

A HIGH-SPEED BRAIN SPELLER USING STEADY-STATE VISUAL EVOKED POTENTIALS

MASAKI NAKANISHI

*Graduate School of Science and Technology
Keio University, Yokohama, Kanagawa, 223-8522, Japan
nakanishi@mitsu.sd.keio.ac.jp*

YIJUN WANG* and YU-TE WANG†

*Swartz Center for Computational Neuroscience
Institute for Neural Computation
University of California, San Diego, La Jolla, California, 92093, USA
*yijun@scn.ucsd.edu
†yute@scn.ucsd.edu*

YASUE MITSUKURA

*Graduate School of Science and Technology
Keio University, Yokohama, Kanagawa, 223-8522, Japan
mitsukura@sd.keio.ac.jp*

TZYY-PING JUNG

*Swartz Center for Computational Neuroscience
Institute for Neural Computation
Center for Advanced Neurological Engineering
Institute of Engineering in Medicine
University of California, San Diego, La Jolla, California, 92093, USA
jung@scn.ucsd.edu*

Implementing a complex spelling program using a steady-state visual evoked potential (SSVEP)-based brain-computer interface (BCI) remains a challenge due to difficulties in stimulus presentation and target identification. This study aims to explore the feasibility of mixed frequency and phase coding in building a high-speed SSVEP speller with a computer monitor. A frequency and phase approximation approach was developed to eliminate the limitation of the number of targets caused by the monitor refresh rate, resulting in a speller comprising 32 flickers specified by eight frequencies (8-15Hz with a 1Hz interval) and four phases (0, 90, 180, and 270 degrees). A multi-channel approach incorporating Canonical Correlation Analysis (CCA) and SSVEP training data was proposed for target identification. In a simulated online experiment, at a spelling rate of 40 characters per minute, the system obtained an averaged information transfer rate (ITR) of 166.91 bits/min across 13 subjects with a maximum individual ITR of 192.26 bits/min, the highest ITR ever reported in EEG-based BCIs. The results of this study demonstrate great potential of a high-speed SSVEP-based BCI in real-life applications.

Keywords: Steady-state visual evoked potential; brain-computer interface; mixed frequency and phase coding; speller.

1. Introduction

In recent years, electroencephalogram (EEG)-based brain-computer interfaces (BCIs) have received increasing attention from researchers in neural engineering, neuroscience, and clinical rehabilitation.¹⁻¹⁰

Although the performance of current BCIs has improved significantly in the past few decades, the current performance level does not support widespread usages. Around the year 2000, the maximum information transfer rate (ITR) of BCI systems was up to 10-25 bits/min.¹ Recent progresses in system design

and data analysis considerably improved the ITR of BCIs. For example, the BCI-based spelling systems that use P300 potential and visual evoked potentials (VEPs) show their advantages of large amount of commands and high communication speed.^{11, 12} Typically, a P300- or VEP-based BCI is capable of reaching more than 30 characters for spelling. The P300 speller (a 6x6 matrix) proposed by Farwell and Donchin,¹³ for example, reached a speed up to 5 characters/min. Jin *et al.*¹⁴ obtained an ITR of 35 bits/min using a 12x7 matrix speller. Recently, in an electrocorticogram (ECoG)-based P300 BCI, a subject achieved an averaged ITR of 69 bits/min and a peak ITR of 113 bits/min.¹⁵ The speed of a P300-based speller is limited by the time during which the characters are intensified and the time between two consecutive characters.¹⁶ In general, VEP-based spellers show a higher communication speed than P300 spellers. Bin *et al.*¹⁷ reported an ITR of 58 bits/min in a 6-target system based on steady-state visual evoked potentials (SSVEPs). Recently, higher ITRs were reported in SSVEP-based BCIs. Nakanishi *et al.*¹⁷ reported an ITR of 95 bits/min in an 8-target system. Chen *et al.*¹⁸ reported an ITR of 105 bits/min in a 45-target system. High ITRs were also reported in BCIs using other types of VEP signals. For example, using a code modulated VEP (c-VEP) paradigm, Bin *et al.*¹⁹ reported an ITR of 108 bits/min in a 32-target speller. By applying adaptive learning and error correction methods, Spüler *et al.*²⁰ obtained an ITR of 144 bits/min using a similar c-VEP paradigm. In an early study using c-VEP, a subject with implanted intracranial electrodes reached communication rates of 10-12 words/min.²¹

Generally, the performance of an SSVEP-based BCI depends on three major factors: (1) stimulus presentation, (2) multiple target coding, and (3) target identification. First, the properties of flickers in stimulus presentation need to be stable and accurate so that the elicited SSVEP signals are robust and reliable for accurate target identification. When using a monitor-based stimulator, the number of targets is always limited by the monitor refresh rate. Recently, the approximation approach proposed by Wang *et al.*²² solved this problem by using varied frequencies in different stimulation cycles. Using this method, any frequency lower than the half of the refresh rate can be accurately rendered on a computer screen.²³ Second, target coding plays an important role in system design and implementation. An

efficient target coding approach can enhance the signal-to-noise ratio (SNR) of VEP and thereby increase the discriminability of different targets. The multiple access methods widely used in telecommunications have been adopted in VEP BCIs.²⁴ For example, the time division multiple access (TDMA), frequency division multiple access (FDMA), and code division multiple access (CDMA) methods have been applied to current VEP BCIs.²⁴ Third, a high-speed VEP BCI requires an efficient target identification method.¹² Recently, a multi-channel approach based on Canonical Correlation Analysis (CCA) has been widely used in SSVEP BCIs.^{16, 25} Compared with power spectrum density (PSD)-based methods using single-channel EEG data, CCA can significantly improve target identification accuracy.²⁵ The CCA-based method has been further improved in several recent studies.^{26, 27, 28}

This study aims to explore the communication speed limit of an SSVEP-based BCI. Currently, most SSVEP BCIs showed ITRs lower than 60 bits/min.¹⁶ The highest ITR reported in SSVEP-based BCIs, which employed frequency coding in system design, was 105 bits/min.¹⁸ Obviously, there is room for improvement. By employing hybrid coding approaches such as mixed frequency and phase coding,²⁹ the number of targets can be increased by integrating frequency and phase information while the discriminability of different targets remains. Furthermore, target identification approaches could be improved by integrating subject specific training data into the CCA-based method. Therefore, by solving the problems of stimulus presentation, multiple target coding, and target identification, the SSVEP-based BCI has potential to achieve a higher communication speed. This study proposes and develops a high-speed SSVEP BCI to support fast spelling at a speed of 40 characters per minute, which corresponds to an ITR up to 200 bits/min, a new world record to the best of our knowledge. To this end, this study first develops an extended approximation approach in visual stimulus presentation to realize accurate coding of both frequency and phase of SSVEPs. The limitation of number of targets caused by the monitor refresh rate can thereby be totally eliminated. A 32-target speller can be implemented by rendering the flickers characterized by eight frequencies (8-15Hz with a 1Hz interval) and four phases (0, 90, 180, and 270 degrees) on a regular LCD monitor. To assess the frequency and phase information

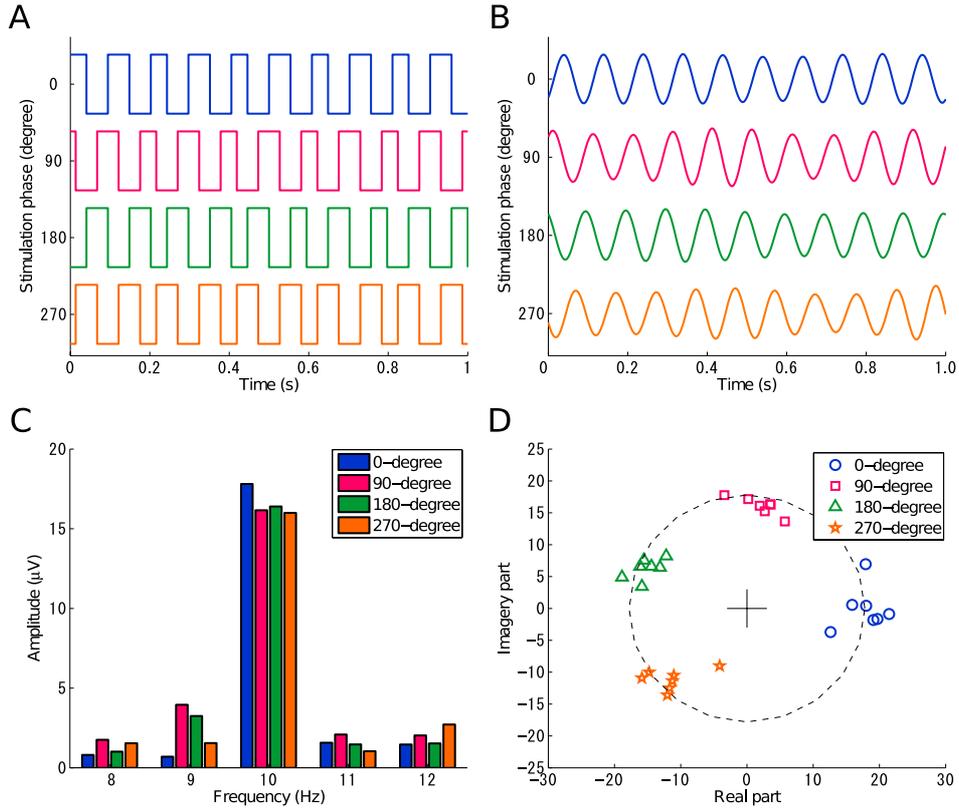


Fig. 1. Time series of quad-phase coded (A) flickering signals, (B) elicited SSVEPs by the 10Hz stimuli presented on an LCD monitor with a 75Hz refresh rate, and (C) amplitude spectra, (D) a scatter diagram of the complex spectra of elicited SSVEPs at 10Hz from a subject.

within short time durations, this study proposes to incorporate CCA and individual SSVEP training data in target identification. The primary novelties and contributions of this study include: (1) a frequency and phase approximation approach for rendering visual flickers using a regular monitor refresh rate, e.g. 60Hz, (2) a 32-target speller design using a mixed frequency and phase coding scheme, and (3) a target identification approach using CCA with reference signals derived from individual SSVEP training data.

The rest of this paper is organized as follows. Section 2 describes the frequency and phase approximation approach, the target identification method using CCA with SSVEP training data, and the experimental setups in the spelling experiment. Section 3 illustrates the signal characteristics of the SSVEP signals and the results of BCI performance in the spelling experiment. Section 4 discusses important factors that affect the system performance and points out some directions for further improvement. Finally, Section 5 concludes the study.

2. Methods

2.1. Visual stimulus design

In the conventional SSVEP-based BCI systems that use a computer monitor to present visual stimuli, a visual flicker that alternates white and black frames at a specified frequency is used to elicit SSVEPs. In this way, the number of flickers that can be presented simultaneously on the screen is always limited by the refresh rate. It is impossible to realize the frequencies by which the refresh rate is not dividable (e.g., an 11Hz flicker under a 60Hz refresh rate) because the white/black reversal needs to occur every 2.73 frames. Wang *et al.*²² proposed an approximation approach that can generate a visual flicker at a flexible frequency by approximating the frequency with variable number of frames in a stimulation cycle. For instance, a flicker at 11Hz under a 60Hz refresh rate can be realized by interleaving five and six frames in a stimulation cycle as ‘1110001110011100011100111...’. Based on this

approach, any stimulus frequency up to the half of the monitor refresh rate can be realized. Therefore, the number of visual stimuli that can be presented on a monitor can be considerably increased.^{18, 22} Importantly, a recent study found that the phase and latency of SSVEPs elicited by the approximation approach are stable across different frequencies,¹⁷ suggesting that the approximation approach could be extended to the phase domain. In this way, the mixed frequency and phase coding approach, which was proved highly efficiently in stimulus coding,²⁹ could be used to implement a BCI speller with a large amount of targets. Generally, to render a visual flicker at frequency f with an initial phase ϕ , a stimulus sequence $s(f, \phi, i)$ can be generated by:

$$s(f, \phi, i) = \text{square} \left[2\pi f \left(\frac{i}{\text{RefreshRate}} \right) + \phi \right] \quad (1)$$

where $\text{square}[\]$ generates a 50% duty cycle square wave with levels 0 and 1, and i indicates the frame index. In this way, visual stimuli tagged by the same frequency but different phases can be realized. Theoretically, the extended approximation approach can thereby implement visual flickers at any frequency (up to the half of the refresh rate) and any phase (from 0 to 360 degrees). Fig. 1A shows time sequences of quad-phase coded flickering signals (0, 90, 180, and 270 degrees) at 10Hz on an LCD monitor with a 75Hz refresh rate. According to the photic driving effect,³⁰ frequency and phase information of the target stimulus can be accurately decoded from the resulting SSVEP signals (Fig. 1B).

2.2. Signal characteristics analysis

The stimulus sequence generated by (1) can accurately encode specified frequency and phase in a visual flicker. The frequency and phase of the resulting SSVEPs can be estimated through calculating the complex spectrum using fast Fourier transform (FFT). An amplitude spectrum value and a phase spectrum value can be obtained at the stimulation frequency in the complex spectrum. The initial phase is the angle between the imaginary and real parts of the complex spectrum value. The amplitude and initial phase of SSVEPs can be estimated as follows:

$$A_x(f) = \text{abs} \left[\frac{1}{N} \sum_{n=1}^N \mathbf{x}(n) e^{-j2\pi \left(\frac{f}{f_s} \right) n} \right]$$

$$\phi_x(f) = \text{angle} \left[\frac{1}{N} \sum_{n=1}^N \mathbf{x}(n) e^{-j2\pi \left(\frac{f}{f_s} \right) n} \right] \quad (2)$$

where $x(n)$ is single-channel EEG signal, f is the stimulation frequency, f_s is the sampling rate, and N is the number of data points. Then, the phase difference (in π radians) between SSVEPs at two stimulating frequencies (f_1 and f_2) can be defined as:

$$\Delta\phi_x(f_1 - f_2) = \phi_x(f_1) - \phi_x(f_2) \quad (3)$$

The response latency t (in milliseconds) between the visual stimulus and the SSVEP response can be derived by measuring phase as a function of stimulating frequency and estimating the slope of the curve:^{31, 32}

$$t = - \frac{\Delta\phi_x(f_1 - f_2)}{2 \times (f_1 - f_2)} \times 1000 \quad (4)$$

2.3. Target identification

2.3.1. Standard CCA-based method

In an SSVEP BCI using mixed frequency and phase coding, target identification can be performed using a two-step strategy: (1) frequency detection, and (2) phase detection. This study first proposes a standard CCA-based method for frequency and phase detection. The frequency of SSVEPs is first detected by the CCA approach, and then the initial phase, which can be obtained through calculating the angle of complex spectrum value at the detected frequency, is compared with reference phases measured from the training data. CCA has been widely used to detect the dominant frequency of SSVEPs in SSVEP-based BCIs.^{16, 25} CCA is a statistical way to measure the linear relationship between two multidimensional variables, which may have some underlying correlation. Considering two multidimensional variable \mathbf{X} , \mathbf{Y} and their linear combinations $\mathbf{x} = \mathbf{X}^T \mathbf{W}_X$ and $\mathbf{y} = \mathbf{Y}^T \mathbf{W}_Y$, CCA finds the weight vectors, \mathbf{W}_X and \mathbf{W}_Y , which maximize the correlation between \mathbf{x} and \mathbf{y} by solving the following problem:

$$\max_{\mathbf{W}_X, \mathbf{W}_Y} \rho(\mathbf{x}, \mathbf{y}) = \frac{E[\mathbf{W}_X^T \mathbf{X} \mathbf{Y}^T \mathbf{W}_Y]}{\sqrt{E[\mathbf{W}_X^T \mathbf{X} \mathbf{X}^T \mathbf{W}_X] E[\mathbf{W}_Y^T \mathbf{Y} \mathbf{Y}^T \mathbf{W}_Y]}}. \quad (5)$$

The maximum of ρ with respect to \mathbf{W}_X and \mathbf{W}_Y is the maximum canonical correlation. Projections onto \mathbf{W}_X and \mathbf{W}_Y are called canonical variants. Here, \mathbf{X} refers to the set of multi-channel EEG signals and \mathbf{Y} refers to the

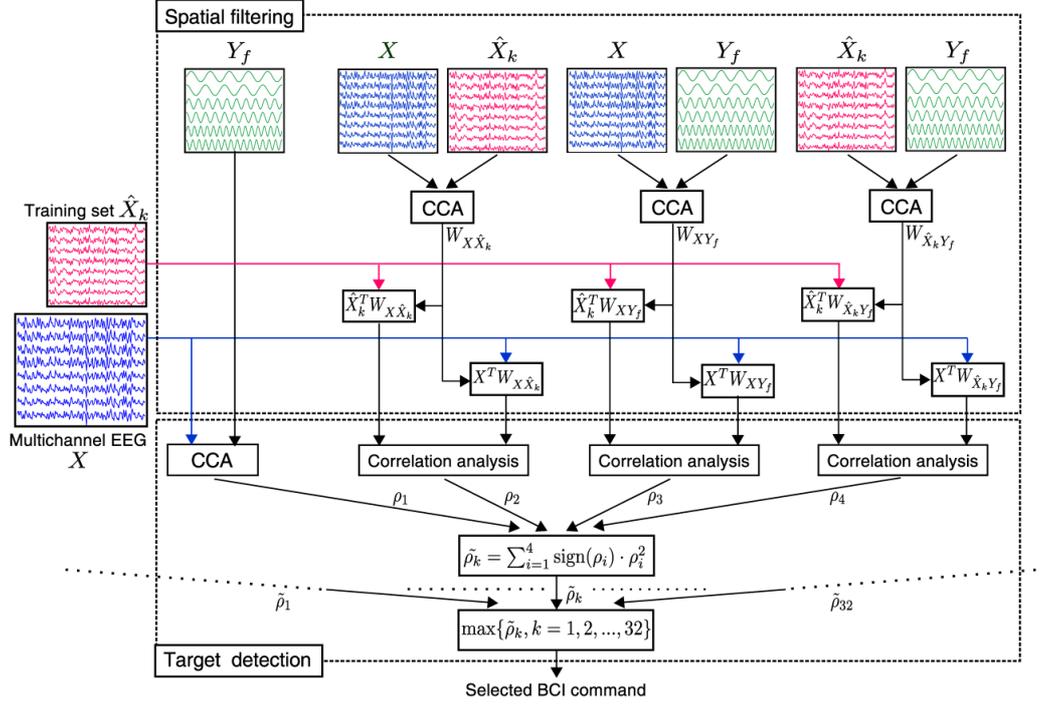


Fig. 2. Flowchart of the proposed target identification method using CCA with SSVEP training data.

set of reference signals that have the same length as \mathbf{X} . The reference signals \mathbf{Y}_f are set as

$$\mathbf{Y}_f = \begin{bmatrix} \sin(2\pi f n) \\ \cos(2\pi f n) \\ \vdots \\ \sin(2\pi N_h f n) \\ \cos(2\pi N_h f n) \end{bmatrix}, n = \frac{1}{f_s}, \frac{2}{f_s}, \dots, \frac{N}{f_s} \quad (6)$$

where f is the target frequency, N_h is the number of harmonics, and N is the number of sampling points. To recognize the frequency of the SSVEPs, CCA calculates the canonical correlation between the multi-channel EEG signals and the reference signals at each stimulus frequency. The frequency of the reference signals with the maximal correlation is selected as the frequency of SSVEPs.

Once the frequency of SSVEP is detected, the complex spectrum value at the detected frequency is calculated by FFT to obtain the initial phase of SSVEP signals from a single channel. The complex spectrum is then projected onto the reference phase directions corresponding to targets at the detected frequency, which can be measured from SSVEP training data.²² The visual stimulus with the maximum projected value is then selected as the target.

2.3.2. CCA with SSVEP training data

Since training data are required for phase detection, this study proposes a CCA-based method, which uses reference signals from SSVEP training data, for target identification in a one-step way. In the proposed method, training reference signals $\hat{\mathbf{X}}_k$ can be obtained by averaging training set \mathbf{X}_k ($k = 1, 2, \dots, K$). Correlation coefficient between projections of test set \mathbf{X} and training reference signals $\hat{\mathbf{X}}_k$ using CCA-based spatial filters (*i.e.*, \mathbf{W}_X in (5)) can be used to identify a target from K visual stimuli. In this study, three canonical coefficients including (1) $\mathbf{W}_{X\hat{\mathbf{X}}_k}$ between test set \mathbf{X} and training reference signals $\hat{\mathbf{X}}_k$, (2) \mathbf{W}_{XY} between test set \mathbf{X} and sine-cosine reference signals \mathbf{Y} , and (3) $\mathbf{W}_{\hat{\mathbf{X}}_k Y}$ between training reference signals $\hat{\mathbf{X}}_k$ and sine-cosine reference signals \mathbf{Y} are used as spatial filters for improving the SNR of SSVEPs. Fig. 2 illustrates the procedure of target identification using the proposed method. The target can be identified by recognizing the training reference signal that maximizes the correlation coefficient. Although the standard CCA-based method cannot discriminate different phases, it still contributes to frequency detection. Therefore, this study considers four different approaches including the standard CCA-based frequency detection method and the three

correlation analysis based methods using CCA with SSVEP training data.

This study develops an ensemble classifier to combine decisions from the four methods described above. A correlation vector ρ is defined as follows:

$$\rho = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} = \begin{bmatrix} \rho(X^T W_{X\hat{X}}, \hat{X}^T W_{X\hat{X}}) \\ \rho(X^T W_{XY}, \hat{X}^T W_{XY}) \\ \rho(X^T W_{X\hat{Y}}, \hat{X}^T W_{X\hat{Y}}) \end{bmatrix} \quad (7)$$

where $\rho(a, b)$ indicates the Pearson's correlation coefficient between a and b . To combine these correlation values, the following weighted correlation coefficient $\tilde{\rho}$ is used as the final feature value for target identification:

$$\tilde{\rho} = \sum_{i=1}^4 \text{sign}(\rho_i) \cdot \rho_i^2 \quad (8)$$

where $\text{sign}()$ is used to remain discriminative information from negative correlation coefficients between test set and training reference signals (*e.g.*, 0 degree vs. 180 degrees, 90 degrees vs. 270 degrees). The training reference signal that maximizes the weighted correlation value is selected as the reference corresponding to the target.

To validate the efficiency of the combined method, this study compared classification performance of the following five methods: (M1) a standard CCA-based method; (M2) a correlation analysis using a spatial filter derived from test set and training reference signals; (M3) a correlation analysis using a spatial filter derived from test set and since-cosine reference signals; (M4) a correlation analysis using a spatial filter derived from training reference signals and sine-cosine reference signals; and (M5) a combined method using the ensemble classifier described in (8).

2.4. Experiment

2.4.1. Experiment 1: Offline experiment

The offline experiment was designed to analyze the characteristics of SSVEPs elicited by the mixed frequency and phase coded visual stimuli and obtain training reference signals and classifiers for the simulated online experiment. In this experiment, 32 visual flickers (3 cm×3 cm) were concurrently rendered on a Dell S2409W 24-inch LCD monitor (Dell Inc.) with a refresh rate of 75Hz. The horizontal and vertical intervals between two stimuli were 2 cm and 3 cm

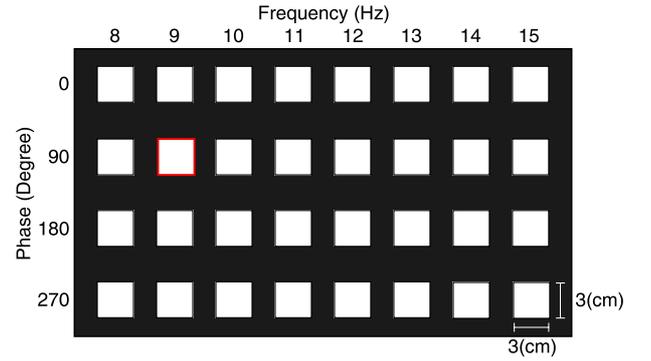


Fig. 3. Presentation of the 32-target visual stimuli using mixed frequency and phase coding.

respectively. As shown in Fig. 3, targets in a 4×8 matrix were specified by four phases (0, 90, 180, 270 degrees) and eight frequencies (8Hz to 15Hz with a 1Hz interval). Different phases and frequencies were used to label rows and columns respectively. The visual stimuli at all frequencies and phases were produced by the approximation approach described in (1). The stimulation program was developed under MATLAB (Mathworks, Inc.) using the Psychophysics Toolbox extensions.³³

Thirteen healthy adults (10 males and 3 females, mean age: 23 years) with normal or corrected-to-normal vision participated in the study. All subjects signed an informed consent form approved by the Research Ethics Committee of Keio University before participating in the experiment. The subjects were seated in a comfortable chair 70 cm in front of the LCD monitor in a dark room. They were asked to gaze at one of the visual stimuli (a target stimulus) for 4 seconds, and 32 targets were indicated in a random order in a run. At the beginning of each trial, a red rectangle marker (see Fig. 3) appeared for 0.5 second at the position of the target stimulus. Subjects were asked to shift their gaze to the target within the same 0.5 second duration. After that, all stimuli started to flicker simultaneously for 4 seconds on the LCD monitor. To reduce eye movement artifacts, subjects were asked to avoid eye blinks during the stimulation period. Seven runs were carried out for each subject. To avoid visual fatigue, subjects could take few minutes of rest between two runs. EEG data were recorded by 16 electrodes over the parietal and occipital areas (FPz, F3, F4, Fz, Cz, P1, P2, Pz, PO3, PO4, PO7, PO8, POz, O1, O2, and Oz) using the g.USBamp (g.tec medical engineering GmbH) with a

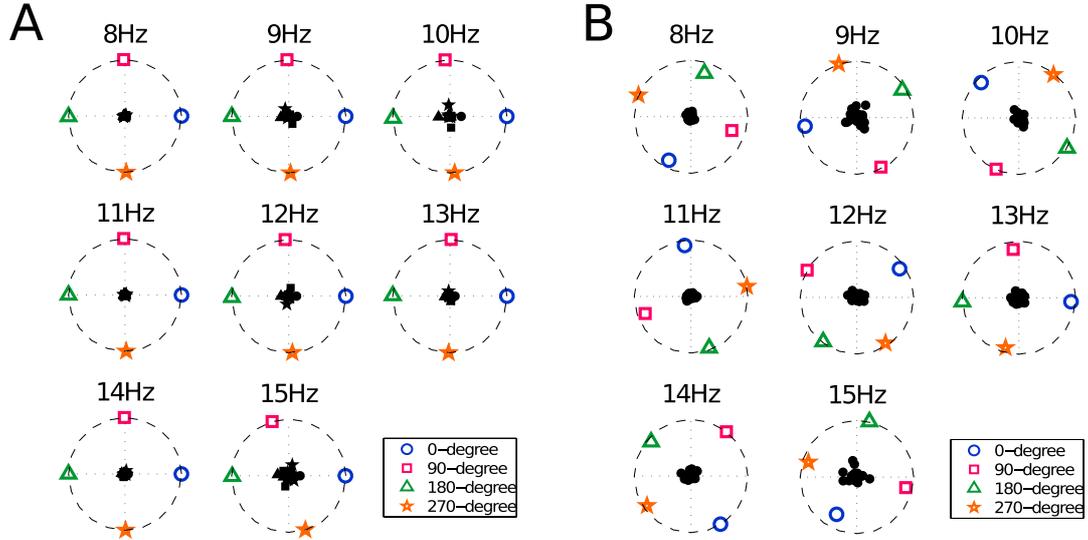


Fig. 4. Complex spectra of (A) stimulus sequences and (B) averaged SSVEPs across all subjects at each stimulation frequency. For each sub-figure, horizontal and vertical axes indicate real and imaginary parts of complex spectra. Black dots indicate spectrum values corresponding to other stimulation frequencies. Dashed circle indicates spectrum values with maximal amplitude.

sampling rate of 512Hz. Event triggers generated by the stimulus program were sent from the parallel port of the computer to the amplifier and recorded on an event channel synchronized to the EEG data. The 4s EEG data epochs synchronized to the visual stimuli were extracted and saved for further analysis.

2.4.2. Experiment 2: Simulated online experiment

To estimate an online performance of the BCI based on the proposed paradigm, a simulated online experiment using the same data recording setup, was conducted after the offline experiments. The subjects were asked to input a sequence with all 32 characters in a run, and repeat the run twice in the experiment. The order of targets was randomized in the task sequence. Unlike the 4s stimulus duration used in the offline experiment, visual stimuli flickered for only 1 second on the LCD monitor in the simulated online experiment. The 1s EEG epochs synchronized to the visual stimuli were used for estimating online BCI performance.

2.5. Data analysis and performance evaluation

In offline and simulated online experiments, all data epochs were referenced using common average reference (CAR) and then band-pass filtered from 7Hz to 50Hz with an infinite impulse response (IIR) filter. Zero-phase forward and reverse IIR filtering was implemented using the `filtfilt()` function in Matlab.

Considering a latency delay in the visual system (see Equation (4)), the data epochs for offline and simulated online experiments were extracted in $[0.12s \ 4.12s]$ and $[0.12s \ 1.12s]$ respectively (time 0 indicated stimulus onset). The 120ms delay was selected towards highest classification accuracy in the offline experiment. To avoid overfitting in CCA, eight electrodes over the parieto-occipital region (PO3, PO4, PO7, PO8, POz, O1, O2, and Oz) were selected for representing the SSVEP data. In the standard CCA using sine-cosine reference signals, the number of harmonics (N_h) was set to 3. The single-channel SSVEP signal recorded from the Oz electrode with a small Laplacian reference was used for phase detection using the standard CCA-based method.

The recorded EEG epochs were classified by the proposed target identification methods described in Section 2.3. For the offline experiment, the classification accuracy was estimated using leave-one-out cross-validation. Training reference signals and classifiers derived from offline data using 1s-long data epochs were applied to analyzing data epochs from the simulated online experiment. In addition to classification accuracy, BCI performance was also evaluated by ITR calculated as follows¹:

$$ITR = \left(\log_2 M + P \log_2 P + (1 - P) \log_2 \left[\frac{1-P}{M-1} \right] \right) * \left(\frac{60}{T} \right) \quad (9)$$

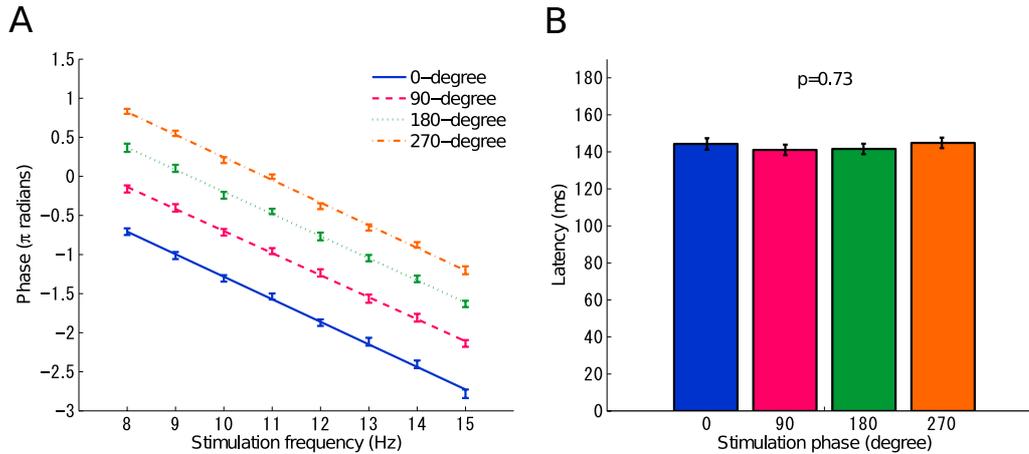


Fig. 5. (A) Averaged phases of SSVEPs at all stimulation frequencies specific to four phases, and (B) estimated latencies for four phases across subjects.

where M is the number of targets, P is the accuracy of target identification, and T (seconds/selection) is the average time for a selection. This study calculated classification performance using different T (Target gazing: 1s, 2s, 3s, 4s; Gaze shifting: 0.5s) in the offline experiment. For the simulated online experiment, T is 1.5s (1s for target gazing and 0.5s for gaze shifting).

3. Results

3.1. Amplitude, phase, and latency

Fig. 1 shows an example of stimulus signals (Fig. 1A) and elicited SSVEPs (Fig. 1B) corresponding to four different phases (0, 90, 180, 270 degrees) at 10Hz from a subject. Fig. 1C shows amplitude spectra of the SSVEPs, which exhibit peak amplitudes at 10Hz for all phases. The scatter diagram of complex spectrum for SSVEP signals (Fig. 1D) indicates four clusters corresponding to four phases. These results indicate that the frequency and phase approximation approach can elicit robust SSVEP signals that encode accurate frequency and phase information. This approach provides a flexible solution for presenting a large number of flickering targets in an SSVEP-based BCI. Fig. 4A and Fig. 4B illustrate the complex spectra of stimulus sequences and averaged SSVEPs across subjects for each stimulation frequency. The SSVEPs were recorded from the Oz electrode. Ideally, spectrum values corresponding to target stimulation signals should appear at 0, 90, 180, 270 degrees with maximal amplitude while those corresponding to non-target

stimulation signals have zero amplitude. Although there were some frequency and phase errors in stimulus sequences (e.g., stimuli at 15Hz) generated by the approximation approach, each target still has distinct frequency and phase characters and thereby can be discriminated from others. The resulting SSVEP signals show similar patterns to the stimulus signals. The change of phases between stimulus signals and SSVEPs is caused by a latency delay in the visual system, which can be estimated by (4). Fig. 5 shows the estimation of phase and latency of SSVEPs. The averaged phases across subjects were plotted in Fig. 5A as a function of stimulus frequency for four phase conditions. All frequency-phase curves fit linear models with a similar slope, indicating that the latency of SSVEP is a constant. Fig. 5B shows the latency estimated by the slope of the linear regression line. The latencies of the SSVEPs for four phase conditions (0, 90, 180, 270 degrees) were 144ms, 141ms, 142ms, and 145ms respectively. The estimated latencies are consistent with results in previous studies.^{34, 29} A one-way analysis of variance (ANOVA) indicates the difference of latency between different phase conditions was not significant ($F(3,48)=0.43$, $p=0.73$). These results suggest that, by extracting distinct frequency and phase information, single-trial SSVEP signals can be accurately discriminated using machine-learning based methods.

3.2. Offline BCI performance

Fig. 6 shows the averaged accuracy (Fig. 6A) and ITR (Fig. 6B) across all subjects for the offline experiments.

Results for different CCA-based methods were calculated with different data lengths from 1s to 4s. The five methods were described in Section 2.3.2. It is evident that the four methods (M_2 , M_3 , M_4 , and M_5) outperformed M_1 under all conditions with different data lengths. The ensemble classification method (M_5) showed the highest classification accuracy and ITR. With 1-s data length, the accuracy increased from $70.95 \pm 17.53\%$ to $88.84 \pm 11.48\%$ (M_1 vs. M_5 , $t=-5.69$, $p < 10^{-3}$), resulting in an ITR increase from 111.65 ± 41.75 bits/min to 161.00 ± 34.23 bits/min (M_1 vs. M_5 , $t=-7.43$, $p < 10^{-5}$). Since the accuracy was relatively high ($>88\%$) with all data lengths, the highest ITR was obtained with the shortest data length (1s). According to (9), when using a longer data length, a minor increase of classification accuracy leads to a significant decrease of ITR. For example, compared with 1-s data length, classification accuracy increased by 4.26% when using 2-s data length ($88.84 \pm 11.48\%$ vs. $93.10 \pm 6.16\%$, paired t-test: $t=-2.81$, $p < 0.05$); however, ITR dropped from 161.00 ± 34.23 bits/min to 104.02 ± 12.49 bits/min (paired t-test: $t=9.36$, $p < 10^{-6}$). These results indicate that the data length of 1s is optimal for achieving a high ITR in an online system using the proposed diagram.

To further compare the performance of these methods, this study calculated classification accuracy and ITR for all five methods using data lengths of 0.5s and 1s. As shown in Fig. 6C and Fig. 6D, for both conditions, the averaged classification accuracy and ITRs using the four methods (M_2 , M_3 , M_4 , and M_5) were significantly higher than the standard CCA-based method. The performance improvement is more significant with 0.5s data length. More specifically, accuracy of M_1 was significantly lower than the other methods (M_1 vs. M_2 : $t=-8.15$, $p < 10^{-5}$; M_1 vs. M_3 : $t=-9.03$, $p < 10^{-5}$; M_1 vs. M_4 : $t=-10.06$, $p < 10^{-6}$; M_1 vs. M_5 : $t=-10.45$, $p < 10^{-6}$) while M_5 significantly outperformed the others (M_5 vs. M_1 :

$t=10.45$, $p < 10^{-6}$; M_5 vs. M_2 : $t=4.81$, $p < 10^{-3}$; M_5 vs. M_3 : $t=8.58$, $p < 10^{-5}$; M_5 vs. M_4 : $t=6.52$, $p < 10^{-4}$). There was no significant difference between M_2 , M_3 , and M_4 (M_2 vs. M_3 : $t=-0.32$, $p=0.75$; M_2 vs. M_4 : $t=-0.18$, $p=0.85$; M_3 vs. M_4 : $t=-0.69$, $p=0.50$). These results proved the efficacy of the combined method based on the standard CCA method and the correlation analysis approaches using CCA-based spatial filtering.

3.3. Simulated online BCI performance

Table 1 lists the classification accuracy and ITR in the simulated online experiment. Results corresponding to frequency-only classification (8 classes), phase-only classification (4 classes), and mixed frequency and phase classification (32 classes) were calculated separately. The frequency-only classification calculated the mean of results from four 8-class frequency classifications corresponding to the four phases. Similarly, the phase-only classification calculated the mean of results from eight 4-class phase classifications corresponding to the eight frequencies. Across all subjects, classification accuracies under the three conditions were all above 90% (Frequency: $95.55 \pm 3.98\%$, Phase: $91.83 \pm 5.68\%$ Mixed: $91.35 \pm 5.69\%$). In general, frequency classification shows higher accuracy than phase classification. Three subjects (Subjects 7, 8, and 10) obtained 100% accuracy in frequency classification. Given a trial length of 1.5 s, these classification accuracies resulted in high ITRs. For mixed frequency and phase coding, an averaged ITR of 166.91 ± 18.50 bits/min (range: 126.34-192.26 bits/min) was obtain across all subjects. The ITRs for frequency and phase coding were 105.64 ± 11.55 bits/min and 59.55 ± 10.85 bits/min respectively.

Fig. 7 illustrates the simulated online performances with different data lengths. Fig. 7A shows that the

Table 1. BCI performance in simulated online experiment.

Subject	Accuracy (%)			ITR (bits/min)		
	Frequency	Phase	Mixed	Frequency	Phase	Mixed
S1	96.88	93.75	92.19	108.47	62.54	168.70
S2	93.75	90.62	90.62	99.49	56.10	163.47
S3	93.75	95.31	93.75	99.49	66.11	174.12
S4	93.75	92.19	92.19	99.49	59.23	168.70
S5	96.88	90.62	90.62	108.47	56.10	163.47
S6	96.88	93.75	92.19	108.47	62.55	168.70
S7	100.00	95.31	95.31	120.00	66.11	179.79
S8	100.00	96.88	96.88	120.00	69.99	185.78
S9	89.06	78.12	78.12	89.06	35.82	126.34
S10	100.00	98.44	98.44	120.00	74.36	192.26
S11	95.31	85.94	85.94	103.82	47.65	148.70
S12	98.44	96.88	96.88	113.60	69.99	185.78
S13	87.50	85.94	84.38	84.22	47.65	144.03
Mean±Std	95.55±3.98	91.83±5.68	91.35±5.69	105.64±11.55	59.55±10.85	166.91±18.50

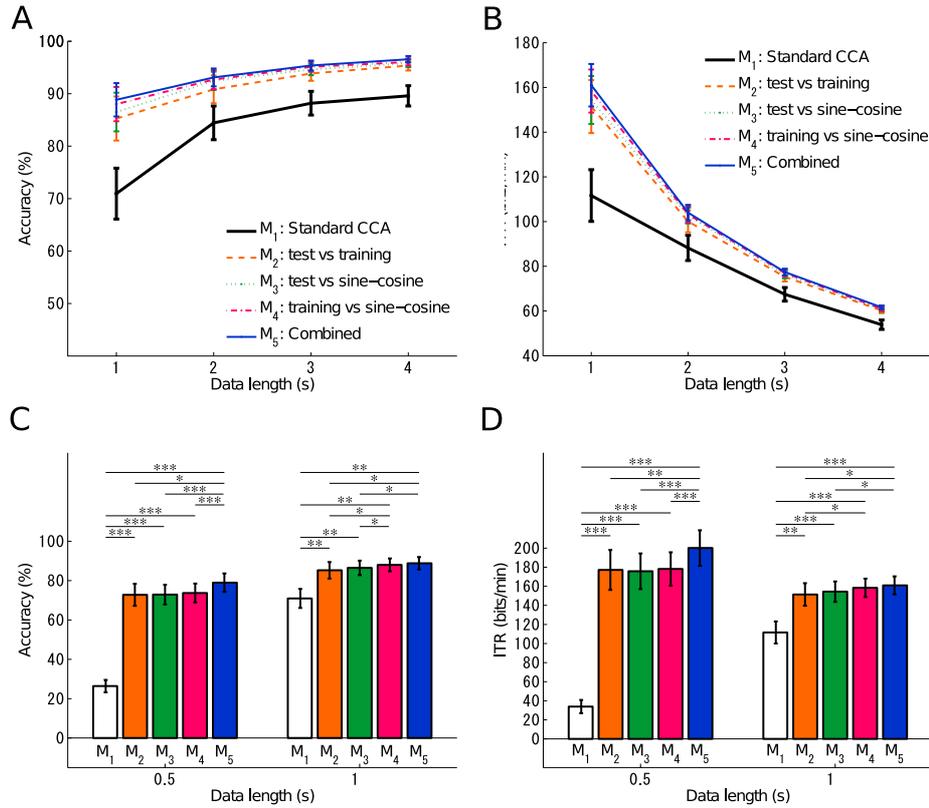


Fig. 6. (A) Target identification accuracy, (B) ITRs as functions of data length (from 1 s to 4 s) calculated by different methods, (C) accuracy, and (D) ITR for data lengths of 0.5 s and 1 s. Five methods include: (M₁) a standard CCA-based method; (M₂) a correlation analysis using spatial filter derived from test set and training reference signals; (M₃) a correlation analysis using spatial filter derived from test set and since-cosine reference signals; (M₄) a correlation analysis using a spatial filter derived from training reference signals and sine-cosine reference signals; (M₅) a combined method. The asterisks indicate significant difference between different methods (*: $p < 0.05$, **: $p < 0.005$, and ***: $p < 0.0005$).

classification accuracy increased as data length increased from 0.1 s to 1 s. The starting point of the data segment was set to time 0 (stimulus onset). Note that, a latency delay of 120ms was used in extracting the data epochs. For example, the data length of 0.1 s corresponds to the time window between 0.12s and 0.22s in real EEG data. Even with a very short data length, classification accuracy was significantly higher than the chance level under all three classification conditions. With short data lengths, the phase-only classification accuracy was relatively high compared with the frequency-only classification accuracy. As shown in Fig. 7B, by reducing target detection time using a shorter data length, ITR could be improved although accuracy decreased. This phenomenon is common under all three classification conditions. The highest ITR appeared at the data length of 0.6 s (Mixed: 203.56 ± 30.74 bits/min; Frequency: 121.58 ± 21.47

bits/min; Phase: 72.01 ± 14.96 bits/min). In practice, a tradeoff between accuracy and ITR should be considered to optimize the data length for practical system use.

4. Discussion

4.1. Information transfer rate

ITR has been widely used in evaluating system performance of BCIs.^{1, 35} In this study, an averaged ITR of 166.91 ± 18.50 bits/min was obtained across 13 subjects. The highest and lowest ITR for individuals was 192.26 bits/min (Subject 10) and 126.34 bits/min (Subject 9) respectively. To our knowledge, an averaged ITR of 166.91 ± 18.50 bits/min is the highest ITR reported in EEG-based BCIs, which broke the previous record of 144 bits/min in a code modulated VEP (c-VEP) based BCI.²⁰

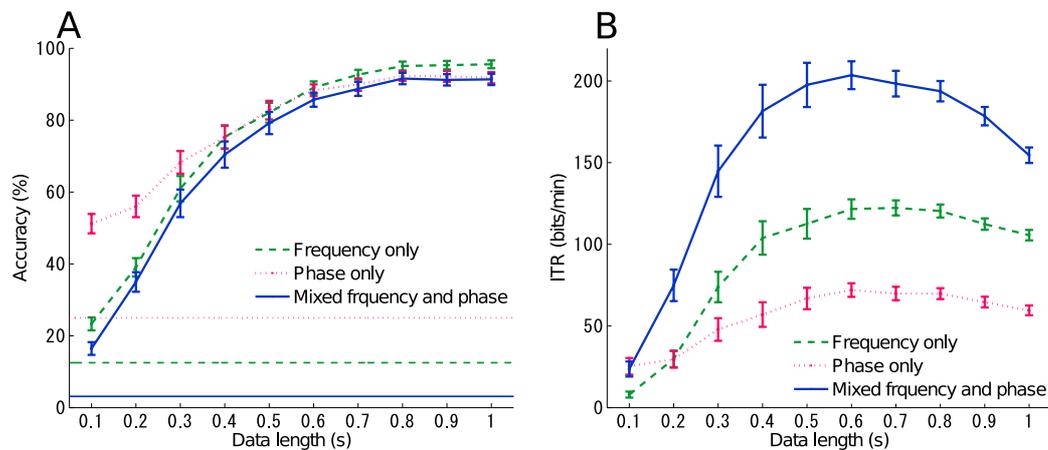


Fig. 7. (A) Target identification accuracy and (B) ITRs as functions of data length. Dashed lines indicate chance-level accuracies in classification (Frequency: 12.5%, Phase: 25%, Mixed: 3.13%).

The mixed frequency and phase coding method studied by Jia *et al.*²⁹ obtained an ITR of 66.5 bits/min using a similar testing paradigm. The distinct performance improvement (~100 bits/min) in the present study is mainly attributed to the following factors: (1) a larger number of targets (32 vs. 15), (2) shorter data length (1 s vs. 2 s), (3) multiple channel analysis approach (CCA vs. bipolar), and (4) a combined method in target identification. The proposed frequency and phase approximation approach significantly facilitates the presentation of a large amount of concurrent stimuli on a computer monitor. The results indicate that the robustness of the SSVEPs elicited by the approximation approach can satisfy the requirements of signal characters for discriminating SSVEPs with a high frequency and phase resolution (1Hz and 90 degrees respectively). The proposed CCA-based multi-channel approach extracted SSVEP signals with a much higher SNR than the single-channel counterparts. The combined method further improved the classification performance, leading to accuracy above 90% when the data length was reduced to 1 s.

The present study achieved a significant performance improvement compared to other existing SSVEP-based BCIs. The first online BCI system using the CCA method obtained an ITR of 58 bits/min.¹⁶ ITRs around 100 bits/min were reported in two recent studies using the frequency approximation approach.^{17,18} The present study extended the frequency approximation approach to the phase domain and thereby further improved the efficiency of information coding in SSVEPs. The proposed system also outperformed the BCIs using code

modulated VEPs.^{19,20} These systems were considered to have highest communication speed in previous BCI studies.¹⁰ For example, the previous record of ITR (144 bits/min) was obtained by a c-VEP-based BCI.²⁰ The present study suggests that the mixed frequency and phase modulation approach could be as efficient as the code modulation approach. A direct comparison between the two methods could be helpful for optimizing the coding methods in VEP-based BCIs.

Although the present system achieved a very high ITR, there is still room for improvement. As shown in Fig. 7, with a shorter time window of 0.6 s, the ITR can further reach 203.56 ± 30.74 bits/min across all subjects and 262.17 bits/min for the best individual (Subjects 8). In real practice, adaptive data length adjustment³⁶ can be applied to reduce the target detection time across trials and subjects. Considering individual difference, selection of subject specific parameters (*e.g.*, channel location, number of harmonics, ban-pass filters, stimulating frequencies) can also be useful for improving individual performance. The classification accuracy could be further improved by incorporating other frequency and phase decoding methods.^{7, 26, 27, 37} Another direction is to increase the number of targets in system design. A wider frequency band and a higher frequency or phase resolution can achieve this goal. For example, a frequency resolution of 0.2Hz was proved feasible in the system using frequency coding.¹⁸ Besides, as discussed by Jia *et al.*²⁹ and Manyakov *et al.*³⁷, the phase resolution could be as low as 60 degrees, which corresponded to six targets per frequency. Further investigations are required to determine an optimal

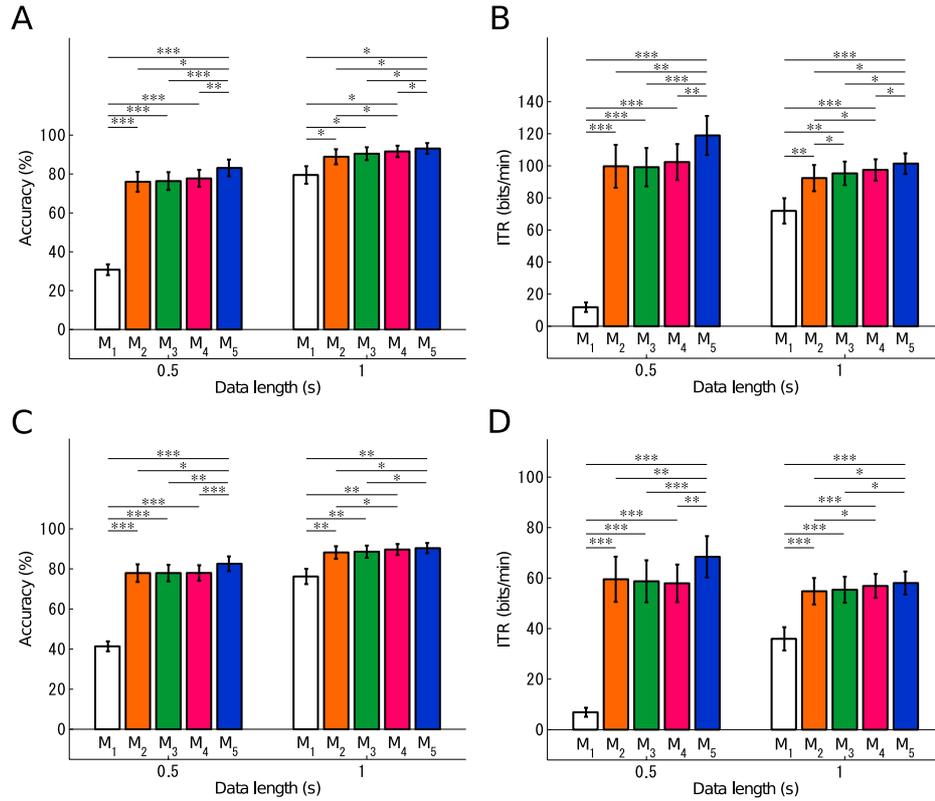


Fig. 8. Target identification accuracy ((A) frequency coding and (C) phase coding) and ITRs ((B) frequency coding and (D) phase coding) calculated by different CCA-based method using data lengths of 0.5 s and 1 s. The asterisks indicate significant difference between different methods (*: $p < 0.05$, **: $p < 0.005$, and ***: $p < 0.0005$).

combination of frequency and phase resolutions in system design. These system parameters should be carefully selected towards an optimal ITR. By combining the above methods, a high-speed BCI that can spell 40 characters per minute could be feasible in online practice.

4.2. The advantages of inclusion of training data

The proposed target identification approach played an important role in the performance improvement. To further explore the underlying mechanisms of the proposed method in characterizing SSVEP signals, this study also compared the proposed target identification method with the standard CCA approach using frequency-coding and phase-coding paradigms separately. Fig. 8 shows target identification accuracy and ITR specific to frequency coding and phase coding using different methods with data lengths of 0.5 s and 1 s. It is evident that the employment of SSVEP training data in the CCA-based approaches can facilitate both

frequency and phase detection. More specifically, the improvement of phase detection seems to be comparable to that of frequency detection. When using a data length of 1 s, the accuracy of phase detection was increased by 14.15% (M₁: 76.20%, M₅: 90.35%, $t = -5.59$, $p < 0.005$), while the accuracy of frequency detection showed an increase of 13.60% (M₁: 79.50%, M₅: 93.10%, $t = -4.64$, $p < 0.05$). Even with a short data length of 0.5 s, the proposed approach can obtain very high classification accuracy, which was significantly higher than the standard CCA-based method (Frequency: 83.10% vs. 30.77%, $t = 13.70$, $p < 10^{-6}$; Phase: 82.55% vs. 41.38%, $t = 14.31$, $p < 10^{-8}$).

It is important to find out why the frequency and phase characters of SSVEPs can be significantly improved by the proposed method. By comparing to other CCA-based methods in the literature, primary advantages of the proposed method can be explained by the following three reasons:

(1) It has been demonstrated that, compared to sine-cosine reference signals, the time domain features can

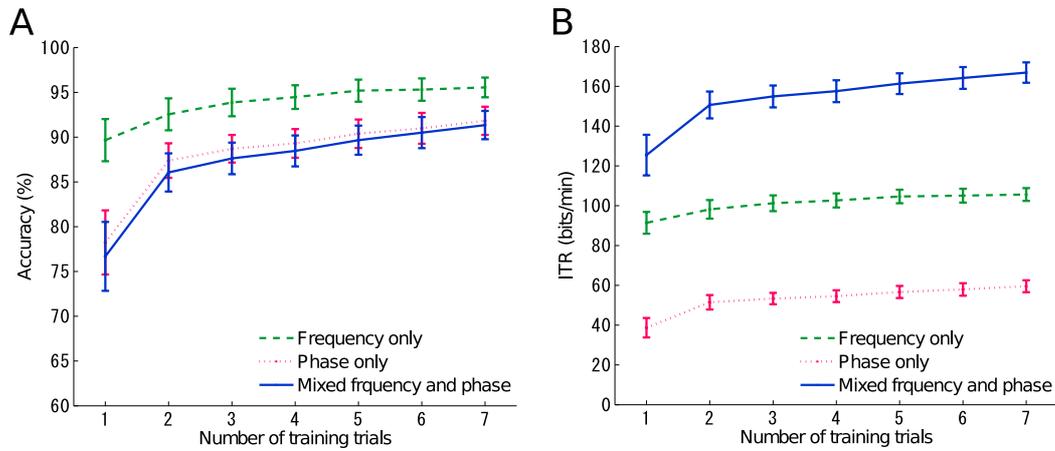


Fig. 9. (A) Target identification accuracy and (B) ITRs as functions of number of training trials.

be better characterized by the SSVEP training data. Zhang *et al.*²⁶ proposed a multi-way CCA method by combining SSVEP training data in frequency detection. The multiset CCA approach has recently been applied to improve the detection of SSVEP.²⁸ Pan *et al.*²⁷ proposed a phase constrained CCA approach by measuring the latency delay in the visual pathway. These methods improved the standard CCA approach. As shown in Fig. 6, the proposed method significantly outperformed the traditional CCA method using sine-cosine reference signals.²⁵

(2) The present method used correlation coefficient instead of the canonical correlation (see Ref. 25 for more details) as features. The canonical correlation could not discriminate the two classes with a 180-degree phase shift, leading to a positive correlation between test data and training reference signals for both classes. The difference between the two classes is presented in weight vectors (W_X and W_Y in Equation (4)). However, correlation coefficients between test data and training reference signals for the two classes show opposite signs. Therefore, the two classes with the same frequency but a 180-degree phase shift can be reliably discriminated with the correlation coefficient calculated with SSVEPs after CCA-based spatial filtering. This process significantly facilitates phase detection.

(3) This study proposed three different strategies to design CCA-based spatial filters for improving the estimation of correlation coefficient between test data and training reference signals. These methods were further combined with the standard CCA method to make final decision in target identification. As shown in

Fig. 6, the combination method was highly efficient and robust for recognizing SSVEPs using mixed frequency and phase coding. An important finding is that the proposed approach can reach high classification accuracy even with a very short data length. Compared with the standard CCA method, the classification accuracy and ITR were significantly improved when using a data length of 0.5 s (Accuracy: 78.98% vs. 26.41%, $t=12.78$, $p<10^{-7}$; ITR: 200.14 bits/min vs. 33.92 bits/min, $t=10.45$, $p<10^{-6}$). The proposed method makes it possible to implement a fast speller that can spell a character in 1-2 seconds.

Note that, the employment of training data in target identification requires additional efforts in data collection before system operation. For frequency detection, the standard CCA method for frequency detection does not require training data. However, since individual variability is relatively high in phase, training data are usually required in phase detection.²⁹ In practice, the number of trials in training data is a key factor for the proposed system. More training data can enhance robustness of system performance, but increase difficulty in system use at the same time. Fig. 9 shows the system performance as a function of the number of training trials. Overall, the performance increased as the size of training data increased. The increase of number of training trials enhanced the performance of phase classification more significantly than that of frequency classification. ITR remained relatively stable when the training data have more than 5 trials (5 trials: 161.34 bits/min; 6 trials: 164.20 bits/min; 7 trials: 166.91 bits/min). Given a trial length of 1.5 s, the training data

with 5 trials for each target can be collected within 5 minutes. If adaptive learning approaches³⁸ are applicable, the training procedure could be further simplified by adopting training data from previous recordings.

4.3. Online implementation

This study estimated the system performance using a simulated online test. This strategy has been widely used in previous BCI studies.^{29, 37, 39} In essence, the simulated online test could be easily converted to an implementation in real time. However, in an online system, BCI performance could be affected by technical issues such as time delays in data transmission, data processing, and feedback presentation. These problems can be solved with real-time platforms and technologies. In addition, online performance could also be affected by user's error-related physiological and psychological changes, as well as distraction from feedback or real-world events during system operation. In this regard, user training could be helpful to improve user's performance in operating an online system. In practice, the real BCI performance needs to be evaluated using a close-loop online system.³⁵

To implement a practical spelling system, the following issues need to be further addressed. First, the high ITR obtained in the present study depends critically on the fast selection speed (1.5 seconds per selection). In an online system, it requires further investigation to prove the feasibility of gaze shifting following a real-time feedback within a 0.5s duration. Second, a BCI for free spelling requires an asynchronous system design. In other words, the system should be capable of discriminating the resting and working states automatically. The ON/OFF switch design⁴⁰ could be a solution to solve this problem. Third, in a VEP BCI, it is important to reduce discomfort and fatigue caused by continuous visual stimulation during system use. The comfortableness of visual stimuli can be improved using different stimulating parameters. For example, visual stimuli using high-frequency flickers⁴¹ and high-duty cycle flickers⁴² have been proved useful to reduce visual discomfort and fatigue. Fourth, a mobile BCI system is required for long-term daily use. The mobile and wireless BCI technology⁴³⁻⁴⁵ can satisfy the requirements of system feasibility and practicality in real-life applications.

4.4. Other applications

This study demonstrates a 32-target spelling program using an SSVEP-based BCI system. Compared to other BCI-based spelling programs such as the P300 speller,¹³ the present system shows higher communication speed and less user variation. Given the spelling speed at 1.5 s per character, the current paradigm can spell 40 characters in a minute. This communication speed makes the proposed system even competitive to other input devices such as eye tracker and touch screen. In addition to the spelling program, the proposed system can also be applied to other applications that require large amount of commands. These applications include phone dialing, complex menu selection, and multi-dimensional movement control. The methodologies developed in this study also provide a practical solution to the newly proposed concept of the brain-display interface technology.⁴⁶

The proposed target identification approach can be adopted to implement SSVEP-based BCIs for other applications that require fewer targets.⁴⁷ For example, a 4-target system can be realized using two frequencies (*e.g.*, 10Hz and 15Hz) and two phases (*e.g.*, 0 and 180 degrees). In this case, data length required for target identification could be reduced to less than 0.5 s. An operation speed at a hundred-millisecond level has potential for many applications under time constraints. For example, during human natural behaviors, visual attention on frequency-tagged objects⁴⁸ could be tracked in near real time. In addition to the communication and control purposes for the disabled patients, the proposed high-speed BCI technology can also be combined with traditional human-computer interfaces to improve behavioral performance of healthy people.^{49, 50}

The methodologies developed in this study can also facilitate visual neuroscience studies that use SSVEP-based frequency tagging (*e.g.*, selective visual attention).^{51, 52} The approximation approach makes it possible to use flexible frequencies and phases to tag multiple targets. The proposed CCA-based target identification method can be applied to improving the efficiency in extracting task-related power and phase modulations. Accordingly, these methods have potential for improving the independent BCI systems based on attentional modulation of SSVEPs.⁵³⁻⁵⁵

5. Conclusion

This study proposed a high-speed brain speller using an SSVEP-based BCI. By adequately addressing technical issues of visual stimulus presentation and target identification, this study showed great potential of the mixed frequency and phase coding approach in implementing a high-speed BCI. As a result, the proposed system that can spell 40 characters per minute (1.5 seconds/character) obtained a record-breaking average ITR of 166.91 bits/min across 13 subjects with a maximum ITR of 192.26 bits/min for a single subject. In addition to the achievement of a high-speed BCI, this study also provides experimental and methodological guidelines for implementing practical applications using the SSVEP-based BCI.

Acknowledgements

Research is supported in part by a gift fund from Swartz Foundation, Office of Naval Research (N00014-08-1215), Army Research Office (under contract number W911NF-09-1-0510), Army Research Laboratory (under Cooperative Agreement Number W911NF-10-2-0022), and DARPA (USDI D11PC20183). M. Nakanishi is also supported by Japan Society for the Promotion of Science.

References

1. J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller and T. M. Vaughan, Brain-computer interfaces for communication and control, *Clin. Neurophysiol.* 113 (2002) 767–791.
2. M. A. Lebedev and M. A. L. Nicolelis, Brain-machine interfaces: past, present and future, *Trends Neurosci.* 29(9) (2006) 536-546.
3. J. J. Daly and J. R. Wolpaw, Brain-computer interfaces in neurological rehabilitation, *Lancet Neurol.* 7(11) (2008) 1032–1043.
4. A. Ortiz-Rosario and H. Adeli, Brain-Computer Interface Technologies: From signal to action, *Rev. Neurosci.* 24(5) (2013) 537-552.
5. W. Y. Hsu, Application of competitive Hopfield neural network to brain-computer interface systems, *Int. J. Neural Syst.* 22(1) (2012) 51–62.
6. M. A. Lopez-Gordo, F. Pelayo, A. Prieto and E. Fernandez, An auditory brain-computer interface with accuracy prediction, *Int. J. Neural Syst.* 22(3) (2012) 1–14.
7. N.V. Manyakov, N. Chumerin, and M.M. Van Hulle, Multichannel decoding for phase-coded SSVEP brain-computer interface, *Int. J. Neural Syst.* 22(5) (2012) 1-7.
8. G.R. Mueller-Putz, C. Pokorny, D.S. Klobassa, and P. Horki, A single-switch brain-computer interface based on passive and imagined movements: towards restoring communication in minimally conscious patients, *Int. J. Neural Syst.* 23(2) (2013) 1-11.
9. J. Li, J. Liang, Q. Zhao, K. Hong, and L. Zhang, Design of wheelchair assistive system directly steered by human thoughts, *Int. J. Neural Syst.* 23(3) (2013) 1-12.
10. G. Rodriguez-Bermudeza, P.J. Garcia-Laencina, and J. Roca-Dorda, Efficient automatic selection and combination of EEG features in least squares classifiers for motor-imagery brain computer interfaces, *Int. J. Neural Syst.* 23(4) (2013) 1-17.
11. R. Fazel-Rezai, B. Z. Allison, C. Guger, E. W. Sellers, S. C. Kleih and A. Kübler, P300 brain computer interface: current challenges and emerging trends, *Front. Neuroeng.* 5(14) (2012).
12. Y. Wang, X. R. Gao, B. Hong, C. Jia and S. Gao, Brain-computer interfaces based on visual evoked potentials - Feasibility of practical system designs, *IEEE EMB Mag.* 27(5) (2008) 64-71.
13. L. A. Farwell and E. Donchin, Talking off the top of your head - toward a mental prosthesis utilizing event-related brain potentials, *Electroencephalogr. Clin. Neurophysiol.* 70(6) (1988) 510-523.
14. J. Jin, B. Z. Allison, E. W. Sellers, C. Brunner, P. Horki, X. Y. Wang and C. Neuper, Optimized stimulus presentation patterns for an event-related potential EEG-based brain-computer interface, *Med. Biol. Eng. Comput.* 49(2) (2011) 181-191.
15. P. Brunner, A. L. Ritaccio, J. F. Emrich, H. Bischof and G. Schalk, Rapid communication with a “P300” matrix speller using electrocorticographic signals (ECoG), *Front. Neurosci.* 5(5) (2011).
16. 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.* 6 (2009) 046002(6pp).
17. M. Nakanishi, Y. Wang, Y. T. Wang, Y. Mitsukura and T. P. Jung, Generating visual flickers for eliciting robust steady-state visual evoked potentials at flexible frequencies using monitor refresh rate, *PLoS One*, accepted.
18. X. Chen, Z. Chen, S. Gao, and X. Gao, A high-ITR SSVEP based BCI speller, *Brain-Comp. Interfaces*, submitted.
19. 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.* 8(2) (2011) 025015(5pp).
20. M. Spuler, W. Rosenstiel and M. Bogdan, Online adaptation of a c-VEP brain-computer interface (BCI) based on error-related potentials and unsupervised learning, *PLoS One* 7(12) (2012) e51077.
21. E. E. Sutter, The brain response interface - communication through visually-induced electrical brain responses, *J. Microcomput. Appl.* 15(1) (1992) 31-45.
22. Y. Wang, Y. T. Wang and T. P. Jung, Visual stimulus design for high-rate SSVEP BCI, *Electron. Lett.* 46(15) (2010) 1057-1058.

23. M. Nakanishi, Y. Wang, Y. T. Wang, Y. Mitsukura and T. P. Jung, An approximation approach for rendering visual flickers in SSVEP-based BCI using monitor refresh rate, in Proc. 35th Ann. Int. IEEE EMBS Conf. (Osaka, Japan, 2013) pp. 2176-2179.
24. G. Bin, X. Gao, Y. Wang, B. Hong and S. Gao, Research frontier: VEP based brain-computer interfaces: time, frequency, and code modulations, *IEEE Comput. Intell. Mag.* 4 (2009) 22–26.
25. Z. Lin, C. Zhang, W. Wu and X. Gao, Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs, *IEEE Trans. Biomed. Eng.* 54(6) (2007) 1172–1176.
26. Y. Zhang, G. Zhou, Q. Zhao, A. Onishi, J. Jin, X. Wang and A. Cichocki, Multiway canonical correlation analysis for frequency components recognition in SSVEP-Based BCIs, in Proc. 18th Int. Conf. Neural Inform. Process. (Shanghai, China, 2011) pp. 287–295.
27. J. Pan, X. Gao, F. Duan, Z. Yan and S. Gao, Enhancing the classification accuracy of steady-state visual evoked potential-based brain-computer interfaces using phase constrained canonical correlation analysis, *J. Neural Eng.* 8(3) (2011) 036027.
28. 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.* 24(2) (2014) 1450013.
29. C. Jia, X. Gao, B. Hong and S. Gao, Frequency and phase mixed coding in SSVEP-based brain-computer interface, *IEEE Trans. Biomed. Eng.* 58(12) (2011) 200–206.
30. D. Regan, *Human Brain electrophysiology: Evoked potentials and Evoked Magnetic Fields in Science and Medicine*, (Elsevier, New York, 1989).
31. D. Regan, Some characteristics of average steady-state and transient response evoked by modulated light. *Electroencephalogr. Clin. Neurophysiol.* 20 (1966) 238–248.
32. H. Spekreijse, O. Estevez and D. Reits, Visual evoked potentials and the physiological analysis of visual processes in man, in *Visual evoked potentials in man: new developments: methodology of pattern visual stimulation: Neurophysiology of visual functions in man: clinical applications* (Clarendon, Oxford, 1977), pp. 16–89.
33. D. H. Brainard, The psychophysics toolbox, *Spat. Vis.* 10(4) (1997) 433–436.
34. F. D. Russo and D. Spinelli, Electrophysiological evidence for an early attentional mechanism in visual processing in humans, *Vision Res.* 39(18) (1999) 2975–2985.
35. P. Yuan, X. Gao, B. Allison, Y. Wang, G. Bin and S. Gao, A study of the existing problems of estimating the information transfer rate in online brain-computer interfaces, *J. Neural Eng.* 10(2) (2013) 026014.
36. Y. Wang, R. Wang, X. Gao, B. Hong and S. Gao, A practical VEP-based brain-computer interface, *IEEE Trans. Neural Syst. Rehabil. Eng.* 14(2) (2006) 234–240.
37. N. V. Manyakov, N. Chumerin, A. Robben, A. Combaz, M. v. Vliet and M. M. V. Hulle, Sampled sinusoidal stimulation profile and multichannel fuzzy logic classification for monitor-based phase-coded SSVEP brain–computer interfacing, *J. Neural Eng.* 10(3) (2013) 036011.
38. P. Shenoy, M. Krauledat, B. Blankertz, R. P. Rao, and K. R. Muller, Towards adaptive classification for BCI, *J. Neural Eng.* 3(1) (2006) R13-R23.
39. Y. Kimura, T. Tanaka, H. Higashi and N. Morikawa, SSVEP-based brain-computer interfaces using FSK-modulated visual stimuli, *IEEE Trans. Biomed. Eng.* 60(10) (2013) 2831-2838.
40. M. Cheng, X. Gao, S. Gao and D. Xu, Design and implementation of a brain-computer interface with high transfer rates, *IEEE Trans. Biomed. Eng.* 49(10) (2002) 1181–1186.
41. Y. Wang, R. Wang, X. Gao and S. Gao, Brain-computer interface based on the high-frequency steady-state visual evoked potential, in Proc. 2005 1st Int. Conf. Neural Interface and Control (Wuhan, China, 2005), pp. 37-39.
42. P. L. Lee, C. L. Yeh, J. Y. S. Cheng, C. Y. Yang and G. Y. Lan, An SSVEP-based BCI using high duty-cycle visual flicker, *IEEE Trans. Biomed. Eng.* 58(12) (2011) 3350-3359.
43. C. T. Lin, L. W. Ko, J. C. Chiou, J. R. Duann, R. S. Huang, S. F. Liang, T. W. Chiu and T. P. Jung, Noninvasive neural prostheses using mobile and wireless EEG, *Proc. IEEE.* 96(7) (2008) 1167-1183.
44. Y. T. Wang, Y. Wang and T. P. Jung, A cell-phone-based brain-computer interface for communication in daily life, *J. Neural Eng.* 8(2) (2011) 025018(5pp).
45. M. Chi, Y. T. Wang, Y. Wang, C. Maier, T. P. Jung, G. Cauwenberghs, Dry and noncontact EEG sensors for mobile brain-computer interfaces, *IEEE Trans. Neural Syst. Rehabil. Eng.* 20(3) (2012) 228–235.
46. Y. Y. Chien, F. C. Lin, C. C. Chou, J. K. Zao, H. Y. Kuo, Y. P. Huang, Y. Wang, T. P. Jung and H. P. D. Shieh, Polychromatic high-frequency steady-state visual evoked potentials for brain-display interaction, *SID Symposium Digest of Technical Papers* 44(1) (2013) 146-149.
47. F. B. Vialatte, M. Maurice, J. Dauwels and A. Cichocki, Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives, *Prog. Neurobiol.* 90(4) (2010) 418–438.
48. S. S. Robertson, S. E. Watamura and M. P. Wilbourn, Attentional dynamics of infant visual foraging, *Proc. Natl. Acad. Sci. USA* 109(28) (2012) 11460-11464.
49. Y. Wang and T. P. Jung, A collaborative brain-computer interface for improving human performance, *PLoS One* 6(5) (2011) e20422.
50. B. J. Lance, S. E. Kerick, A. J. Ries, K. S. Oie and K. McDowell, Brain-computer interface technologies in the coming decades, *Proc. IEEE* 100 (2012) 1585-1599.
51. S. T. Morgan, J. C. Hansen and S. A. Hillyard, Selective attention to stimulus location modulates the steady-state

visual evoked potential, *Proc. Natl. Acad. Sci. USA* 93(10) (1996) 4770–4774.

52. S. K. Andersen and M. M. Muler, Behavioral performance follows the time course of neural facilitation and suppression during cued shifts of featureselective attention, *Proc. Natl. Acad. Sci. USA* 107(31) (2010) 13878–13882.
53. S. P. Kelly, E. C. Lalor, C. Finucane, G. McDarby and R. B. Reilly, Visual spatial attention control in an independent brain-computer interface, *IEEE Trans. Biomed. Eng.* 52(9) (2005) 1588–1596.
54. B. Z. Allison, D. J. McFarland, G. Schalk, S. D. Zheng, M. M. Jackson and J. R. Wolpaw, Towards an independent brain-computer interface using steady state visual evoked potentials, *Clin. Neurophysiol.* 119 (2008) 399–408.
55. D. Zhang, A. Maye, X. Gao, B. Hong, A. K. Engel and S. Gao, An independent brain-computer interface using covert non-spatial visual selective Attention, *J. Neural Eng.* 7(1) (2011) 016010.