An Idle-State Detection Algorithm for SSVEP-Based Brain–Computer Interfaces Using a Maximum Evoked Response Spatial Filter

Dan Zhang  
Department of Psychology, Tsinghua University  
Room 334, Ming Zhai Building  
Beijing 100084, P. R. China  
dzhang@tsinghua.edu.cn

Bisheng Huang  
Department of Computer Science and Technology  
Tsinghua University, Beijing 100084, P. R. China

Wei Wu  
School of Automation Science and Engineering  
South China University of Technology, No. 3 Building  
Wushan Road, Tianhe District, Guang Zou City, P. R. China

Department of Psychiatry and Behavioral Sciences  
Stanford University, Palo Alto, CA 94304, USA

Siliang Li  
Department of Computer Science and Technology  
Tsinghua University, Beijing 100084, P. R. China

Accepted 10 July 2015  
Published Online 5 August 2015

Although accurate recognition of the idle state is essential for the application of brain–computer interfaces (BCIs) in real-world situations, it remains a challenging task due to the variability of the idle state. In this study, a novel algorithm was proposed for the idle state detection in a steady-state visual evoked potential (SSVEP)-based BCI. The proposed algorithm aims to solve the idle state detection problem by constructing a better model of the control states. For feature extraction, a maximum evoked response (MER) spatial filter was developed to extract neurophysiologically plausible SSVEP responses, by finding the combination of multi-channel electroencephalogram (EEG) signals that maximized the evoked responses while suppressing the unrelated background EEGs. The extracted SSVEP responses at the frequencies of both the attended and the unattended stimuli were then used to form feature vectors and a series of binary classifiers for recognition of each control state and the idle state were constructed.

EEG data from nine subjects in a three-target SSVEP BCI experiment with a variety of idle state conditions were used to evaluate the proposed algorithm. Compared to the most popular canonical correlation analysis-based algorithm and the conventional power spectrum-based algorithm, the proposed algorithm outperformed them by achieving an offline control state classification accuracy of 88.0 ± 11.1% and idle state false positive rates (FPRs) ranging from 7.4 ± 5.6% to 14.2 ± 10.1%, depending on the specific idle state conditions. Moreover, the online simulation reported BCI performance close to practical use: 22.0 ± 2.9 out of the 24 control commands were correctly recognized and the FPRs achieved as low as approximately 0.5 event/min in the idle state conditions with eye open and 0.05 event/min in the idle state condition with eye closed. These results demonstrate the potential of the proposed algorithm for implementing practical SSVEP BCI systems.

Keywords: Steady-state visual evoked potentials; idle state detection; spatial filtering; maximum evoked response.
1. Introduction

The steady-state visual evoked potential (SSVEP)-based brain–computer interface (BCI) is one of the most popular BCI paradigms. A typical SSVEP BCI operates on the attentional modulation of the SSVEPs: the attended flickering stimulus elicits enhanced SSVEP responses at the corresponding frequency over occipital brain regions; recognition of the attention target is then achieved by analyzing the spectrum of the recorded electroencephalograms (EEGs). Although SSVEP BCI is known for its high classification accuracy and high information transfer rate (ITR), most studies have reported only the performances under control conditions, i.e. when the subjects were actively operating the BCI system for command outputs. In a practical BCI scenario, however, accurate detection of the idle state is crucial for good user experience: when the subjects are relaxing, reading books, watching TVs, etc. (i.e. in the idle state), frequent unintended command outputs may not only annoy the user but also put the user at risk of being in danger in certain circumstances.

The idle state is difficult to model, as it may include a variety of different mental activities. Hereby, discrimination between the control state and the idle state is usually achieved by modeling the control state only: Efforts have been devoted to designing spatial filtering technique that could effectively extract the SSVEP responses from the multi-channel EEG data. To date, the most popular spatial filtering technique is canonical correlation analysis (CCA), which seeks a spatially weighted combination of the multichannel EEG data that maximizes the multivariate correlation between the brain activities and a group of ideal sine and cosine signals at the stimulation frequency and its harmonics. Since the stimulation frequencies are explicitly used for deriving the spatial filters, CCA could be considered as a ‘supervised’ method. The extracted SSVEP responses can then be compared to an empirically established threshold for idle state detection: a below-threshold response indicates that the subject is not paying sufficient attention to the corresponding steady-state flicker, i.e. the idle state. If multiple steady-state frequencies are present, the frequency with the strongest above-threshold response is regarded as the user intention.

Recently, increasing interest has been drawn toward exploring neurophysiologically relevant information for BCI classifications. One study on motor imagery BCIs employed independent component analysis (ICA) to obtain motor-related spatial filters that could effectively extract the motor imagery related mu rhythm source. Although ICA is an ‘unsupervised’ method, BCI results comparable to the ‘supervised’ common spatial patterns algorithm were obtained. In an event related potential (ERP)-based visual motion speller (the N200 speller) study, it has been suggested that idle state performance can be enhanced by constructing a feature space including not only the attention target related ERPs but also the ERPs elicited by the nontarget stimuli. The nontarget stimuli were assumed to have weaker ERPs that decreased as a function of its spatial distance to the attention target, known as the eccentricity effect. The combination of both target and nontarget ERPs provided a better model of the control state, thus allowing a better discrimination of the idle state. In the field of SSVEP BCIs, it has been shown that the performance of idle state detection can be further improved by constructing idle-specific algorithmic modules based on measures of brain activities beyond SSVEPs, i.e. the spectral complexity, the approximate fractal entropy, the low frequency EEG power, etc.

In the present study, we proposed a new algorithmic framework for implementing a practical SSVEP BCI system, with a special emphasis on the idle state detection. The contribution of the proposed algorithm is two-folded: first, in contrast to the conventional approaches taking only the frequency with the maximal SSVEP response (as extracted by CCA), we constructed a feature space including SSVEP responses at all stimulation frequencies; second and more importantly, a novel spatial filtering technique named maximum evoked response (MER) spatial filter was introduced. Compared to the ‘supervised’ CCA method, the MER spatial filter is an ‘unsupervised’ method that is regarded to be more neurophysiologically oriented. To evaluate the proposed algorithm, an SSVEP BCI experiment including the control condition and a variety of idle state conditions was conducted. The performance of the proposed algorithm was also compared with the conventional CCA method as well as the...
classical power spectrum-based method without spatial filtering.

2. Experimental Setup

2.1. Subjects and data recording

Nine subjects (four females, undergraduate students from Tsinghua University), aged from 20 to 21 years old, participated as paid volunteers. All of them showed normal or corrected to normal eyesight. All subjects were naïve to both BCI experiments and steady-state visual stimulation, and they gave informed consent.

Electroencephalographs were recorded from 15 electrodes (Fz, F3/4, C3/4, T3/4, T5/6, Pz, P3/4, Oz, O1/2) according to the 10–20 systems, with forehead ground at Fz and reference at Cz. A portable EEG amplifier (NeuSen.Uamp, Neuracle, China) was used for data recording at a sampling rate of 1000 Hz. Electrode impedances were kept below 10 kΩ for all electrodes.

2.2. Stimulation paradigms

The stimuli were displayed on an LCD monitor (22-inch, DELL, USA) with 60 Hz refreshing rate, as illustrated in Fig. 1(a). Subjects were seated about 60 cm from the monitor. Three flickering white squares and a non-flickering gray square were presented on the black screen, serving as four possible attention targets. All squares subtended 3.8 × 3.8° of visual angle and their center-to-center distances were 6.2°. The flickering frequencies of the three squares were set to 12, 15 and 20 Hz (5, 4 and 3 frames per cycle, with duty cycles of 40%, 50% and 67%, respectively). The squares switched their colors between white (on) and black (off). The nonflickering square, however, was kept static with a gray color (RGB: 10, 10, 10) to be visible on the black screen.

2.3. Experimental procedure

The subjects participated in four sessions. During the first session (for training the models), the subjects were cued to overtly attend to one of the four squares in a trial-by-trial manner. At the beginning of each trial, all the squares were kept static for 1 s before flickering. The to-be-attended square was cued by a yellow arrow that was presented throughout the trial. Each trial lasted for 8 s, including the first 1 s for cuing and the remaining 7 s with the flickering stimulation. The inter-trial intervals were kept constant at 4 s. Each square was attended 8 times in a random order, summing up to 32 trials. The whole session lasted about 6.3 min. During the condition when the subjects were cued to the nonflickering square, no SSVEP would be elicited. Hereby this condition was considered as one idle state with subjects’ eyes still fixating on the computer screen, termed on-screen idle state.

To mimic the other commonly encountered real-world idling situations, the subjects were instructed to read a book and relax with their eyes closed, during the second and third session, respectively. The subjects were allowed to move their head and body, and they had conversations with the experimenter from time to time. The duration of the two sessions were set to the same as the first session. For these two sessions, the subjects were in idle states with their eyes off the screen. Here these two conditions were termed off-screen idle state and eye-closed idle state, respectively. The fourth session was identical to the first session, the subjects executed the overt attention task for 32 trials (24 control state trials and eight on-screen idle state trials). The procedure was summarized in Fig. 1(b).

The experiment was carried out in a normal office environment at a normal illuminance level. No feedback was given throughout the
The Idle State Detection Algorithm

3.1. The MER spatial filter

The goal of the MER method is to learn a spatial filter that maximized the brain responses evoked by the steady-state visual stimulation while suppressing the background activity. Here we argue that the optimal representation of the evoked responses, however, should not be the combination of ideal sine and cosine waveforms at the stimulation frequencies, which are explicitly used as the reference signals in the CCA method. The use of ideal sine or cosine waveforms as the reference signals for the spatial filter optimization relies upon the assumption that the frequency responses at the stimulation frequency over all EEG channels are relevant to SSVEPs. This assumption is especially problematic when the subjects have their eyes closed and the steady-state stimulations are within the alpha range: the increase of alpha power by eye closing may lead CCA to erroneously assign large spatial filter weights to the channels with large alpha power. To resolve this issue, we propose that the multi-channel recorded brain activity temporally averaged across trials is a better representation of the SSVEPs per se. By using such a temporal averaging criterion, the contribution of the channels outside the parietal-occipital regions is expected to be significantly reduced, as they only contain the spontaneous EEG activities that are not phase-locked to the stimulations and hence are minimized through temporal averaging. Therefore, the computation of the MER spatial filter method can be implemented as follows.

Denote the single-epoch multichannel EEG data by $X_i$ with a dimension of $N$ (channels) by $T$ (samples), where $i$ is the epoch index. As discussed above, the representation of the SSVEPs at stimulation frequency $f$ is the temporal average of the EEGs in the epochs when the subject is attending to the square flickering at frequency $f$:

$$Y = \frac{1}{M} \sum_{i} X_i, \quad (1)$$

where $M$ denotes the number of epochs to be averaged. Each epoch data $X_i$ can then be represented as the summation of the ‘evoked response’ and random background noise:

$$X_i = Y + n_i, \quad (2)$$

MER seeks a spatial filter weight $w$ that maximizes the representation of SSVEP component $w^T Y$ and minimizes the representation of the noise component $w^T n_i$. This problem can be solved by transforming the data into the covariance space:

$$C_X = C_Y + C_n, \quad (3)$$

where $C$ indicates the covariance matrix. $C_X$ and $C_Y$ are computed by

$$C_X = \frac{1}{N} \sum_{i} X_iX_i^T, \quad (4)$$

$$C_Y = YY^T. \quad (7)$$

The covariance of the single-epoch data can be factorized into the product of eigenvectors and eigenvalues via eigenvalue decomposition

$$C_X = U_0 \sum \Sigma_0 U_0^T, \quad (5)$$

where $U_0$ is the matrix of eigenvectors and $\Sigma$ is the diagonal matrix of the eigenvalues. Thus the covariance matrix $C_X$ can be transformed with matrix $P$ into an identity matrix:

$$P = \sum U_0^T, \quad (6)$$

$$P \cdot C_X \cdot P^T = I_N. \quad \text{The same transformation can be applied to the right-hand side of Eq. (3):}$$

$$I_N = P \cdot C_X \cdot P^T = P \cdot C_Y \cdot P^T + P \cdot C_n \cdot P^T = S_Y + S_n. \quad (7)$$

The transformed matrices $S_Y$ and $S_n$ can also be factorized into their eigenvectors and eigenvalues. One important property of $S_Y$ and $S_n$ is that they share common eigenvectors

$$S_Y = U \cdot \sum_{Y} U^T, \quad (8)$$

$$S_n = U \cdot \sum_{n} U^T. \quad (9)$$

Therefore, $\sum_{Y} = \sum_{n} = I_N$. If the eigenvalues in $\Sigma_Y$ is arranged in the descending order, then the first few eigenvectors correspond to...
the spatial filters that maximize the representation of the reference signal $Y$ in single-epoch data $X$.

Alternatively, the above linear transformation can be obtained via a generalized eigenvalue decomposition based on solving the following optimization problem:

$$\max_{w} w^T C_Y w \quad \text{s.t.} \quad C_X w = \lambda C_Y w.$$  \hspace{1cm} (10)

The number of spatial filters to be retained is then decided by the contribution of the first few eigenvalues. In the present study, only the first spatial filter, corresponding to the largest eigenvalue, was employed since it accounted for over 30% of the total power on average and no significant improvement of the BCI classification performance was obtained by including additional spatial filters.

To extract SSVEP responses at different stimulation frequencies, the EEG recordings from the first session were categorized into different attention conditions. Only the data when the subjects were attending to frequency $f$ were used for deriving the MER spatial filter at frequency $f$, resulting in eight trials (7 s per trial) per condition. To increase the power of temporal averaging, the trial-structured data were further segmented into 1 s epochs, leading to 48 epochs per condition (6 s per trial used; the first second of all trials were considered as the preparation time for SSVEP to fully build up, hereby excluded). Note that the stimulation frequency information was not required for deriving the MER spatial filters.

3.2. Modeling the idle state

The flowchart of the proposed algorithm is illustrated in Fig. 2(a). Three MER spatial filters corresponding to the three stimulation frequencies were obtained using data under the three attention conditions (i.e., attending to 12/15/20 Hz flickering squares), respectively. The epoched data were first filtered simultaneously by the three spatial filters. To get a complete view of the response pattern, here the three spatially filtered data were subject to spectral analyses and the spectral responses at the three

Fig. 2. (a) Flow diagram of the proposed MER algorithm. (b) Schema of the CCA algorithm for comparison. (c) Schema of the PS algorithm for comparison.
stimulation frequencies were extracted to constitute a nine-dimensional feature vector (three spatial filters × three stimulation frequencies). Three binary support vector machine (SVM) classifiers (12 Hz versus idle, 15 Hz versus idle, 20 Hz versus idle) were trained by using the data epochs when the subjects were attending to one of the flickering squares (control state) versus the condition when they were attending to the nonflickering square (on-screen idle state). The SVM classifiers were trained for posterior probability estimation and these probability values (the probability of a given epoch to be in one specific control state) were used as outputs. The idle state is identified when the maximum output of the three classifiers is below a certain threshold \( p_{th} \) and one specific control state is identified when its corresponding classifier yields the largest and above-threshold probability values. Here a probability value of 0.5 was used for the reported classification results.

Output

\[
\begin{align*}
\text{Output} & = \left\{ \begin{array}{ll}
\text{idle state, } & \max(p_1, p_2, p_3) \leq p_{th}, \\
\text{control state, } & p_i = \max(p_1, p_2, p_3) > p_{th}.
\end{array} \right.
\end{align*}
\] (11)

All data analyses were carried out in Matlab (the Mathworks, US) and the SVM classifiers were implemented using libSVM. A linear kernel was employed for SVMs and a five-fold cross-validation procedure was employed to determine the hyperparameters of the SVMs.

3.3. Evaluation of the proposed algorithm

To evaluate the performance of the proposed algorithm, two algorithms were additionally implemented for comparison: a CCA-based algorithm and a power spectrum (PS)-based algorithm.

The CCA algorithm was selected, as it is the most popular spatial filtering technique used in SSVEP BCIs. The flowchart of the CCA algorithm is shown in Fig. 2(b). Unlike the MER method that uses all training epochs to derive the spatial filters, the CCA spatial filters were generated for each individual epoch. In the original proposal of the CCA algorithm, the spectral responses of one given EEG data epoch at the three stimulation frequencies were reflected by the CCA correlation coefficients between the data epoch and three reference signals, respectively. The three reference signals were generated as below, including both the fundamental frequency and its second harmonics:

\[
\text{REF}_f(t) = \begin{bmatrix}
\sin 2\pi f_t \\
\cos 2\pi f_t \\
\sin 2\pi(2f)t \\
\cos 2\pi(2f)t
\end{bmatrix}
\] (12)

The recognition of one specific control state is then decided by taking the frequency with the largest CCA correlation coefficient (for simplicity, only the first CCA coefficient per frequency was considered). However, as the idle state detection was not included in those studies, here we made a simple and straightforward extension to enable a close match to the proposed MER algorithm: the spectral responses at the three stimulation frequencies were extracted from the CCA spatially filtered data, by the three CCA spatial filters respectively. Therefore, a nine-dimensional feature vector (three spatial filters × three stimulation frequencies) was formed to train three binary SVM classifiers (12 Hz versus idle, 15 Hz versus idle, 20 Hz versus idle) in a similar way as the proposed MER-based algorithm. Likewise, the classification was carried out using the strategy given in Eq. (11).

The PS algorithm was employed to represent the baseline performance without applying any spatial filters. Here the spectral responses at the three stimulation frequencies were extracted from all EEG channels, thus forming a 45-dimensional feature vector (3 frequencies × 15 channels). The classification procedure was similar to other two algorithms, as shown in Fig. 2(c).

Before calculating the BCI performance, we compared the spatial preference of the MER spatial filters and the CCA spatial filters. To this aim, the MER spatial filter weights were transformed to their absolute values (large positive and negative values indicate similar importance of the electrode) and then rescaled to have a maximal value of 1. In this way, an average value (across subjects) close to 1 indicates that the corresponding electrode is deemed important by the MER method. Although the CCA spatial weights in the conventional CCA BCI algorithm were obtained for each epoch separately, we took the spatial filter weights averaged over all epochs with the same attentional tasks to represent the weighting of different electrodes by the CCA algorithm. The averaged CCA filter weights
were then processed in the same way as the MER weights. As the MER spatial filters were obtained by maximizing evoked responses, we hypothesized that the MER spatial filter weights would be more similar to the original SSVEP topographies than the CCA filter weights. The most prominent spatial filters, i.e. the 1st MER spatial filters and the 1st CCA spatial filters per stimulation frequency, were extracted for comparison.

For all the algorithms, the data from the first session was used for both learning the spatial filters and training the classifiers. Offline classifications were performed on the data from the following three sessions. For the second session (off-screen idle state) and the third session (eye-closed idle state), false positive rate (FPR) was used to assess the idle state detection performance. For the fourth session (control state and on-screen idle state), both classification accuracy during the control state and FPR during the idle state were provided. The EEG recordings data were segmented into 1 s epochs for training and testing. According to the attentional task of the subjects, the epochs were categorized into control epochs, on-screen idle epochs, off-screen idle epochs and eye-closed idle epochs. Note that the classification results during the fourth session were used to represent the control state classification performances instead of cross-validated results within the subjects, the epochs were categorized into control epochs and on-screen idle epochs. According to the attentional task of the subjects, the epochs were categorized into control epochs, on-screen idle epochs, off-screen idle epochs and eye-closed idle epochs. Note that the classification results during the fourth session were used to represent the control state classification performances instead of cross-validated results within the subjects, the epochs were categorized into control epochs and on-screen idle epochs.

The possible online performances during the control state were also assessed by calculating the ITR using the following equation:

\[
\text{ITR} = \frac{60}{T_p} \left( \log_2 C + P \cdot \log_2 P + (1 - P) \cdot \log_2 \left( \frac{1 - P}{C - 1} \right) \right) \text{ (bit/min)},
\]

where \( P \) is classification accuracy, \( C \) the number of classes and \( T \) the time required for computing the output. \( T = 3 \text{s} \).

The possible online performances during the control state were also assessed by calculating the ITR using the following equation:

\[
\text{ITR} = \frac{60}{T_p} \left( \log_2 C + P \cdot \log_2 P + (1 - P) \cdot \log_2 \left( \frac{1 - P}{C - 1} \right) \right) \text{ (bit/min)},
\]

where \( P \) is classification accuracy, \( C \) the number of classes and \( T \) the time required for computing the output. \( T = 3 \text{s} \).

The simulated online control accuracy was characterized by the number of trials with missing or false control state output: a trial was regarded as correctly recognized as long as there was only one positive output during the 7 s flickering period and this output matched the attention target. Any positive results during the previous defined idle state periods and the inter-trial interval periods were considered as false positives. The number of false positives was then divided by the total idle state time to obtain a practical FPR reported in event per minute (session duration: 6.3 min).

4. Results

The SSVEPs elicited at all three frequencies under their attended conditions showed similar spatial topographies over the occipital areas (Fig. 3(a)). The extracted MER component clearly illustrated the overt attentional effect: the attended flickering stimulus elicited a stronger response at the stimulation frequency and its harmonics, with more prominent responses at the fundamental frequencies (Fig. 3(b)). For each individual subject, the corresponding MER spatial filter weights showed consistently large weights over the occipital areas, with a certain degree of inter-subject variations (Fig. 3(c)). Only 12 Hz spatial filters were shown. Interestingly, the spatial preference plots (Fig. 3(d)) showed that...
the MER methods preferred a spatial topography mainly over the occipital areas, while the CCA preference covered a relatively larger area including part of the centro-parietal areas. The MER preference topography, therefore, was more similar to the SSVEP topography at 12Hz, as shown in Fig. 3(a).

The grand-average power spectra of electrode Oz and the 1st 12Hz MER components under different idle state conditions are shown in Figs. 4(a) and 4(b). The SSVEP responses at 12Hz were significantly enhanced in the MER component, compared with the spontaneous activity, especially the prominent alpha power during the eye-closed idle state condition. The time–frequency plots from subject S5 at electrode Oz are shown in Figs. 4(g)–4(l). Clear peaks at the attended frequencies were observed during the control state conditions (Figs. 4(g)–4(i)). During the idle states, however, no peaks at the stimulation frequencies were seen, even for the on-screen idle condition (Figs. 4(j)–4(l)). Although the eye-closed idle condition was associated with a strong spectral power around the alpha band (8–12Hz), this power peak was neither narrow-banded nor temporally stationary (Fig. 4(k)). The single-epoch maximal probability values (i.e. the maximal output of the three control versus idle classifier, Eq. (11)) from one representative subject (S2) using the three algorithms are shown in Figs. 4(c)–4(e). A clear separation of the control state and different types of idle states were observed in Fig. 4(c), demonstrating the effectiveness of the MER algorithm. CCA could also recognize a majority of the idle state epochs correctly, with relatively more overlaps with the control state outputs (Fig. 4(e)). The PS algorithm failed to detect the idle state, especially in the eye-closed idle state condition, possibly due to the increased alpha power over parietal-occipital areas (Fig. 4(d)). The ROC curves corresponding to these probability value plots are given in Fig. 4(f): the MER algorithm yielded the largest AUC. The ROC results for all subjects are summarized in Table 1. The AUC were significantly larger for the MER algorithm than the other two algorithms ($p < 0.05$ for all pairwise comparisons between MER and CCA/PS), indicating better discriminability of the control and noncontrol states of the proposed method.
Fig. 4. (a) Grand-average single-epoch spectrum during the three idle state conditions and the attending 12 Hz condition. (b) Grand-average single-epoch spectrum of the 1st 12Hz MER component. (c)–(e) The single-epoch maximal probability values from one representative subject (S2) using the three algorithms. (f) The ROC curves corresponding to the results shown in (c)–(e). (g)–(l) Time–frequency plots for different control and idle state conditions, data from electrode Oz of subject S5.
Averaged the idle state detection. The average accuracy for the MER algorithm outperformed the other two algorithms is listed in Tables 2 and 3. The proposed closed idle states (on-screen/off-screen/eye-closed: 0.92 ± 0.07) declined significantly in the eye-closed condition different idle state conditions: the AUCs by CCA other two algorithms were quite sensitive to the algorithm. Moreover, while the AUC results of the MER algorithm were relatively stable (on-screen/off-screen/eye-closed: 0.94 ± 0.05 (SD)/0.03 ± 0.07/0.95 ± 0.05, one-way ANOVA p > 0.5), the results of the other two algorithms were quite sensitive to the different idle state conditions: the AUCs by CCA declined significantly in the eye-closed condition (on-screen/off-screen/eye-closed: 0.81 ± 0.14/0.86 ± 0.11/0.80 ± 0.18); the AUCs by PS showed relatively low values in both the off-screen and eye-closed idle states (on-screen/off-screen/eye-closed: 0.83 ± 0.09/0.66 ± 0.13/0.70 ± 0.17).

The BCI performance for all the three algorithms is listed in Tables 2 and 3. The proposed MER algorithm outperformed the other two algorithms in both the control state classification and the idle state detection. The average accuracy for recognizing different control states was significantly higher (MER versus CCA: 88.0 ± 11.1 versus 79.9 ± 15.1%, p < 0.05; MER versus PS: 88.0 ± 11.1% versus 74.8 ± 13.9%, p < 0.05). Moreover, the MER algorithm yielded less than 15% FPRs on average for all idle state conditions (on-screen/off-screen/eye-closed: 12.0 ± 8.2%/14.2 ± 10.1%/7.4 ± 5.6%), while the other two algorithms resulted in much higher FPRs (CCA on-screen/off-screen/eye-closed: 41.4 ± 24.1%/31.1 ± 18.3%/42.4 ± 24.1%; PS on-screen/off-screen/eye-closed: 36.8 ± 19.1%/70.6 ± 20.3%/55.8 ± 20.3%). The estimated ITRs show a similar result: the MER algorithm yielded a significant higher ITR than the other two algorithms (MER versus CCA: 60.7 ± 24.3 bit/min versus 45.4 ± 27.3 bit/min, p < 0.05; MER versus PS: 60.7 ± 24.3 bit/min versus 35.0 ± 18.6 bit/min, p < 0.05), while the ITRs of

<table>
<thead>
<tr>
<th>Subject no.</th>
<th>Control versus on-screen idle</th>
<th>Control versus off-screen idle</th>
<th>Control versus eye-closed idle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MER</td>
<td>CCA</td>
<td>PS</td>
</tr>
<tr>
<td>S1</td>
<td>0.99</td>
<td>0.95</td>
<td>0.90</td>
</tr>
<tr>
<td>S2</td>
<td>0.91</td>
<td>0.90</td>
<td>0.79</td>
</tr>
<tr>
<td>S3</td>
<td>0.98</td>
<td>0.97</td>
<td>0.92</td>
</tr>
<tr>
<td>S4</td>
<td>0.83</td>
<td>0.69</td>
<td>0.68</td>
</tr>
<tr>
<td>S5</td>
<td>0.91</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>S6</td>
<td>0.96</td>
<td>0.64</td>
<td>0.73</td>
</tr>
<tr>
<td>S7</td>
<td>1.00</td>
<td>0.98</td>
<td>0.96</td>
</tr>
<tr>
<td>S8</td>
<td>0.92</td>
<td>0.65</td>
<td>0.83</td>
</tr>
<tr>
<td>S9</td>
<td>0.94</td>
<td>0.73</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Averaged 0.94 ± 0.05 0.81 ± 0.14 0.83 ± 0.09 0.92 ± 0.07 0.86 ± 0.11 0.66 ± 0.13 0.95 ± 0.05 0.80 ± 0.15 0.70 ± 0.17

Table 2. Control state classification accuracies and ITRs (bit/min).

<table>
<thead>
<tr>
<th>Subject no.</th>
<th>MER</th>
<th>CCA</th>
<th>PS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy (%)</td>
<td>ITR</td>
<td>Accuracy (%)</td>
</tr>
<tr>
<td>S1</td>
<td>96.5</td>
<td>79.9</td>
<td>88.1</td>
</tr>
<tr>
<td>S2</td>
<td>80.6</td>
<td>40.9</td>
<td>75.0</td>
</tr>
<tr>
<td>S3</td>
<td>98.6</td>
<td>87.9</td>
<td>99.3</td>
</tr>
<tr>
<td>S4</td>
<td>68.0</td>
<td>19.2</td>
<td>52.8</td>
</tr>
<tr>
<td>S5</td>
<td>89.6</td>
<td>60.0</td>
<td>86.1</td>
</tr>
<tr>
<td>S6</td>
<td>95.1</td>
<td>75.2</td>
<td>79.2</td>
</tr>
<tr>
<td>S7</td>
<td>99.6</td>
<td>87.9</td>
<td>95.8</td>
</tr>
<tr>
<td>S8</td>
<td>78.5</td>
<td>37.1</td>
<td>61.8</td>
</tr>
<tr>
<td>S9</td>
<td>88.9</td>
<td>58.3</td>
<td>80.6</td>
</tr>
</tbody>
</table>

Averaged 88.0 ± 11.1 60.7 ± 24.3 79.9 ± 15.1 45.4 ± 27.3 74.8 ± 13.9 35.0 ± 18.6
Idle-State Detection Algorithm for SSVEP BCI

Table 3. FPRs under different idle state conditions.

<table>
<thead>
<tr>
<th>Subject no.</th>
<th>On-screen idle FPR</th>
<th>Off-screen idle FPR</th>
<th>Eye-closed idle FPR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MER (%)</td>
<td>CCA (%)</td>
<td>PS (%)</td>
</tr>
<tr>
<td>S1</td>
<td>0.0</td>
<td>10.4</td>
<td>27.1</td>
</tr>
<tr>
<td>S2</td>
<td>10.4</td>
<td>33.3</td>
<td>27.1</td>
</tr>
<tr>
<td>S3</td>
<td>18.8</td>
<td>50.0</td>
<td>29.2</td>
</tr>
<tr>
<td>S4</td>
<td>12.5</td>
<td>47.9</td>
<td>56.3</td>
</tr>
<tr>
<td>S5</td>
<td>16.7</td>
<td>27.1</td>
<td>35.4</td>
</tr>
<tr>
<td>S6</td>
<td>18.8</td>
<td>79.2</td>
<td>75.0</td>
</tr>
<tr>
<td>S7</td>
<td>0.0</td>
<td>6.3</td>
<td>8.3</td>
</tr>
<tr>
<td>S8</td>
<td>8.3</td>
<td>60.4</td>
<td>33.3</td>
</tr>
<tr>
<td>S9</td>
<td>22.9</td>
<td>58.3</td>
<td>39.6</td>
</tr>
</tbody>
</table>

Averaged: \( \frac{12.0 \pm 8.2}{41.4 \pm 21.1} \) \( \frac{36.8 \pm 19.1}{14.2 \pm 10.1} \) \( \frac{31.1 \pm 18.3}{70.6 \pm 20.3} \) \( \frac{7.4 \pm 5.6}{42.4 \pm 24.1} \) \( 55.8 \pm 20.3 \)

CCA and PS did not differ significantly (CCA versus PS: \( 24.5 \pm 27.3 \) bit/min versus \( 35.0 \pm 18.6 \) bit/min, \( p = 0.06 \)).

One example of the simulated online output, taken from one average subject (S5), is illustrated in Fig. 5. The proposed MER algorithm showed the best match with the expected output, during both the control condition and the shown idle condition. The complete results are given in Tables 4 and 5. On average, \( 22.0 \pm 2.9 \) trials were correctly recognized out of the 24 trials by MER, which was significantly better than CCA and PS (\( 19.0 \pm 4.7 \) and \( 17.9 \pm 5.2 \) trials, respectively, \( p < 0.05 \)). The better recognition accuracies also resulted in higher simulated online ITRs: MER results show a significantly higher ITR than CCA and PS (MER versus CCA: \( 23.8 \pm 8.7 \) bit/min versus \( 15.8 \pm 10.0 \) bit/min, \( p < 0.05 \); MER versus PS: \( 23.8 \pm 8.7 \) bit/min versus \( 13.2 \pm 7.8 \) bit/min, \( p < 0.05 \)). The simulated online ITRs were lower than the offline estimations.
Table 4. Simulated online control state classification results (number of trials correctly recognized/24 trials in total) and ITRs (bit/min).

<table>
<thead>
<tr>
<th>Subject no.</th>
<th>MER</th>
<th>CCA</th>
<th>PS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hit</td>
<td>ITR</td>
<td>Hit</td>
</tr>
<tr>
<td>S1</td>
<td>23</td>
<td>25.9</td>
<td>22</td>
</tr>
<tr>
<td>S2</td>
<td>21</td>
<td>18.3</td>
<td>20</td>
</tr>
<tr>
<td>S3</td>
<td>24</td>
<td>31.7</td>
<td>24</td>
</tr>
<tr>
<td>S4</td>
<td>15</td>
<td>5.1</td>
<td>9</td>
</tr>
<tr>
<td>S5</td>
<td>21</td>
<td>18.3</td>
<td>21</td>
</tr>
<tr>
<td>S6</td>
<td>24</td>
<td>31.7</td>
<td>19</td>
</tr>
<tr>
<td>S7</td>
<td>24</td>
<td>31.7</td>
<td>23</td>
</tr>
<tr>
<td>S8</td>
<td>23</td>
<td>25.9</td>
<td>14</td>
</tr>
<tr>
<td>S9</td>
<td>23</td>
<td>25.9</td>
<td>19</td>
</tr>
</tbody>
</table>

Averaged  **22.0 ± 2.9**  **23.8 ± 8.7**  19.0 ± 4.7  15.8 ± 10.0  17.9 ± 5.2  13.2 ± 7.8

Table 5. Simulated online FPR results.

<table>
<thead>
<tr>
<th>Subject no.</th>
<th>On-screen idle FPR (event/min)</th>
<th>Off-screen idle (event/min)</th>
<th>Eye-closed idle (event/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MER</td>
<td>CCA</td>
<td>PS</td>
</tr>
<tr>
<td>S1</td>
<td>0.00</td>
<td>1.14</td>
<td>3.70</td>
</tr>
<tr>
<td>S2</td>
<td>0.00</td>
<td>2.84</td>
<td>5.12</td>
</tr>
<tr>
<td>S3</td>
<td>0.85</td>
<td>1.14</td>
<td>1.42</td>
</tr>
<tr>
<td>S4</td>
<td>0.00</td>
<td>8.25</td>
<td>11.69</td>
</tr>
<tr>
<td>S5</td>
<td>0.28</td>
<td>2.47</td>
<td>3.70</td>
</tr>
<tr>
<td>S6</td>
<td>1.14</td>
<td>8.82</td>
<td>7.96</td>
</tr>
<tr>
<td>S7</td>
<td>0.00</td>
<td>0.57</td>
<td>1.42</td>
</tr>
<tr>
<td>S8</td>
<td>0.00</td>
<td>8.53</td>
<td>4.55</td>
</tr>
<tr>
<td>S9</td>
<td>1.99</td>
<td>3.98</td>
<td>7.39</td>
</tr>
</tbody>
</table>

Averaged  **0.47 ± 0.71**  **4.17 ± 3.43**  **5.15 ± 3.17**  **0.40 ± 0.76**  **1.78 ± 1.82**  **10.64 ± 5.15**  **0.05 ± 0.11**  **4.42 ± 3.79**  **6.61 ± 4.25**

due to the additional time-threshold criterion, as at least 3 s were required to one BCI command output. More importantly, the FPR by MER was lower than 0.5 event/min in both the on-screen idle state and the off-screen idle state conditions and an even lower FPR of around 0.05 event/min was achieved in the eye-closed idle state condition.

5. Discussions and Conclusions

The proposed MER method is capable of yielding more neurophysiologically plausible spatial filters than the conventional spatial filtering methods such as CCA, since it takes into account the inter-trial consistency of the elicited SSVEPs, which is believed to be the underlying neural mechanism of event-related brain responses.\(^{32-34}\) More specifically, while the inter-trial amplitude consistency was similarly emphasized in both the MER method and conventional Fourier-based methods, the inter-trial phase coherence was largely ignored in these conventional methods. In addition, MER explicitly emphasized the stationarity of the SSVEP spatial pattern across different epochs: the optimization equation (Eq. (10)) sought for the same spatial filter weights for both the single-epoch and temporally averaged data. The temporally stationary brain response showing both the amplitude and phase consistency across trials is likely to be a better representation of the ‘true’ SSVEPs. Indeed, the MER spatial filter weights conformed better with the spatial distribution of the SSVEP spectral power compared with the CCA filters. Also, the spectra of the extracted MER components showed spectral peak at the SSVEP frequencies and the attentional modulation of the SSVEPs were clearly observed.
Idle-State Detection Algorithm for SSVEP BCI

(Fig. 3(b)). More importantly, MER produced higher classification accuracies during the control state and significantly less false positives in different idle state conditions, compared with both the CCA and the power spectrum-based algorithms. These results demonstrated that the MER spatial filters provide a more precise description of the SSVEP responses, thus facilitating both control state classification and idle state detection.

The design of the proposed algorithm is different from previous SSVEP BCI algorithms in three major aspects. First, contrary to the use of only the responses at the attended frequencies, here responses at all stimulation frequencies, including the attended and unattended ones, were employed to form the feature vector for classification. Second, multiply binary classifiers, each for one control state versus the idle state, were constructed. Such a design was able to enhance the detection rate of the idle state, by emphasizing the differences between the control state and the idle state. A similar idea has been previously exploited for recognizing the idle state in the N200 speller and promising results were reported. The proposed algorithm is expected to achieve better performances when applying to SSVEP BCI systems with more targets, as the control state can be more precisely described by the combination of SSVEP responses at more stimulation frequencies.

Third, although the proposed algorithm needs to be trained using EEG data collected in a synchronous manner (i.e. for inter-trial phase consistency), it can be applied for both synchronous and asynchronous SSVEP BCIs, as the classification only relies on the spectral powers at the stimulation frequencies without using the phase information. Notably, the simulated online classification in the present study was carried out in an asynchronous way: the simulated BCI output one BCI command every second continuously.

The proposed algorithm was evaluated under a variety of different idle state conditions. The on-screen idle state condition was intended to represent the situations when the subjects need to read information on screen but not controlling the BCI. The other two idle state conditions, namely the off-screen idle state and eye-closed idle state, were similar to real-world situations when the subjects were doing something else off the screen, including relaxation. The online simulation revealed less than one false positive event every two minute during the on-screen idle state and off-screen idle state conditions. More importantly, the eye-closed condition, which was a challenging task for most previous SSVEP BCI algorithms, showed an even lower FPR (0.05 event/min) and seven out of the nine subjects showed no false alarm during the 6.3 min experiment session. The BCI results were in accordance with the power spectrum analyses as shown in Figs. 4(a) and 4(b): the MER method selectively enhanced the evoked responses at the stimulation frequency, thereby providing better discrimination between the control state and the idle state.

Nevertheless, the proposed algorithm could be further improved before moving onto an online system. First, additional strategies may be incorporated into the present system, such as the time-threshold strategy already applied in the online simulation of the present study. In the present study, a 3 s time threshold was applied and proved to be effective. For instance, although the 1 s epoch-based classification accuracy during the control state was 78.5% for subject S8, the online simulation recognized 23 out of 24 events correctly with less than one false positive per minute during all idle state conditions. An optimal proper selection of the time-threshold remained to be systematically investigated in order to achieve a better balance between system speed and performance. Also, the threshold of the probability value for discriminating the control state and the noncontrol state in the binary SVM classifiers were set to 0.5 in the present study, which might not be the optimal choice. For a practical system, misclassification during the control state may be more acceptable than false positives during the idle state, as these two types of errors may have different consequences on the user experience. Therefore, a proper selection of the threshold may be highly dependent on the specific application scenario and the relative risk of false positives versus false negatives. In addition, hybrid BCI technique may be considered as well, especially for subjects with relatively low performances (e.g. subject S2 and S4 in the present study). Besides, the performance of the proposed algorithm was only evaluated with young, healthy subjects: further studies are needed to verify the performances for elder people as well as severely motor disabled people.

Although promising results were obtained for the proposed MER method, the MER criterion also
poses several limitations on its applications. First, the proposed method cannot be used for a zero-training system, as a training phase is necessary for collecting EEG data for the calculation of the MER spatial filters. Second, the training stage needs to include trials with stimulation phases precisely aligned at the steady-state frequencies, in order to obtain reliable ‘evoked responses’ (i.e. the temporal average brain responses across trials), which is the most critical component for the MER method. Third, the MER method may not be applicable for most ERP-based BCIs using the P300 component,\textsuperscript{6,39} as the P300 latency is not sufficiently stable across trials, which may result in poor inter-trial phase coherence.

In summary, a new algorithmic framework together with a novel MER spatial filtering method was developed for the purpose of implementing practical SSVEP BCI systems. Focusing on idle state detection, the proposed MER method aimed at extracting more neurophysiologically plausible spatial filters that could maximize the evoked steady-state visual responses while suppressing unrelated background EEGs. Both the offline classification and the online simulation showed the effectiveness of the proposed algorithm: compared with the most popular CCA algorithm and the conventional power spectrum algorithm, the performance of the MER algorithm was not only significantly better but also close to practical use. These results demonstrate the practicability of the proposed algorithm for SSVEP BCI applications in real-world scenarios.

Acknowledgments
This work was supported by the National Science Foundation of China under Grant #61101151 and #61403144, Tsinghua University Initiative Scientific Research Program, and the Capital Health Research and Development Special 2011-4032-01.

References