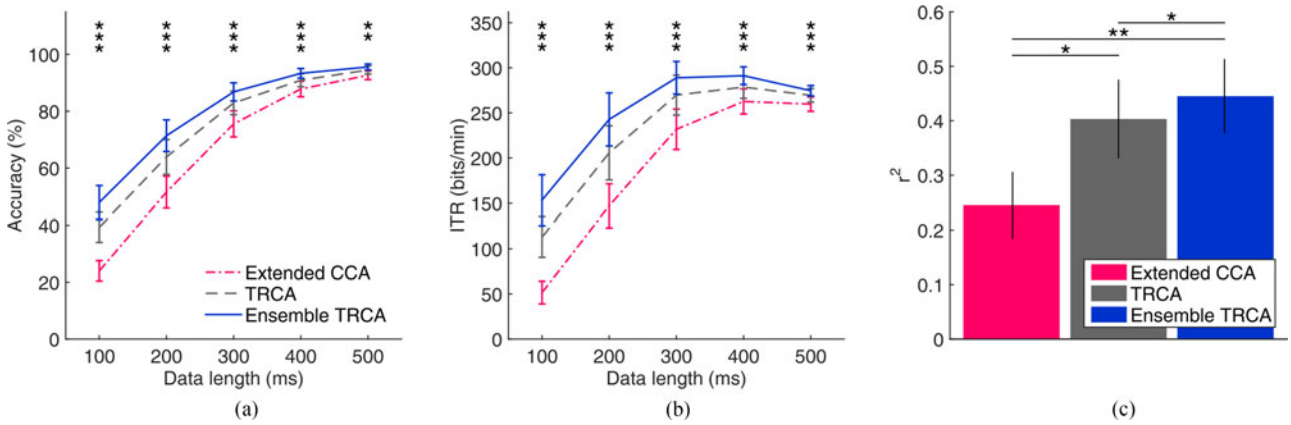

Fig. 4. Performance comparison of three methods using recorded SSVEPs. (a) Averaged classification accuracy, (b) simulated ITRs across subjects as a function of data length (d). (c) An example of r-square values for SSVEPs at 9.6 Hz for each method (data length 300 ms). The error bars indicate standard errors. The asterisks in the subfigure (a) and (b) indicate significant difference between the three methods obtained by one-way repeated measures ANOVAs, and those in the subfigure (c) indicate significant difference between each pair of the two methods by paired t-tests ($*p<0.05$, $**p<0.01$, $***p<0.001$).

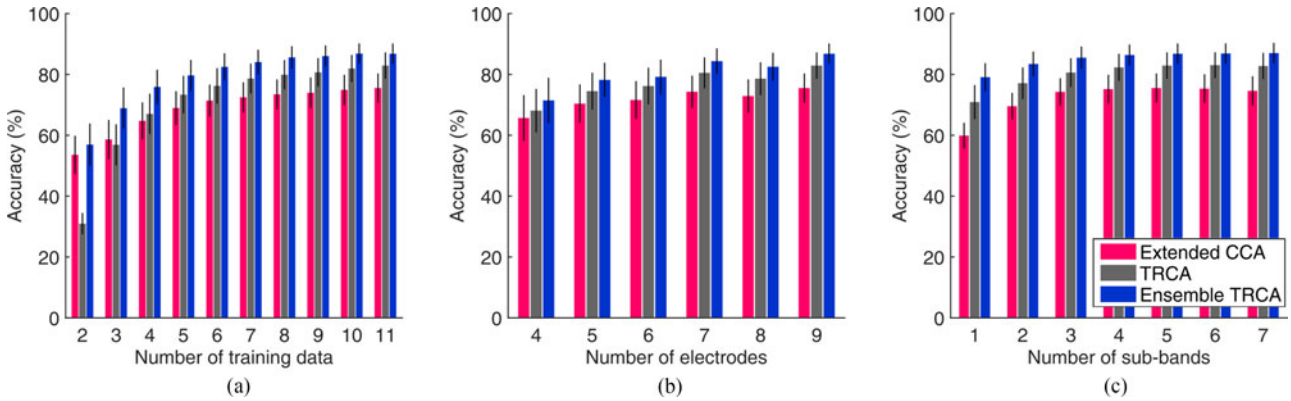

Fig. 5. Averaged classification accuracy with (a) different numbers of training trials (N_t), (b) different numbers of electrodes (N_c), and (c) different numbers of sub-bands (N_m) across subjects with 300ms-long epochs. The error bars indicate standard errors.

algorithm ($F(2,22) = 49.36$, $p < 0.001$) and the number of sub-bands ($F(6,66) = 43.85$, $p < 0.001$). There were also significant interactions of algorithms and each parameter (N_t : $F(18,198) = 21.32$, $p < 0.001$; N_c : $F(10,110) = 3.65$, $p < 0.001$; N_m : $F(12,132) = 6.61$, $p < 0.001$).

C. Online BCI Performance

Table I lists the results of the online cue-guided BCI experiment. Towards the highest ITR obtained in the offline analysis, a fixed stimulus duration of 300 ms and a gaze shifting duration of 500 ms were used for all subjects, leading to the spelling speed of 800 ms per character (i.e., 75 characters per minute). The parameters optimized by the grid-search were used for all subjects. The averaged accuracy across all subjects was 89.83 ± 6.07%, leading to an averaged ITR of 325.33 ± 38.17 bits/min. Across individuals, the minimal and maximal ITR was 263.00 bits/min (s6 and s9) and 376.58 bits/min (s1 and s15) respectively. Table II lists the results of the free spelling experiment. Due to increased gaze shifting time (0.7 – 1.2 s), the mean spelling rate was 36.17 ± 11.02 characters per minute, which was much lower than 75 characters per minute in the cue-guided task. The mean ITR was 198.67 ± 50.48 bits/min across the ten subjects. These results demonstrated the efficacy of the ensemble TRCA-based method in an online SSVEP-based BCI.

IV. Discussions

Improving communication speed of BCIs is a critical issue for various applications including BCI spellers. To the best of our knowledge, an online ITR of 325.33 ± 38.17 bits/min is the highest ITR reported in EEG-based BCI studies. Compared with the ITR of 267 bits/min reported in [2] and in this study, the improvement of performance in the present study can be attributed to the difference of the target identification algorithm, since the other elements such as stimulus design and the number of electrodes in these two studies are identical. Recent studies have suggested that spatial filtering techniques are effective to remove background EEG signals, leading to enhanced SNR and discriminability of SSVEPs [13], [14]. In this study, the discriminability of SSVEPs was evaluated by r-squares. The r-square values were calculated with correlation

TABLE I
RESULTS OF ONLINE CUED-GUIDED EXPERIMENTS

Subject	No. of trials (Correct/Incorrect)	Accuracy [%]	ITR [bits/min]
s1	195/5	97.50	376.58
s2	183/17	91.50	333.98
s3	194/6	97.00	372.67
s4	189/11	94.50	354.30
s5	167/33	83.50	285.28
s6	159/41	79.50	263.00
s7	161/39	80.50	268.46
s8	174/26	87.00	305.80
s9	159/41	79.50	263.00
s10	169/31	84.50	291.04
s11	181/19	90.50	327.52
s12	185/15	92.50	340.59
s13	177/23	88.50	314.95
s14	168/32	84.00	288.15
s15	195/5	97.50	376.58
s16	185/15	92.50	340.59
s17	189/11	94.50	354.30
s18	188/12	94.00	350.80
s19	192/8	96.00	365.12
s20	183/17	91.50	333.98
Mean ± STD	-	89.83 ± 6.07	325.33 ± 38.17

coefficients corresponding to the target stimulus (i.e., $\rho_{n=\tau}$) and the maximal feature values corresponding to non-target stimuli (i.e., $\rho_{n\neq\tau}$) obtained through the three methods [18]. A two-way repeated measures ANOVA showed significant main effects of methods ($F(2, 22) = 25.53$, $p < 0.001$) and targets ($F(39, 429) = 3.43$, $p < 0.001$), and significant interaction between them ($F(78, 858) = 1.82$, $p < 0.001$). Fig. 4(c) depicts examples of averaged *r*-square values for SSVEPs at 9.6 Hz across all subjects. Paired *t*-tests showed the *r*-square of the TRCA-based methods are significantly higher than that of the extended CCA-based methods (TRCA vs. Extended CCA: $p = 0.038$; Ensemble TRCA vs. Extended CCA: $p = 0.008$). The results indicate that the TRCA-based spatial filtering could increase the distance between target and non-target feature values by removing background EEG activities, resulting in improved accuracy in target identification. The ensemble TRCA method also shows a significantly higher *r*-square value than the TRCA method ($p = 0.013$), indicating that the integration of multiple spatial filters can further improve the SNR of SSVEPs. The ensemble approach can also be applied to other spatial filtering approaches in detecting multi-class SSVEPs.

The direct comparison of performance of the three methods revealed the effectiveness of the proposed methods. With the TRCA-based methods, the data length could be reduced toward the highest performance compared with the extended CCA-based method (Extended CCA: 400 ms; TRCA: 400 ms; Ensemble TRCA: 300 ms). For all the methods, the data length could be further optimized by applying dynamic stopping approaches, which adaptively determine a selection time in each single trial [33], [34]. In the online cue-guided experiments, all of the subjects could perform the target-selection task with a speed of 800 ms per character (Target gazing time: 300 ms, Gaze shifting time: 500 ms) using the ensemble TRCA-based method. Note that subjects needed to be well trained to realize the speed of 0.8 s per character. Unexperienced users, therefore, might require a longer gaze switching duration (e.g., 1 – 2 seconds) in free spelling tasks, resulting in lower ITR than the one in the cue-guided task. The free-spelling experiment performed in this study achieved 198.67 ± 50.48 bits/min with a gaze switching time of 0.7 – 1.2 s. Importantly, all subjects performed the free spelling task with the stimulation duration of 0.3 s, which is the shortest in all the reported online SSVEP-based BCIs. That is, the communication rate might approach to the theoretical ITR obtained in the cue-guided experiment once they get familiar with the layout of visual stimuli. This study used the same design of visual stimulator as the one used in our previous study for a direct comparison of target identification methods [2], however, the performance could be further improved by optimizing the parameter of stimuli for the TRCA-based method. For example, the framework of designing an SSVEP-based BCI proposed by Chen *et al.* [2] suggested running a grid-search to optimize a phase interval and stimulus duration.

All the methods used in this study require acquiring individual training data prior to an online operation. The number of training data is an important parameter in target identification. As shown in Fig. 5(a), the accuracy increased when the number of training trials increased. Assuming 10 trials for each of the 40 visual stimuli, the training data with a trial length of 0.8 s (Target gazing time: 0.3 s; Gaze shifting time: 0.5 s) can be acquired within approximately 5 minutes excluding break time. Several template-transferring approaches (e.g., transfer across different subjects [35] or different recording sessions [36]) have been proposed to reduce the training time. Another approach that updates template signals in the online operation proposed by Yuan *et al.* might be used to reduce the training time and enhance the performance [33]. These approaches could be applied to the TRCA-based methods for reducing the training time. An asynchronous system design, in which users can switch between control and non-control states of the system at their own pace, is also important for implementing a more natural communication tool. It could be realized by simply adding another visual stimulus [11] or combining SSVEPs with other EEG features [37] to control the on/off states of BCI.

The present design of visual stimulator uses stimuli flickering in the alpha band (e.g., 8 – 13 Hz) to achieve high performance. However, visual flickers at higher frequencies above the flicker fusion threshold might be a better choice to reduce visual fatigue and discomfort for practical applications. In general, the amplitude of SSVEPs decreases as stimulus frequency increases [12]. The proposed methods might have great potential to improve the detection accuracy of the high-frequency SSVEPs. In addition, the TRCA-based spatial filtering method could be applied to enhance SNRs of other time-locked EEG components such as event-related potentials (ERPs), which have been widely used in the fields of BCI, cognitive neuroscience, and clinical applications. In particular, P300-based BCIs generally require an averaging process across multiple trials to increase detection accuracy, leading to slow communication speed and low ITR [38]. The number of trials required in the averaging process could be reduced by employing the TRCA-based spatial filter.

TABLE II
RESULTS OF ONLINE FREE-SPELLING EXPERIMENTS

Subject	Trial length [s]	No. of trials (Correct/Incorrect)	Spelling rate [cpm]	ITR [bits/min]
s1	1.2 (0.3 + 0.9)	56 (49/7)	37.50	205.88
s5	1.2 (0.3 + 0.9)	44 (43/1)	47.73	252.27
s8	1.2 (0.3 + 0.9)	56 (49/7)	37.50	205.88
s10	1.4 (0.3 + 1.1)	42 (42/0)	42.86	228.08
s15	1.5 (0.3 + 1.2)	104 (73/31)	16.15	114.70
s16	1.5 (0.3 + 1.2)	68 (55/13)	24.71	144.30
s17	1.0 (0.3 + 0.7)	48 (45/3)	52.50	279.26
s18	1.5 (0.3 + 1.2)	42 (42/0)	40.00	212.88
s19	1.5 (0.3 + 1.2)	46 (44/2)	36.52	193.36
s20	1.5 (0.3 + 1.2)	64 (53/11)	26.25	150.06
Mean ± STD	-	-	36.17 ± 11.02	198.67 ± 50.48

The subjects were asked to input 'HIGH SPEED BCI' three times without visual cues (42 characters in total). 'Backspace' was used to remove an incorrect input. For trial length, the two values in brackets correspond to stimulation duration and gaze shifting time respectively. The gaze shifting time could vary between subjects (from 0.7 to 1.2 seconds).

V. CONCLUSION

A novel SSVEP detection method using the TRCA-based spatial filters was proposed and evaluated in this paper. The offline analysis results showed that the proposed method significantly outperformed the extended CCA-based method in terms of classification accuracy and ITR. The feasibility of the proposed method in online applications was also validated by a cue-guided spelling task and a free-spelling task. To our knowledge, the online ITR of 325.33 ± 38.17 bits/min in the cue-guided spelling task was the highest ITR reported in EEG-based BCIs. Since the design of the SSVEP-based BCI can be utilized in many kinds of applications, this study will encourage more real-life BCI applications for communication and control.

ACKNOWLEDGMENT

The authors would like to thank Dr. H. Tanaka for his advice.

REFERENCES

[1] S. Gao *et al.*, "Visual and auditory brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1436–1447, May 2014.
[2] X. Chen *et al.*, "High-speed spelling with a noninvasive brain-computer interface," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 44, pp. E6058–E6067, 2015.
[3] Y. Wang *et al.*, "Brain-computer interfaces based on visual evoked potentials: feasibility of practical system design," *IEEE Eng. Med. Biol. Mag.*, vol. 27, no. 5, pp. 64–71, Sep. 2008.
[4] M. Nakanishi *et al.*, "A high-speed brain speller using steady-state visual evoked potentials," *Int. J. Neural Syst.*, vol. 24, no. 6, 2014, Art. no. 1450019.
[5] Y. Kimura *et al.*, "SSVEP-based brain-computer interfaces using FSK-modulated visual stimuli," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2831–2838, Oct. 2013.
[6] E. Yin *et al.*, "A hybrid brain-computer interface based on the fusion of P300 and SSVEP scores," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 23, no. 4, pp. 693–710, Jul. 2015.
[7] Y. Wang *et al.*, "Visual stimulus design for high-rate SSVEP BCI," *Electron. Lett.*, vol. 46, no. 15, pp. 1057–1058, 2010.
[8] M. Nakanishi *et al.*, "Generating visual flickers for eliciting robust steady-state visual evoked potentials at flexible frequencies using monitor refresh rate," *PLoS One*, vol. 9, no. 6, 2014, Art. no. e99235.
[9] X. Chen *et al.*, "A high-ITR SSVEP-based BCI speller," *Brain-Comp. Interfaces*, vol. 1, no. 3/4, pp. 181–191, 2014.
[10] 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.
[11] M. Cheng *et al.*, "Design and implementation of a brain-computer interface with high transfer rate," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 10, pp. 1181–1186, Oct. 2002.
[12] Y. Wang *et al.*, "A practical VEP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 234–239, Jun. 2006.
[13] Z. Lin *et al.*, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 6, pp. 1172–1176, Jun. 2007.
[14] G. Bin *et al.*, "An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method," *J. Neural Eng.*, vol. 6, 2009, Art. no. 046002.
[15] Y. Zhang *et al.*, "Multiway canonical correlation analysis for frequency components recognition in SSVEP-based BCIs," in *Proc. 18th Int. Conf. Neural Inform. Process.*, 2011, pp. 287–295.
[16] 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, Nov. 2013.
[17] Y. Zhang *et al.*, "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis," *Int. J. Neural Syst.*, vol. 24, no. 4, 2014, Art. no. 1450013.
[18] 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.
[19] M. Nakanishi *et al.*, "A comparison study of canonical correlation analysis based methods for detecting steady-state visual evoked potentials," *PLoS One*, vol. 10, no. 10, 2015, Art. no. e140703.
[20] J. Onton *et al.*, "Imaging human EEG dynamics using independent component analysis," *Neurosci. Biobehav. Rev.*, vol. 30, no. 6, pp. 808–822, 2006.
[21] S. Makeig *et al.*, "Independent component analysis of electroencephalographic data," in *Advances in Neural Information Processing Systems*, D. Touretzky *et al.*, Eds., Cambridge, MA, USA: MIT Press, 1996, pp. 145–151.
[22] T.-P. Jung *et al.*, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiol*, vol. 37, pp. 163–178, 2000.
[23] T.-P. Jung *et al.*, "Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects," *Clin. Neurophysiol.*, vol. 111, pp. 1745–1758, 2000.
[24] A. Delorme *et al.*, "Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis," *NeuroImage*, vol. 34, no. 4, pp. 1443–1449, 2007.
[25] H. Takana *et al.*, "Task-related component analysis for functional neuroimaging and application to near-infrared spectroscopy data," *NeuroImage*, vol. 64, pp. 308–327, 2013.
[26] H. Takana *et al.*, "Task-related oxygenation and cerebral blood volume changes estimated form NIRS signals in motor and cognitive tasks," *NeuroImage*, vol. 94, pp. 107–119, 2014.
[27] D. Regan, "Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine," *Brit. J. Ophthalmol.*, vol. 74, no. 4, p. 255, 1990.

[28] F. B. Vialatte *et al.*, "Steady-state visual evoked potentials: Focus on essential paradigms and future perspectives," *Prog. Neurobiol*., vol. 90, pp. 418–438, 2010.
[29] D. H. Brainard, "The psychophysics toolbox," *Spat. Vis.*, vol. 10, pp. 433–436, 1997.
[30] X. Chen *et al.*, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain-computer interface," *J. Neural Eng.*, vol. 12, 2015, Art. no. 046008.
[31] R. Srinivasan *et al.*, "Steady-state visual evoked potentials: distributed local sources and wave-like dynamics are sensitive to flicker frequency," *Brain Topography*, vol. 18, no. 3, pp. 167–187, 2006.
[32] 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. 1–13, 2009.
[33] E. Yin *et al.*, "A dynamically optimized SSVEP brain-computer interface (BCI) speller," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1447–1546, Jun. 2015.
[34] M. Nakanishi *et al.*, "A dynamic stopping method for improve performance of steady-state visual evoked potential based brain-computer interfaces," in *Proc. 37th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2015, pp. 1057–1060.
[35] P. Yuan *et al.*, "Enhancing performance of SSVEP-based brain-computer interface via exploiting inter-subject information," *J. Neural Eng.*, vol. 12, 2015, Art. no. 046006.
[36] M. Nakanishi *et al.*, "Session-to-session transfer in detecting steady-state visual evoked potentials with individual training data," in *Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience*, Part I, LNAI 9743, D. D. Schmorrow *et al.*, Eds. Switzerland: Springer, 2016, pp. 253–260.
[37] G. Pfurtscheller *et al.*, "The hybrid BCI," *Front. Neurosci.*, vol. 4, 2010, Art. no. 42.
[38] L. A. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, pp. 510–523, 1988.