IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 13, 2019, accepted June 20, 2019, date of publication June 26, 2019, date of current version July 15, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2925078*

# Maximum Signal Fraction Analysis for Enhancing Signal-to-Noise Ratio of EEG Signals in SSVEP-Based BCIs

**QINGGUO WEI[1], SHAN ZHU[1], YIJUN WANG[2], XIAORONG GAO[3], HAI GUO[1], AND XUAN WU[1]**

[1]Department of Electronic Information Engineering, School of Information Engineering, Nanchang University, Nanchang 330031, China
[2]State Key Laboratory on Integrated Optoelectronics, Institute Semiconductors, Chinese Academy of Science, Beijing 100083, China
[3]Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China

Corresponding author: Qingguo Wei (wqg07@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61663025 and Grant 61365013, and in part by the Jiangxi Provincial Department of Science and Technology under Grant 20151BBE5006.

**ABSTRACT** Various improved canonical correlation analysis (CCA) methods were developed for enhancing the performance of steady-state visual evoked potential (SSVEP)-based brain–computer interfaces (BCIs). Among them, the method combining CCA spatial filters from sine-cosine references and individual templates yielded the highest performance. However, the CCA aims to optimize the correlation between two sets of variables rather than the signal-to-noise ratio (SNR) of the SSVEP signals, upon which the performance of an SSVEP-based BCI depends mainly. In this paper, a novel algorithm, namely, maximum signal fraction analysis (MSFA), is proposed for creating spatial filters based on individual training data. The spatial filter for a specific stimulus target is estimated by directly maximizing the averaged SNR of the observed signals across multiple trials. An individual template is calculated for each target by averaging training signals of multiple trials. Target recognition is based on template matching between filtered template signals and a single-trial testing signal. Classification performance of the MSFA-based method was evaluated on a benchmark dataset and compared with that of the CCA-based methods. The results suggest that the proposed MSFA method significantly outperforms the CCA-based methods in terms of classification accuracy, and thus, it has great potential to be applied in the real-life SSVEP-based BCI systems.

**INDEX TERMS** Brain-computer interface, steady-state visual evoked potential, maximum signal fraction analysis, canonical correlation analysis, signal-to-noise ratio.

## I. INTRODUCTION

A brain-computer interface (BCI) is a non-muscular communication channel between the brain and an external device [1]. Such a channel can help people with severe motor disabilities communicate with the external world and thus improve their quality of lives. In the past decades, electroencephalogram (EEG)-based BCIs have attracted attention in the field of neural engineering and clinical rehabilitation. BCI systems can be built using several paradigms such as motor imagery, visual evoked potentials (VEP) and P300 event-related potentials. Among them, the steady-state VEP (SSVEP)-based BCIs have become popular due to the advantages of high communication speed, low user variation and litter user training [2].

Great progresses have been achieved in the study of VEP-based BCIs in recent years [3]–[14]. However, creating a robust and practical BCI system remains a challenging problem [2]. The main reason is that the observed SSVEP signals are very noisy and exhibit low signal-to-noise ratio (SNR). In the EEG signals recorded on the scalp, the useful signal for target recognition is generated from task-related brain activities, while the noise is derived from spontaneous brain activities and other task-unrelated sources, which include power line interference, eye movements and eye blinks, etc. Some noise sources, e.g. eye blinks, produce voltage changes of much higher amplitude than the task-related brain activities. The low SNR of EEG signals severely limits the practical applications of SSVEP-based BCIs.

The associate editor coordinating the review of this manuscript and approving it for publication was Aysegul Ucar.

A widely used metric for BCI performance is the information transfer rate (ITR) defined by Wolpaw *et al.* [1]. According to the ITR metric, the three parameters to affect BCI performance are the number of selectable targets, the accuracy of target detection and the average time taken for one selection. Among the three parameters, the second one is the most important because high classification accuracy is necessary and essential for a practical BCI regardless of its type and number of selections. Thereby, in the study of SSVEP-based BCIs, much effort was made towards improving target recognition algorithms [8]. Power spectrum density analysis (PSDA) was a widely used method for detecting target frequency of SSVEP signals from a single electrode channel [13]–[14]. To make use of information from multiple channels, minimum energy combination (MEC) [15] and canonical correlation analysis (CCA) [16]–[17], as spatial filtering methods, were employed for target recognition. The standard CCA algorithm estimates spatial filters between a single-trial testing signal and sine-cosine reference signals. Although the algorithm exhibits better robustness than PSDA, it is susceptible to noise including the spontaneous EEG activities. To enhance the SNR of SSVEP signals, individual training data were incorporated in CCA algorithm and subsequently, diverse improved methods were developed for SSVEP frequency detection [7], [18]–[21]. The multiway CCA (MwayCCA) and the L1-reguralized MwayCCA (L1-MCCA) were first proposed to optimize sine-cosine reference signals of multiple standard CCA processes using training data [18], [19]. The multi-set CCA (MsetCCA) was proposed to optimize the reference signals with common features in multiple training trials [20] and exhibited better classification performance than the MwayCCA and the L1-MCCA. Nakanish *et al.* proposed an extended CCA method that combines CCA-based spatial filtering and template matching for detecting the frequency of SSVEP signals [6], [22]. Among the above methods, the extended CCA method uses CCA process to separately estimate multiple spatial filters, and the frequency detection is based on correlation between two spatially filtered signals. A comparison study showed that the extended CCA method yielded the best classification performance among all CCA-based methods [22].

Green et al. presented a linear transformation, maximum signal fraction analysis (MSFA), for maximizing the SNR of a set of variables [23]. MSFA was first used for reducing noise in multispectral satellite images [23] and achieved better denoising effect than principal component analysis (PCA). Subsequently, MSFA was applied to blind source separation [24], [25] and artifact removal [26], [27] of biomedical signals. As a kind of spatial filtering algorithm, MSFA can be used for removing noise of multichannel SSVEP signals. Different from the CCA algorithm, which estimates spatial filters by maximizing the underlying correlation between two sets of variables, MFSA estimates spatial filters by directly maximizing the SNR of one set of variables. Thereby, MSFA is expected to provide more robust feature signals when it is used to spatially filter multichannel SSVEP signals.

The detection accuracy of a BCI depends completely upon the SNR of the extracted feature signals used for classification. In this study, we proposed an MSFA-based algorithm for target recognition in order to enhance the performance of SSVEP-based BCIs. To the best of our knowledge, the MSFA-based analysis has never been used in BCI studies. The performance of MSFA was evaluated with a 40-target benchmark SSVEP data set recorded from 35 subjects [28]. In the performance evaluation, the CCA-based methods were used for comparing the performance with the proposed MSFA-based algorithm.

## II. MATERIALS AND METHODS

### A. DATA ACQUISITION AND PREPROCESSING

A publicly available benchmark SSVEP data set [28] was used in this study. The data set was acquired from 35 healthy subjects (18 males and 17 females, aged from 17 to 34 years, mean 22 years) while they performed a cue-guided target selecting task. Among these subjects, eight had experience using SSVEP-based BCIs, and the others were naïve. The user interface was a BCI speller consisting of 40 visual stimulus targets, which were coded using a joint frequency and phase modulation (JFPM) method [29]. The stimulus frequencies ranged from 8 Hz and 15.8 Hz with an interval of 0.2 Hz, whereas the phases ranged between 0 rad and $1.5\pi$ rad with an interval of $0.5\pi$ rad. The data set can currently be downloaded at https://pan.baidu.com/s/1qYhyLtE.

The EEG data were collected with a Synamps2 system (Neuroscan Inc.) at a sampling rate of 1000 Hz. Sixty-four electrodes according to an extended international 10/20 system were used for recording EEG. The reference electrode was positioned at vertex (Cz). Electrode impedances were kept below 10 kΩ during the experiment. Event trigger signals yielded by the stimulus program were recorded on an event channel synchronized to the EEG data. The experiment consisted of six blocks, each of which included 40 trials corresponding to all 40 targets cued in a random order. Each trial started with a visual cue, which lasted 0.5 s and achieved by reddening the target to be attended. Subjects were instructed to shift their gaze to the target as soon as possible. Subsequently, all stimulus targets flicked concurrently for 5 s. After the end of visual stimulation, the screen was blank for 0.5 s and then the next trial began. Subjects were asked to avoid eye blinks during the stimulus period. To avoid visual fatigue, they were allowed a short rest of several minutes between two consecutive blocks.

The continuous EEG data were segmented into epochs of length 6 s including 0.5 s before stimulus onset, 5 s for visual stimulus and 0.5 s after stimulus offset. These epochs were resampled to 250 Hz for offline analysis. More information about the data set can be referred to [28].

Data preprocessing was conducted before the spatial filtering for feature extraction. All data epochs were temporally

filtered in the frequency band 7-90 Hz with a Chebyshev Type I infinite impulse response (IIR) filter. Forward and backward filtering was performed to avoid phase distortion. Considering a latency delay in the visual system, the filtered epochs were extracted in [0.64 s, (0.64+$d$) s], where $d$ indicates the data length used for target recognition.

### B. TARGET RECOGNITION BASED ON CCA

CCA is a multivariable statistical technique whose goal is to examine the underlying correlation between two sets of random variables. It seeks two weight vectors to maximize the correlation between the two variables [30], [31]. Given two multidimensional variables $X$ and $Y$, their respective linear combinations can be denoted as $x = X^T W_x$ and $y = Y^T W_y$. CCA finds the two weight matrices $W_x$ and $W_y$ by maximizing the correlation $\rho$ between the two combinations as follows:

$$\begin{aligned}\rho &= \arg\max_{W_x, W_y} \frac{E[x^T y]}{\sqrt{E[x^T x]E[y^T y]}} \\ &= \arg\max_{W_x, W_y} \frac{E[W_x^T X Y^T W_y]}{\sqrt{E[W_x^T X X^T W_x]}\sqrt{E[W_y^T Y Y^T W_y]}}\end{aligned} \tag{1}$$

where $\rho$ is a one-dimensional vector whose elements are arranged in descending order and $E$ denotes the operation of mathematical expectation. The first columns of $W_x$ and $W_y$, denoted respectively as $w_x$ and $w_y$, corresponding to the maximal value in $\rho$, are used as spatial filters in SSVEP-based BCIs.

CCA has two input signals. In SSVEP-based BCI applications, the input $X$is a single-trial testing signal, whereas the input $Y$ is a reference signal that can be either a frequency specific sine-cosine reference signal (named SC CCA hereafter) or an individual template signal (named IT CCA hereafter) yielded by averaging training signals across multiple trials.

#### 1) SC CCA

For the target with stimulus frequency $f_n, n = 1, 2, \ldots, N_f$, the sine-cosine reference signals, $Y_n \in R^{2N_h \times N_s}$, are artificially created as follows [16], [17]:

$$Y_n = \begin{pmatrix} \sin(2\pi f_n t) \\ \cos(2\pi f_n t) \\ \cdots \\ \sin(2\pi N_h f_n t) \\ \cos(2\pi N_h f_n t) \end{pmatrix}, \quad t = \frac{0}{F_s}, \frac{1}{F_s}, \ldots, \frac{N_s - 1}{F_s} \tag{2}$$

where $N_h$ and $N_s$ denote the number of harmonics and the number of sampling points in a single trial respectively, and $F_s$ is the sampling rate of EEG signals.

In the SC CCA, the maximal correlation coefficients $\rho_n(1), n = 1, 2, \ldots, N_f$ are used as classification features for target recognition. $\rho_n(1)$ are calculated between a testing signal $X \in R^{N_c \times N_s}$ and each reference signal $Y_n, n = 1, 2, \ldots, N_f$. The frequency of the testing trial is decided as the frequency of the reference signal with the maximum correlation:

$$f_t = \max_f \rho_f(1), \quad f = f_1, f_2, \ldots, f_{N_f} \tag{3}$$

#### 2) IT CCA

In the IT CCA method [21], [22], the reference signals are individual templates yielded by averaging training signals across multiple trials from the same target, i.e. $\bar{\chi}_n = (1/N_t)\sum_{t=1}^{N_t} \chi_n^t$, where $N_t$ is the number of training trials. By replacing the sine-cosine reference signals $Y_n$ with individual templates $\bar{\chi}_n$, the procedure for target recognition is the same as that of SC CCA.

#### 3) EXTENDED CCA

The extended CCA (named EX CCA hereafter) is the combination of SC CCA and IT CCA [6], [22]. Different from the two CCA methods, which employ the maximal values of canonical correlation $\rho$ in (1) as classification features, the EX CCA utilizes the correlation coefficients between different projected signals. The combination method makes use of the following three spatial filters for computing projection signals: (1) $w_x(X, \bar{\chi}_n)$ between a testing signal $X$ and the individual template $\bar{\chi}_n$; (2) $w_x(X, Y_n)$ between a testing signal $X$ and the sine-cosine reference signals $Y_n$; (3) $w_{\bar{\chi}_n}(\bar{\chi}_n, Y_n)$ between the individual template $\bar{\chi}_n$ and the sine-cosine reference signals $Y_n$. A correlation vector $\hat{r}_n$ containing five correlation coefficients for a stimulus target/frequency $n$, $n = 1, 2, \ldots, N_f$ is calculated between different projection vectors as follows:

$$\hat{r}_n = \begin{bmatrix} r_{n,1} \\ r_{n,2} \\ r_{n,3} \\ r_{n,4} \\ r_{n,5} \end{bmatrix} = \begin{bmatrix} corr(X^T w_x(X, Y_n), Y_f^T w_y(X, Y_n) \\ corr(X^T w_x(X, \bar{\chi}_n), \bar{\chi}_n^T w_X(X, \bar{\chi}_n) \\ corr(X^T w_x(X, Y_n), \bar{\chi}_n^T w_X(X, Y_n) \\ corr(X^T w_{\bar{\chi}_n}(\bar{\chi}_n, Y_n), \bar{\chi}_n^T w_{\bar{\chi}_n}(\bar{\chi}_n, Y_n) \\ corr(\bar{\chi}_n^T w_x(X, \bar{\chi}_n), \bar{\chi}_n^T w_{\bar{\chi}_n}(X, \bar{\chi}_n) \end{bmatrix} \tag{4}$$

where $corr(a, b)$ denotes Pearson correlation coefficient between two vectors $a$ and $b$. Finally, the weighted correlation coefficient $r_n$ is used as feature for target recognition:

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

where the function $sign(\cdot)$ is used in order to retain discriminative information from negative correlation coefficients. Finally, target recognition for the testing trial is completed as follows:

$$T = \arg\max_n r_n, \quad n = 1, 2, \ldots, N_f \tag{6}$$

### C. TARGET RECOGNITION BASED ON MSFA

MSFA is a kind of signal processing method that is used for estimating a set of weight vectors of multichannel signals so that they can be integrated into a one-dimensional signal [23]. The weight vectors are found by solving an optimization problem that maximizes the SNR of a set of variables. Assume that the observed multichannel EEG signal

$X$ is derived from a set of source signals $S$ corrupted by additive noise $N$, i.e. $X = S + N$, $X, S, N \in R^{N_c \times N_s}$, where $N_c$ and $N_s$ are the number of channels and the number of sampling points respectively. Then the optimization problem is represented as:

$$\rho = \max_{w \neq 0} \frac{||Sw||_2}{||Nw||_2} = \max_{w \neq 0} \frac{w^T S S^T w}{w^T N N^T w} \tag{7}$$

where $w \in R^{N_c \times 1}$ is the weight vector and the superscript $T$ denotes transpose operation. The set of solutions to the above problem is defined recursively by constraining all solutions $w_i, i = 1, 2, \ldots, N_c$ to be orthogonal with respect to the weighted inner product $w_i^T S S^T w_j = 0$ for all $i \neq j$. The solutions to the optimization problem are the generalized eigenvectors [23] defined by

$$SS^T w = \lambda NN^T w \tag{8}$$

where $\lambda_i, i = 1, 2, \ldots, N_c$ is the eigenvalue corresponding to the eigenvector $w_i$. These eigenvalues and eigenvectors can be found by eigen decomposition of $SS^T$ and $NN^T$ using Matlab function eig, i.e. $[W, D] = eig(SS^T, NN^T)$, where $W = [w_1, w_2, \ldots, w_{N_c}]$ is a eigenvector matrix and $D = diag(\lambda_1, \lambda_2, \ldots, \lambda_{N_c})$ is a diagonal matrix of eigenvalues. Note that the eigenvalues in $D$ are arranged in ascending order.

In SSVEP-based BCIs, the weight vector $w_{N_c}$ corresponding to the maximum feature value in $D$ can be used as a spatial filter for a stimulus target. Assume that there are $N_t$ training trials for a specific stimulus target $n$, $\chi_n^t, t = 1, 2, \ldots, N_t$. The signal component of each training trial is considered invariant and can be obtained by trial averaging, $S_n^0 = \sum_{t=1}^{N_t} \chi_n^t$. Thereby, $\chi_n = [\chi_n^1, \chi_n^2, \ldots, \chi_n^{N_t}]$, $S_n = [S_n^0, S_n^0, \ldots, S_n^0]$, $N_n = \chi_n - S_n$. Replacing $S$ and $N$ in Equation (7) with $S_n$ and $N_n$, the spatial filter for target $n$ can be obtained. With the spatial filters for all stimulation targets, two projection vectors can be computed by spatially filtering the individual template for a target and a single-trial testing signal respectively. Thus, a feature value is yielded by calculating the Pearson correlation between the two projection vectors as:

$$r_n = corr(X^T w_n, \bar{\chi}_n^T w_n) \tag{9}$$

Finally, the target recognition is done by comparing the feature values from all targets and the target with maximal feature value is decided as the one the subject is gazing at using Equation (6).

### D. ENSEMBLE SPATIAL FILTER AND FILTER BANK ANALYSIS

Recently, two effective techniques were proposed for enhancing the performance of SSVEP-based BCIs, i.e. ensemble spatial filter [5], [32], [33] and filter bank [7], [8], [34] analysis. The former is a method that incorporates the spatial filters from all stimulation targets to increase the SNR of SSVEP signals, whereas the latter is a temporal method that decomposes an SSVEP signal into multiple sub-band signals and combines the useful information embedded in harmonic components for the same purpose.

#### 1) ENSEMBLE SPATIAL FILTER

Since a scalp recorded EEG signal is the mixture of multiple source signals, the spatial filters from all stimulus targets, each of which combines SSVEP signals from different channels, should be similar to each other within the same frequency bands [35], [36]. Thereby, integrating these spatial filters can significantly improve the performance of spatial filtering. An ensemble spatial filter is constructed as $W = [w_1, w_2, \ldots, w_{N_f}]$. Then Equation (9) for computing correlation coefficient for the $n$th stimulus target is modified as:

$$r_n = corr(X^T W, \bar{\chi}_n^T W) \tag{10}$$

#### 2) FILTER BANK ANALYSIS

According to the study [8], filter bank analysis decomposes the informative frequency band of 8∼88 Hz into 10 different sub-bands, with each sub-band ranging between $m \times 8$ Hz and 88 Hz. The sub-band signals are acquired by applying zero-phase infinite impulse response (IIR) filters of Chebyshev Type I. Assume that the template signals and a single-trial testing signal for the $m$th sub-band are denoted as $\bar{\chi}_n^m \in R^{N_f \times N_c \times N_s}$ and $X^m \in R^{N_c \times N_s}$ respectively. In the sub-band, the feature value derived from the $n$th target can be calculated as $r_n^m = f(\bar{\chi}_n^m, X^m)$, where $f$ denotes a kind of spatial filtering algorithm such as the extended CCA and MSFA. The weighted sum of squared feature values for all sub-bands is used for the classification feature and calculated as:

$$r_n = \sum_{m=1}^{N_m} a(m) \cdot (r_n^m)^2 \tag{11}$$

where $N_m$ is the number of sub-bands and $a(m) = m^{-1.25} + 0.25$ is the weight coefficient defined in [8]. Finally, the target corresponding to a single-trial testing signal is recognized using Equation (6).

#### 3) THE APPLICATION OF FILTER BANK AND ENSEMBLE SPATIAL FILTER

Because of their respective characteristics, the three CCA-based methods, SC CCA, IT CCA and EX CCA, can only be combined with filter bank analysis, and the resulting methods are named SC FBCCA, IT FBCCA and EX FBCCA respectively. MFSA, however, can be combined with either filter bank or ensemble analysis or both, the resulting methods are named FBMSFA, EN MSFA and EN FBMSFA respectively.

Flowchart of the proposed MSFA-based methods for target recognition is shown in Fig. 1. The flowchart is divided into training stage and testing stage. The task in the training stage is to estimate a spatial filter and a template signal for each target using training trials, whereas that in the testing stage is to classify a testing trial to a specific target/frequency

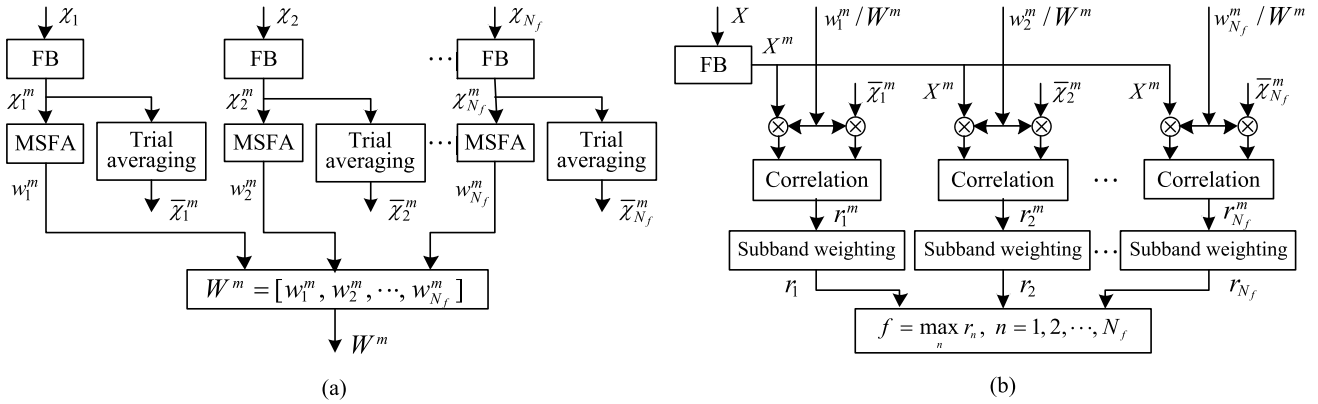


**FIGURE 1.** Flowchart of the proposed MSFA-based methods for target recognition: (a) Training stage: The task is to estimate a spatial filter and a template signal for each target using training trials; (b) Testing stage: The task is to classify a testing trial to a specific target using the templates and spatial filters of all targets or an ensemble spatial filter yielded in the training stage.

using the templates and spatial filters of all targets yielded in the training stage. This flowchart is a comprehensive one, which can include or exclude the filter bank and/or ensemble analysis. In the case that the ensemble analysis is not applied, only the spatial filters $w_n, n = 1, 2, \ldots, N_f$ are used in the testing stage. In the case that the filter bank analysis is not applied, only the first sub-band (8 Hz-88 Hz) is used in the two stages.

### E. PERFORMANCE COMPARISON

Different spatial filtering algorithms can be evaluated under four conditions, i.e. not using both filter bank and ensemble spatial filter analysis, using either of them and using both of them simultaneously. Since the extended CCA (EX CCA) is actually one kind of ensemble analysis method, a reasonable comparison of BCI performance should be made among MSFA, SC CCA and IT CCA, among FBMSFA, SC FBCCA and IT FBCCA, between EN MSFA and EX CCA, and between EN FBMSFA and EX FBCCA.

In this study, three evaluating indicators, i.e. recognition accuracy of targets, simulated ITR and r-square, are employed for comparing BCI performance. The first two were estimated by leave-one-out cross validation, in which 5 blocks were used as training data and 1 block was used as testing data. For the estimation of simulated ITR, the 0.5 s gaze-shifting time was included in the time for target selection. The third indicator was calculated with the feature value yielded by the target stimulus and the maximal feature value yielded by non-target stimuli [1], [21].

## III. RESULTS

The CCA and MSFA based methods have a total of five parameters, i.e. the length of data windows, the number of channels, the number of training trials, the number of harmonics and the number of sub-bands. Except when each parameter was targeted for performance analysis, they were fixed to 0.8 s, 9, 5, 5 and 5 respectively if applicable. The first parameter was selected in line with the study [4], whereas the latter two were determined according to the studies [4], [5]. EEG data from nine channels over occipital region (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, O2) were used for this study and the number of channels was selected in turn from the channel subset when it was used as a variable parameter.

First of all, we compared the BCI performance among MSFA, SC CCA and IT CCA, and among FB MSFA, SC FBCCA and IT FBCCA. The first triad of algorithms does not incorporate filter bank analysis, whereas the second does. Fig. 2 shows the averaged classification accuracy and simulated ITR across all subjects yielded by each of the six algorithms at nine data lengths, which ranged from 0.2 s to 1 s with the interval of 0.1 s. The results in each column were derived from the same one triad. From Fig. 2(a) and 2(b), it is observed that the accuracy from each algorithm increased monotonically with data lengths. From Fig. 2(c) and 2(d), it is observed that for the four algorithms that used training data, the ITR first increased and then decreased with data lengths, whereas for the two algorithms that did not use training data, the ITR increased monotonically. This suggests that the sine-cosine references based CCA methods need more data to achieve their highest ITRs. In terms of both accuracy and ITR, MSFA outperformed IT CCA and SC CCA, and FBMSFA outperformed IT FBCCA and SC FBCCA. One way repeated measures analysis of variance (ANOVA) revealed that there were significant differences in both accuracy and ITR among each triad of algorithms at all data lengths. The results of statistical analyses are reported in Table 1. Moreover, post hoc paired $t$-tests indicated that significant differences existed between any two algorithms in each triad at each data length with all $p$ values small than 0.001.

Fig. 3 shows averaged classification accuracy across all subjects yielded by the six algorithms at different number of channels. From the figure, it is clearly seen that as the number of channels increased, the classification accuracy behaved an upward trend. In the first column, MSFA performed much better than IT CCA and SC CCA, whereas in the second, FBMSFA performed much better than IT FBCCA and

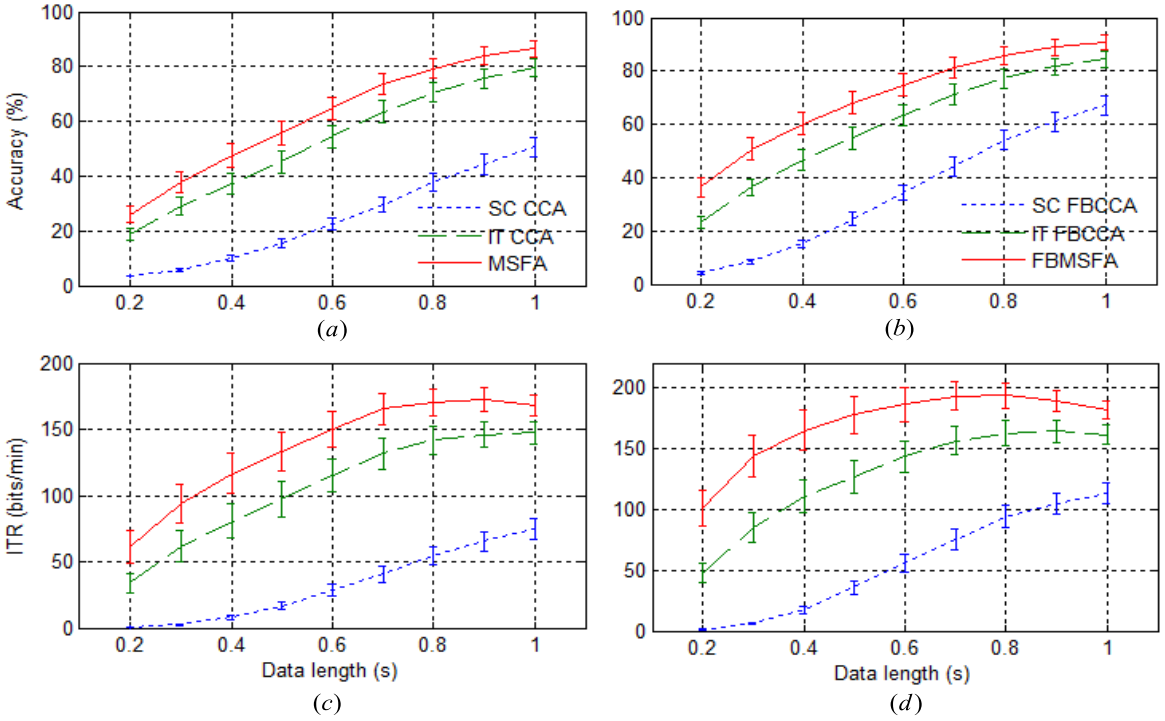


**FIGURE 2.** Averaged classification accuracy (a, b) and simulated ITR (c, d) across all subjects yielded by six spatial filtering algorithms at nine different data lengths. The error bars denote standard errors.

**TABLE 1.** Statistical analysis of significant differences in classification accuracy and simulated ITRs between SC CCA, IT CCA, and MSFA (Acc-1, ITR-1) and between SC FBCCA, IT FBCCA, and FBMSFA (Acc-2, ITR-2) at nine data lengths. *P* denotes the significance level of differences.

| Data length | | 0.2 s | 0.3 s | 0.4 s | 0.5 s | 0.6 s | 0.7 s | 0.8 s | 0.9 s | 1 s |
|---|---|---|---|---|---|---|---|---|---|---|
| Acc-1 | F(2, 68) | 52.66 | 76.48 | 110.12 | 127.12 | 208.84 | 284.27 | 298.39 | 251.05 | 185.8 |
| | $p$ | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 |
| ITR-1 | F(2, 68) | 24.47 | 37.7 | 55.62 | 70.6 | 114.79 | 173.37 | 234.87 | 255.14 | 224.63 |
| | $p$ | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 |
| Acc-2 | F(2, 68) | 85.39 | 124.22 | 179.6 | 225.79 | 278.21 | 264.89 | 174.03 | 145.75 | 100.28 |
| | $p$ | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 |
| ITR-2 | F(2, 68) | 43.91 | 68.69 | 103.42 | 139.79 | 188.08 | 249.2 | 211.31 | 193.8 | 135.44 |
| | $p$ | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 |

SC FBCCA. One way repeated-measures ANOVA showed that at each number of channels, there were significant differences in accuracy among the first/second triad of algorithms ($N_c = 3 : F(2, 68) = 81.56/88.14, p < 0.001$; $N_c = 4 : F(2, 68) = 169.22/124.84, p < 0.001$; $N_c = 5 : F(2, 68) = 167.15/115.53, p < 0.001$; $N_c = 6 : F(2, 68) = 182.48/113.36, p < 0.001$; $N_c = 7 : F(2, 68) = 194.73/104.36, p < 0.001$; $N_c = 8 : F(2, 68) = 191.34/88.11, p < 0.001$; $N_c = 9 : F(2, 68) = 184.15/93.53, p < 0.001$).

Fig. 3 shows averaged classification accuracy across all subjects yielded by the six algorithms at different number of channels. From the figure, it is clearly seen that as the number of channels increased, the classification accuracy behaved an upward trend. In the first column, MSFA performed much better than IT CCA and SC CCA, whereas in the second, FBMSFA performed much better than IT FBCCA and SC FBCCA. One way repeated-measures ANOVA showed that at each number of channels, there were significant differences in accuracy among the first/second triad of algorithms ($N_c = 3 : F(2, 68) = 81.56/88.14, p < 0.001$; $N_c = 4 : F(2, 68) = 169.22/124.84, p < 0.001$; $N_c = 5 : F(2, 68) = 167.15/115.53, p < 0.001$; $N_c = 6 : F(2, 68) = 182.48/113.36, p < 0.001$; $N_c = 7 : F(2, 68) = 194.73/104.36, p < 0.001$; $N_c = 8 : F(2, 68) = 191.34/88.11, p < 0.001$; $N_c = 9 : F(2, 68) = 184.15/93.53, p < 0.001$).

Fig. 4 illustrates averaged classification accuracy across all subjects yielded by the six spatial filtering algorithms at different numbers of training trials. For the two algorithms,

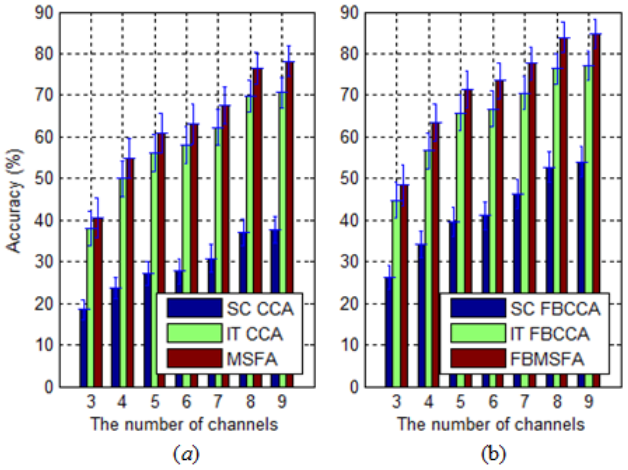


**FIGURE 3.** Averaged classification accuracy across all subjects yielded by the six spatial filtering algorithms at different number of channels. The error bars indicate standard errors.

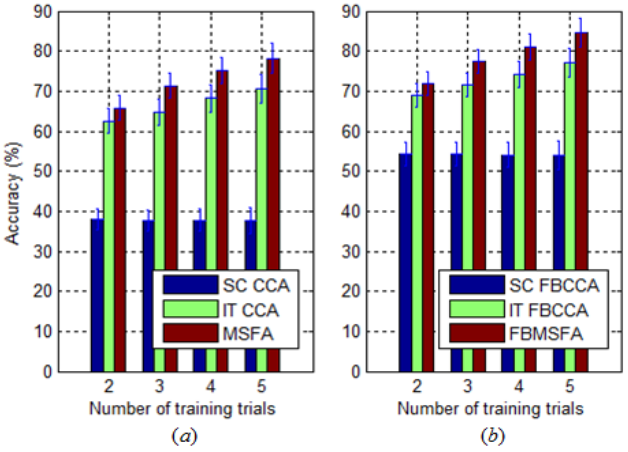


**FIGURE 4.** Averaged classification accuracy across all subjects yielded by the six spatial filtering algorithms at different number of training trials. The error bars indicate standard errors.

SC CCA and SC FBCCA, the classification accuracy remained unchanged regardless of the number of training trials because they did not use training data. For the other four algorithms, the classification accuracy increased monotonically with the number of training trials. At all numbers of training trials, MSFA outperformed IT CCA and SC CCA, whereas FBMSFA outperformed IT FBCCA and SC FBCCA. One way repeated-measures ANOVA exhibited that at each number of training trials, there were significant differences in accuracy between the first/second triad of algorithms ($N_t = 2 : F(2, 68) = 338.24/103.57, p < 0.001$; $N_t = 3 : F(2, 68) = 332.36/152.04, p < 0.001$; $N_t = 4 : F(2, 68) = 339.42/205.67, p < 0.001$; $N_t = 5 : F(2, 68) = 276.14/170.78, p < 0.001$).

r-squares were utilized to explore the discriminability of features for target recognition. Fig. 5 shows an example of the averaged $r$-square values across all subjects for SSVEPs at 10.2 Hz for the two triads of algorithms. One way repeated-measures ANOVA showed significant differences among the

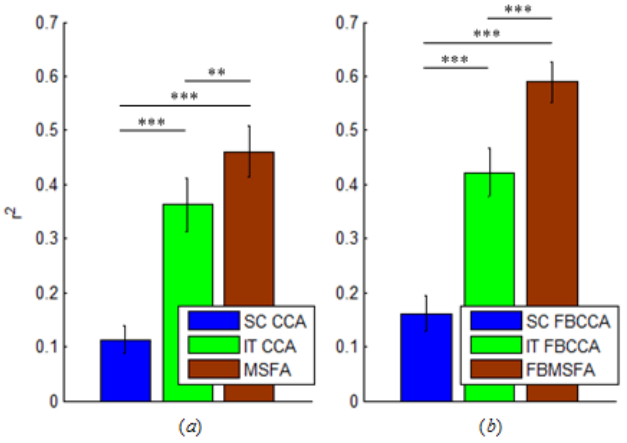


**FIGURE 5.** An example of averaged $r$-square value across all subjects for SSVEPs at 10.2 Hz obtained from each of the six algorithms. The error bars denote standard errors. The asterisks indicate the significant differences between two algorithms in each triad. ∗∗ indicates $p < 0.01$ and ∗ ∗ ∗ indicates $p < 0.001$.

first/second triad of algorithms ($F(2, 68) = 24.64/55.67$, $p < 0.001$). Post hoc paired $t$-tests showed that the r-square achieved by MSFA was significantly higher than those achieved by SC CCA and IT CCA, whereas that achieved by FBMSFA was significantly higher than those achieved by SC FBCCA and IT FBCCA. These results suggest that comparing the three algorithms in each triad, MSFA and FBMSFA could increase the distance of feature values between targets and non-targets.

Comparing results of the two columns in Fig. 2, Fig. 3, Fig. 4 and Fig. 5 respectively, the three algorithms in the second column outperformed those in the first column due to the incorporation of filter bank analysis. Paired $t$-tests showed that there were significant differences between the two triads of algorithms in accuracy and ITR at each data length with $p < 0.001$, in accuracy at each number of channels with $p < 0.001$, in accuracy at each number of training trials with $p < 0.001$, and in r-square with $p < 0.001$. All of these results fully illustrate that the filter bank analysis could enhance the discriminability of the feature signals used for target recognition.

Secondly, we compared the BCI performance between EN MSFA and EX CCA, and between EX FBMSFA and EX FBCCA. Different from the above-mentioned six algorithms, all the four algorithms incorporate the analysis of ensemble spatial filters. Fig. 6 shows averaged classification accuracy and simulated ITR across all subjects yielded at different data lengths. The classification accuracy yielded by each algorithm in the first row increased monotonically with data lengths. The ITR yielded by each algorithm in the second row, however, first increased and then decreased. The results yielded by EN MSFA and EN FBMSFA were much better than those yielded by EX CCA and EX FBCCA respectively, especially for short data lengths. Paired $t$-tests revealed that at each data length, there were significant differences in accuracy and ITR between each pair of algorithms with all $p$ values smaller than 0.001.

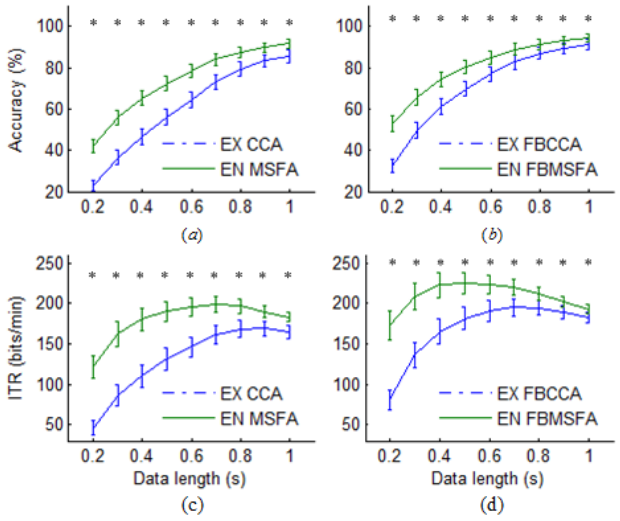


**FIGURE 6.** The averaged classification accuracy (a, b) and simulated ITR (c, d) across all subjects yielded by the four spatial filtering algorithms at nine different data lengths. ∗ indicates the significant difference between each pair of algorithms by paired $t$-test ($p < 0.001$). The error bars denote standard errors.

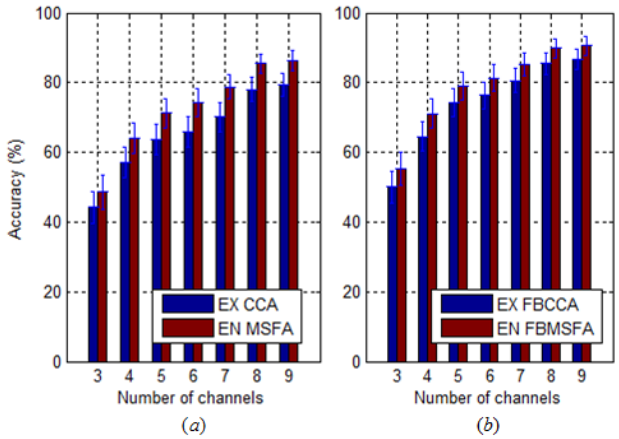


**FIGURE 7.** Averaged classification accuracy across all subjects yielded by the four spatial filtering algorithms at different number of channels. The error bars indicate standard errors.

Fig. 7 illustrates averaged classification accuracy across all subjects yielded by the four spatial filtering algorithms at different number of channels. Clearly, the classification accuracies of all the algorithms increased consistently with the number of channels. However, EN MSFA and EN FBMSFA were superior to EX CCA and EX FBCCA respectively. Paired t-test indicated that significant differences in accuracy existed between EN MSFA and EX CCA, and between EN FBMSFA and EX FBCCA at each number of channels with all $p$ values smaller than 0.001.

Fig. 8 shows averaged classification accuracy across all subjects yielded by the four spatial filtering algorithms at different number of training trials. Likewise, EN MSFA and ENFBMSFA outperformed EX CCA and EX FBCCA respectively at each number of training trials. Pared $t$-tests revealed that there were significant differences in accuracy between

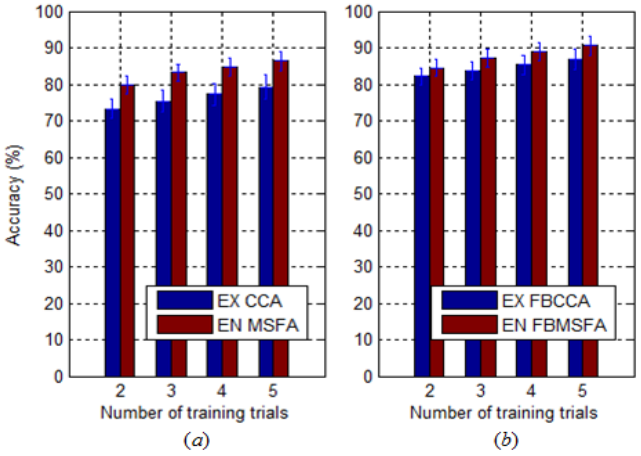


**FIGURE 8.** Averaged classification accuracy across all subjects yielded by the four spatial filtering algorithms at different number of training trials. The error bars indicate standard errors.

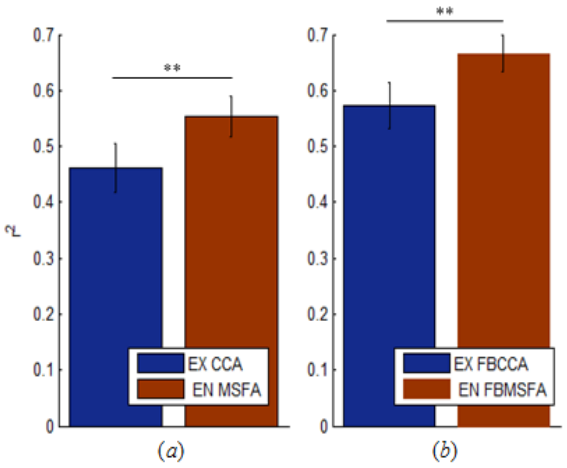


**FIGURE 9.** An example of the averaged $r$-square values for SSVEPs at 10.2 Hz derived from each of the four spatial filtering algorithms. The error bars denote standard errors. The asterisks indicate the significant differences between each pair of algorithms. ∗∗ indicates $p < 0.01$.

each of two pairs of algorithms at each number of training trials with all $p$ values smaller than or equal to 0.001.

Fig. 9 depicts an example of the averaged $r$-square values for SSVEPs at 10.2 Hz for the four spatial filtering algorithms. Paired $t$-tests showed that the r-squares achieved by EN MSFA and EN FBMSFA were significantly higher than those achieved by EX CCA and EX FBCCA respectively. These results suggested that EN MSFA and EN FBMSFA could increase the distance of feature values between targets and non-targets compared to EX CCA and EX FBCCA respectively.

Comparing results of the two columns in Fig. 6, Fig. 7, Fig. 8 and Fig. 9 respectively, it is clearly seen that the two algorithms in the second column were superior to those in the first column due to the incorporation of filter bank analysis. Paired $t$-tests showed that there were significant differences between the two pairs of algorithms in accuracy and ITR at each data length with $p < 0.001$, in accuracy at each number of channels with $p < 0.001$, in accuracy at each

number of training trials with $p < 0.001$, and in r-square with $p < 0.001$. All of these results fully illustrate that the combination of ensemble spatial filter and filter bank analysis is more effective than the use of these two analyses alone.

## IV. DISCUSSIONS

Three parameters affecting the performance of an SSVEP-based BCI are data length, the number of channels and the number of training trials, which are determined to a large extent by the spatial filtering algorithm used for target recognition. These three parameters determine the communication speed of a BCI system, its convenience of use and visual fatigue of users. The data length required for achieving the highest ITR is different for different algorithms. It is shown in Fig. 2 that for the first triad of algorithms, MSFA achieved the highest ITR at a data length of 0.9 s, whereas SC CCA and IT CCA did not achieve their highest ITRs in the time range; For the second triad of algorithms, FB MSFA and IT MSFA yielded the highest ITRs at a data length of 0.8 s and 0.9 s respectively, whereas SC FBCCA did not yield the highest ITR in the time range. It is shown in Fig. 6 that for the first pair of algorithms, EN MSFA and EX CCA reached the highest ITRs at a data length of 0.7 s and 0.9 s respectively; for the second pair of algorithms, EN FBMSFA and EX FBCCA reached the highest ITRs at a data length of 0.4 s and 0.7 s respectively. Thereby, the MSFA-based methods could use shorter data length for target recognition than the CCA-based methods under the same conditions. By replacing the latter with the former, the communication speed of BCIs can be improved substantially.

Using the same number of channels, classification accuracies achieved by different algorithms were different. Fig. 3 shows that at each number of channels, the accuracy yielded by MSFA was much higher than that yielded by SC CCA and IT CCA, whereas the accuracy yielded by FBMSFA was much higher than that yielded by SC FBCCA and IT FBCCA. Fig. 7 shows that at each number of channels, the accuracy yielded by EN MSFA was much higher than that yielded by EX CCA, whereas the accuracy yielded by EN FBMSFA was much higher than that yielded by EX FBCCA. Thereby, the two MSFA-based methods could achieve the identical accuracy with fewer channels than the two CCA-based methods. By replacing the latter with the former, the convenience of BCI use can be improved accordingly.

Using the same number of training trials, classification accuracies achieved by different algorithms were also different. Fig. 4 shows that at each number of training trials, the accuracy yielded by MSFA was much higher than that yielded by SC CCA and IT CCA, whereas the accuracy yielded by FBMSFA was much higher than that yielded by SC FBCCA and IT FBCCA. Fig. 8 shows that at each number of training trials, the accuracy yielded by EN MSFA was much higher than that yielded by EX CCA, whereas the accuracy yielded by EN FBMSFA was much higher that yielded by EX FBCCA. Thereby, the two MSFA-based methods could achieve the identical accuracy with fewer training trials than the two CCA-based methods. By replacing the latter with the former, the fatigue of BCI users can be decreased considerably.

So far, SSVEP-based BCIs do not yet support widespread usage in spite of great advances. The main factors to limit their practical applications may be attributed to low communication speed, inconvenient to use and brain fatigue, all of which are associated with the spatial filter used for feature extraction. CCA is a successful spatial filtering algorithm widely used in SSVEP-based BCIs. Compared to CCA, MSFA has more robust anti-noise performance and its efficiency has been fully verified in this study. As a result, substituting CCA with MSFA can enhance the performance of SSVEP-based BCIs.

## V. CONCLUSION

A novel target detection method using MSFA for spatial filtering was proposed to enhance the performance of SSVEP-based BCIs in this study. MSFA can learn spatial filters from training data by effectively removing background noise and other artifacts. With a benchmark data set from 35 subjects, the MSFA-based methods were comprehensively compared with the CCA-based methods under various conditions. The results demonstrated that MSFA achieved superior classification performance and high robustness to noise, contributing to the practical applications of SSVEP-based BCIs.

## REFERENCES

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtschller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, Jun. 2002.
[2] S. Gao, Y. Wang, X. Gao, and B. Hong, "Visual and auditory brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1436–1447, May 2014.
[3] C. Yang, X. Han, Y. Wang, R. Saab, S. Gao, and X. Gao, "A dynamic window recognition algorithm for SSVEP-based brain-computer interfaces using a spatio-temporal equalizer," *Int. J. Neural Syst.*, vol. c, no. 10, 2018, Art. no. 1850028.
[4] Y. Zhang, E. Yin, F. Li, Y. Zhang, T. Tanaka, Q. Zhao, Y. Cui, P. Xu, D. Yao, and D. Guo, "Two-stage frequency recognition method based on correlated component analysis for SSVEP-based BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 7, pp. 1314–1323, Jul. 2018. doi: 10.1109/TNSRE.2018.2848222.
[5] 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.
[6] M. Nakanishi, Y. Wang, Y. Wang, Y. Mitsukura, and T. Jung, "A high-speed brain speller using steady-state visual evoked potentials," *Int. J. Neural Syst.*, vol. 24, no. 6, p. 1450019, 2014.
[7] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T. P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain–computer interface," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 44, pp. E6058–E6067, Nov. 2015.
[8] 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, 2015, Art. no. 046008sss.
[9] E. Yin, Z. Zhou, J. Jiang, F. Chen, Y. Liu, and D. Hu, "A novel hybrid BCI speller based on the incorporation of SSVEP into the P300 paradigm," *J. Neural Eng.*, vol. 10, no. 2, 2013, Art. no. 026012.
[10] Q. Wei, Y. Liu, X. Gao, Y. Wang, C. Yang, Z. Lu, and H. Gong, "A novel c-VEP BCI paradigm for increasing the number of stimulus targets based on grouping modulation with different codes," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 6, pp. 1178–1187, Jun. 2018.

[11] Y. Liu, Q. Wei, and Z. Lu, "A multi-target brain-computer interface based on code modulated visual evoked potentials," *Plos One*, vol. 13, no. 8, 2018, Art. no. e0202478.
[12] 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, 2011, Art. no. 025015.
[13] M. Chen, 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.
[14] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical vep-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 234–239, Feb. 2006.
[15] O. Friman, I. Volosyak, and A. Graser, "Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 742–750, Apr. 2007.
[16] 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.
[17] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method," *J. Neural Eng.*, vol. 6, no. 4, 2009, Art. no. 046002.
[18] Y. Zhang, G. Zhou, Q. Zhao, A. Onishi, and A. Cichocki, "Multiway canonical correlation analysis for frequency components recognition in SSVEP-based BCIs," in *Proc. 18th Int. Conf. Neural Inform. Process.*, Nov. 2011, pp. 287–295.
[19] 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.
[20] 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.
[21] 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 Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Chicago, IL, USA, Aug. 2014, pp. 3037–3040.
[22] 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, 2015, Art. no. e140703.
[23] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 1, pp. 65–74, Jan. 1988.
[24] D. R. Hundley, M. J. Kirby, and M. Anderle, "Blind source separation using the maximum signal fraction approach," *Signal Process.*, vol. 82, no. 10, pp. 1505–1508, Oct. 2002.
[25] X. Lu, X. Li, M.-S. Fu, and H. Wang, "Robust maximum signal fraction analysis for blind source separation," *IET Signal Process.*, vol. 11, no. 8, pp. 969–974, Oct. 2017.
[26] H. Wang, "Temporally local maximum signal fraction analysis for artifact removal from biomedical signals," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4919–4925, Sep. 2010.
[27] C. W. Anderson, J. N. Knight, T. O'Connor, M.J. Kirby, and A. Sokolov, "Geometric subspace methods and time-delay embedding for EEG artifact removal and classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 142–146, Jun. 2006.
[28] Y. Wang, X. Chen, X. Gao, and S. Gao, "A benchmark dataset for SSVEP-based brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1746–1752, Oct. 2017.
[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 Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Chicago, IL, USA, Aug. 2014, pp. 3993–3996.
[30] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, nos. 3–4, pp. 321–377, 1936.
[31] O. Friman, J. Cedefamn, P. Lundberg, M. Borga, and H. Knutsson, "Detection of neural activity in functional MRI using canonical correlation analysis," *Magn. Reson. Imag.*, vol. 45, pp. 323–330, 2001.
[32] 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.
[33] H. Takana, T. Hiroki, and H. Sato, "Task-related oxygenation and cerebral blood volume changes estimated from NIRS signals in motor and cognitive tasks," *NeuroImage*, vol. 94, pp. 107–119, Jul. 2014.
[34] R. Islam, K. I. Molla, M. Nakanishi, and T. Tanaka, "Unsupervised frequency-recognition method of SSVEPs using a filter bank implementation of binary subband CCA," *J. Neural Eng.*, vol. 14, no. 2, 2017, Art. no. 026007.
[35] 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 Topography*, vol. 18, no. 3, pp. 167–187, Mar. 2006.
[36] 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, Jul. 2009.