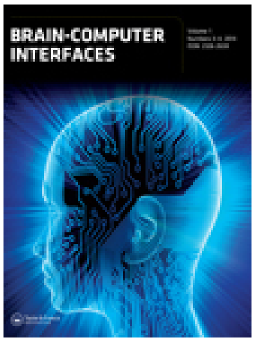

Brain-Computer Interfaces

ISSN: 2326-263X (Print) 2326-2621 (Online) Journal homepage: https://www.tandfonline.com/loi/tbci20

A high-ITR SSVEP-based BCI speller

Xiaogang Chen, Zhikai Chen, Shangkai Gao & Xiaorong Gao

Brain-Computer Interfaces, 2014
Vol. 1, Nos. 3–4, 181–191, http://dx.doi.org/10.1080/2326263X.2014.944469

A high-ITR SSVEP-based BCI speller

Xiaogang Chen, Zhikai Chen, Shangkai Gao and Xiaorong Gao*

Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, People's Republic of China

(*Received 19 June 2014; accepted 10 July 2014*)

Spelling is an important application of brain-computer interfaces (BCIs). Previous BCI spellers were not suited for widespread use due to their low information transfer rate (ITR). In this study, we constructed a high-ITR BCI speller based on the steady-state visual evoked potential (SSVEP). A 45-target BCI speller was implemented with a frequency resolution of 0.2 Hz. A sampled sinusoidal stimulation method was used to present visual stimuli on a conventional LCD screen. The online results revealed that the proposed BCI speller had a good performance, reaching a high average accuracy (84.1% for 2 s stimulation time; 90.2% for 3 s stimulation time) and the corresponding high ITR (105 bits/min for 2 s stimulation time, 82 bits/min for 3 s stimulation time) during the low-frequency stimuli, while 88.7% and 61 bits/min were achieved for a 4 s time window during the high-frequency stimuli.

Keywords: EEG; BCI; SSVEP; speller; ITR

1. Introduction

Brain-computer interfaces (BCIs) provide a new direct communication pathway for individuals with severe neuromuscular diseases.[1] In order to improve the quality of their lives, a variety of attempts to build BCIs have been made in the last decades, such as moving a mouse cursor,[2] spelling characters,[3,4] operating a wheelchair [5] and controlling a home environment.[6] Although the number of BCI application types is extensive, spelling receives more and more attention.

Signals employed for noninvasive BCI systems include slow cortical potentials, P300, sensorimotor rhythms, steady-state visual evoked potentials (SSVEPs), and other possible oscillatory brain activities.[7] So far, most BCI spellers have been set up based on the P300 component, which is an event-related potential (ERP) elicited by an oddball paradigm.[8] In P300 spellers, characters are arranged in a matrix array, where the rows and columns are intensified in a random sequence. Once users focus their attention on the target character, a strong P300 component will be elicited when that row or column is intensified.

Additionally, motor imagery based BCI spellers were studied.[9–11] When individuals imagine moving a certain part of the body, such as a hand, a foot, or the tongue, specific brain rhythms of the motor cortex will be induced.[12,13] These rhythms are detected and mapped onto different commands. Unlike P300 spellers, the number of such available direct commands is no more than five; therefore, we cannot arrange different imaginary states for all characters directly. Therefore, motor imagery spellers usually involve a particular strategy of graphical user interface. For example, Blankertz et al. [10] proposed the Hex-o-Spell, which was built with two imaginary states (such as imaging right hand movement and imaging right foot movement). In addition, motor imagery based BCI systems require a long training period.

Recently, some studies have reported that the BCI spellers can also be realized by the use of SSVEP, which is a periodic response evoked by a visual stimulus oscillating at a constant frequency.[14–17] The majority of the existing SSVEP-based BCI systems are gaze-dependent because of the high signal-to-noise ratio (SNR) in EEG recordings and the high information transfer rate (ITR) desired in BCI performance.[6,18–20] In a typical gaze-dependent SSVEP-based BCI system, multiple targets flicker at different frequencies.[21,22] When users focus their attention on one of the targets, an SSVEP occurs in the visual cortex at the flickering frequency and its harmonics. The target can then be identified by detecting the dominant frequency of the SSVEP. Compared to BCIs based on P300 and motor imagery, this BCI has the advantages of easy system configuration, little user-training and high ITR.[22–25]

In general, SSVEPs in the low- and medium-frequency ranges have larger amplitudes than in the high-frequency range, which makes the detection of SSVEPs easier and produces higher accuracy and higher ITR. However, SSVEPs induced by high-frequency stimuli have specific advantages such as a great decrease in the user's fatigue and risk of photosensitive epileptic seizures.[26–32] Given the respective advantages of low- and high-frequency

*Corresponding author. Email: gxr-dea@tsinghua.edu.cn

stimuli, SSVEPs induced by both low- and high-frequency stimuli were employed as the communication medium respectively in this study.

In addition to considering the choice of stimulus frequencies, it is also important to consider the visual stimulator of SSVEP-based BCI systems. The SSVEP-evoking frequencies can be generated by a light-emitting diode (LED), cathode ray tube (CRT) monitors and liquid crystal display (LCD) screens.[33,34] For LED stimuli, an additional hardware device is required. By contrast, LCD/CRT stimuli, which are much preferred, can be easily implemented and configured as they mainly rely on software developments. To generate stable frequencies on LCD/CRT screens, frame-based coding is the most commonly used method,[15,17,23] in which the number of available stimulus frequencies is limited by the refresh rate of the LCD/CRT. For example, to produce 6 Hz stimulus frequency on a 60 Hz LCD screen, there are 10 frames per period, and suitable frequencies within the alpha frequency band (8–13 Hz) are merely 8.57 Hz (seven frames per period), 10 Hz (six frames per period) and 12 Hz (five frames per period). Therefore, SSVEP-based BCI spellers usually use a decision-tree method or a strategy of moving the cursor position to select the target.[15,17] Though these methods allow users to spell characters within only a small number of stimulus frequencies, these systems cannot achieve a high ITR due to their relatively slow communication speed.

This study aimed to develop a high-ITR SSVEP-based BCI speller, which allowed users to spell one target character per selection. Our BCI speller made use of a novel sampled sinusoidal stimulation method [35] to realize 45 targets on an LCD screen. Both offline analyses and subsequent online tests demonstrated that this BCI speller was full of feasibility and practicality.

2. Experiment 1: SSVEP speller with low-frequency stimulation

2. 1. Materials and methods

2.1.1. Subjects

Ten healthy subjects (three males and seven females, one left-handed, aged 22–28 years) volunteered to participate in this study. All had normal or corrected to normal vision. Each subject signed his or her written informed consent prior to the experiment and was paid for his or her participation. Only two of the subjects had some experience with SSVEP-based BCI experiments, the other subjects were naive to BCI experiments.

2.1.2. Data acquisition

EEG data were acquired using the Neuroscan Synamps2 system at a sampling rate of 1000 Hz. Electrodes were placed according to the international 10-20 system. The reference electrode was placed at the vertex. Electrode impedances were kept below 10 kΩ.

2.1.3. Experimental procedures

Stimuli were presented on a 23.6-inch LCD screen with a resolution of 1920 × 1080 pixels and a refresh rate of 60 Hz at a viewing distance of approximately 70 cm. Matlab and Psychophysics Toolbox Version 3 (PTB-3) [36] controlled the stimuli presentation. In this proposed BCI speller, participants were presented with a 5×9 matrix containing 45 characters (26 letters from the English alphabet, 10 digits, and nine special symbols). When typing, the subjects could delete the misspelled character using the ‘backspace’ key. Each stimulus was presented within a 150 × 150 pixels square and the space between stimuli was 50 pixels wide. Figure 1 shows the distribution of the 45 targets on the computer screen.

The experiment was divided into an offline session and an online session, carried out on two separate days. The offline session consisted of two blocks, each of which contained 45 trials. One trial lasted 6 s. During the first 5 s stimulus period, a red triangle appeared below one flickering target, indicating the focus target. After stimulus offset, the screen was left blank for 1 s before the next trial began. Within a block, the red triangle appeared at each flickering target randomly without repetition. No feedback was provided during the offline experiments.

The online session consisted of two conditions each containing two blocks. Each block consisted of 45 trials. The two conditions differed in the time period for one target detection (3 s for condition 1, 2 s for condition 2).

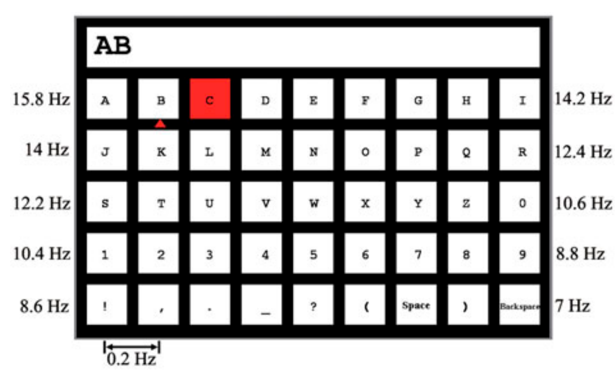

Figure 1. (Color online) Distribution of 45 targets on screen and the corresponding low-frequency stimulation arrangement. Forty-five frequencies, ranging from 7 to 15.8 Hz, with a frequency interval of 0.2 Hz are set on the targets from ‘backspace’ (7 Hz) to ‘A’ (15.8 Hz) in a lateral sequence. For the high-frequency stimulation, these 45 frequencies start from 35.6 to 44.4 Hz with the same frequency resolution and order.

During the stimulus period, a red triangle appeared below one flickering target, indicating the focus target. Subsequently, a feedback indication lasted for 0.3 s before the next trial began. In each block, the subjects were required to input 45 characters in a random order without repetition. During the feedback period, both visual and auditory feedback was available. One of the 45 characters was presented above the matrix, indicating the classification results. If the classification result was consistent with the focus target, one short beep was sounded simultaneously with the visual feedback presentation. Additionally, another target was marked by a red square to indicate the next target.

2.1.4. *Visual stimulation*

In this study we used a novel sampled sinusoidal stimulation method to realize visual stimulus presentation for eliciting SSVEP responses. Forty-five flickering targets were displayed on an LCD screen. The flickering frequencies of the targets were modulated as follows.

Suppose the stimulus frequency of the i-th target is f_i. Then the stimulus intensity of the i-th target on the screen is given by

$$Stim\,(n,f_i) = \frac{1}{2}\{1 + \sin[2\pi f_i(n/R)]\} \tag{1}$$

where sin() generates a sine wave, and$R$ is the screen refresh rate, which can also be considered as the sampling rate of the sinusoidal modulation, and$n$ is the frame index. The dynamic range of $stim\,(n,f_i)$ is from 0 to 1 where 0 represents black and 1 represents white. One advantage of this method is that a stimulus at any frequency up to half the refresh rate can be obtained. Figure 2 shows the time series and frequency spectra of the 7 Hz stimulus, as an example. The frequency spectrum shows a global amplitude maximum at 7 Hz and an additional local maximum in the region of 60±7 Hz. The frequencies of these 45 targets start from 7 to 15.8 Hz with a frequency resolution of 0.2 Hz. Previous studies have reported that SSVEP responses can be obviously observed in this frequency range.[6,18]

2.1.5. *Data analysis*

Canonical correlation analysis (CCA) has been successfully used for channel selection and frequency detection in SSVEP-based BCIs.[23,37] CCA is a multivariable statistical method to measure the linear relationship between two sets of variables which may have some underlying correlation. Considering two sets of variables$X$ and Y, CCA attempts to find a pair of linear transforms W_X and $W_Y$to maximize the correlation between $x = X^T W_Y$ and $y = X^T W_Y$ by solving the following optimization problem:

$$\max_{W_X, W_Y} \rho(x,y) = \frac{E[W_X^T XY^T W_Y]}{\sqrt{E[W_X^T XX^T W_X]E[W_Y^T YY^T W_Y]}} \tag{2}$$

The maximum of ρ with respect to W_X and W_Y is the maximum canonical correlation. Here, Xdenotes the set of multi-channel EEG data and Y_i denotes to the set of reference signals at the i-th stimulus frequency $f_i\,(i = 1, 2, \cdots, K)$, which is constructed by a series of sine-cosine signals. Furthermore, Y has the same length as$X$. The sine-cosine reference signals Y_i are set as

$$Y_i = \begin{pmatrix} \sin(2\pi f_i n) \\ \cos(2\pi f_i n) \\ \vdots \\ \sin(2\pi N_h f_i n) \\ \cos(2\pi N_h f_i n) \end{pmatrix}, \quad n = \frac{1}{f_s}, \frac{2}{f_s}, \cdots, \frac{N}{f_s} \tag{3}$$

where N_h is the number of harmonics, N is the number of sampling points, and f_s is the sampling rate. Here, the reference signals were composed of sinusoid and cosinusoid pairs at the same frequency of the stimulus and its

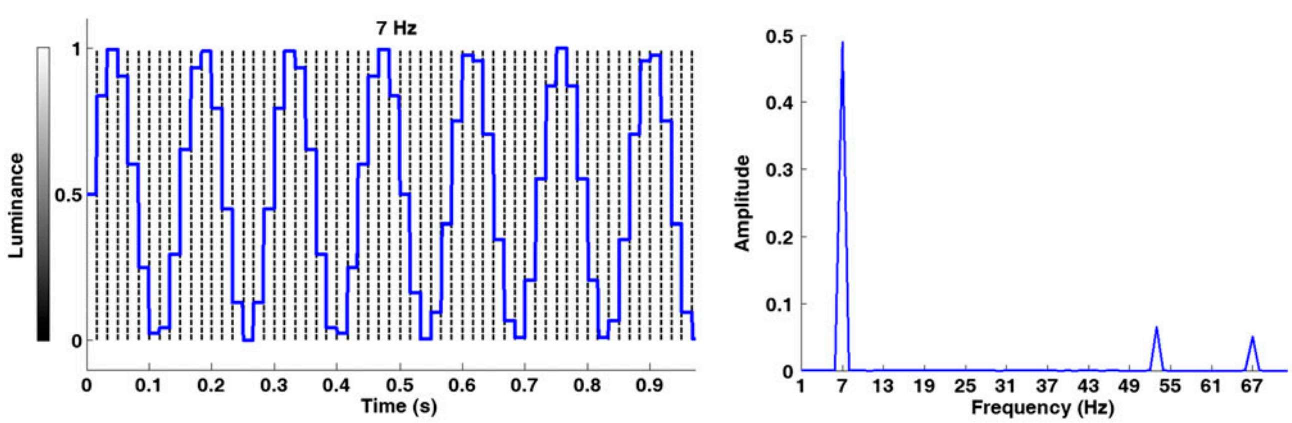

Figure 2. (Color online) Time series and frequency spectra of stimulus signal at 7 Hz.

second harmonics (i.e. $N_h = 2$). CCA calculates the canonical correlation between the set of multi-channel EEG data and sine-cosine reference signals at each stimulus frequency.

The user's command C is recognized as

$$C = \max_{f_i} \rho_i, \quad i = 1, 2, \cdots, K \tag{4}$$

In the offline analysis, the data epochs were extracted in [0s 5s] (time 0 indicated stimulus onset). No epoch artifact rejection procedure was implemented and all epochs were kept for analysis. A fast Fourier transform (FFT) algorithm was used for these 5-s segments starting at the beginning of stimulation and amplitude spectra were extracted for each electrode. The duration of analysis was 5 s so that a frequency resolution of 0.2 Hz could be achieved. Additionally, the SNR was calculated for each stimulus frequency. At each stimulus frequency, the SNR was computed as a ratio between the power of a given frequency and the average power of its eight adjacent frequencies [20]:

$$\mathrm{SNR}_m(f_n) = \mathrm{P}_m(f_n) \Big/ \left(\frac{1}{8} \sum_{\substack{q=-4 \\ q \neq 0}}^{4} \mathrm{P}_m(f_{n+q}) \right) \tag{5}$$

where $\mathrm{P}_m(f_n)$ is the Fourier spectrum at a given frequency f_n in channel m.

The CCA method was applied to each epoch, and the CCA weight matrix W_X of each segment was calculated. The absolute values of weight coefficients were averaged across all subjects and the absolute values of weight coefficients were first averaged across all segments of each subject. Since a larger absolute value made a larger contribution in the weight matrix, optimized channels could be selected according to the average weight matrix.[23] Figure 3 shows the average topographic map of CCA weights. According to the spatial pattern, nine electrode sites (P1, Pz, P4, PO5, POz, PO4, O1, Oz and O2) with strong CCA weights were chosen to calculate the classification accuracy in the offline and online experiments. Additionally, in order to investigate the influence of the length of the time window on the classification accuracy, CCA was subsequently implemented on each single segment of the nine-channel EEG data with different epoch times varying from 1 s to 5 s.

In the online experiments, we used the same classification method as the offline experiments. To evaluate the online performance of the BCI system, we calculated ITR as well as the classification accuracy. The ITR defined by Wolpaw et al. [1,38] was calculated via [39,40]

$$ITR = \frac{60}{T}\left[\log_2 N + p\log_2 p + (1-p)\log_2\left(\frac{1-p}{N-1}\right)\right] \tag{6}$$

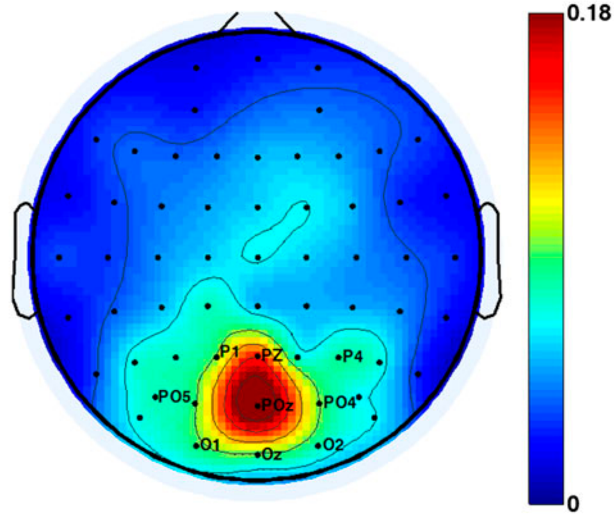

Figure 3. (Color online) The average topographic maps of CCA weights and channels selected in the online experiments.

where N is the number of targets, p is the mean accuracy averaged over all targets and T (seconds/target) is the time for a selection.

2. 2. *Results*

2.2.1. *Offline results*

Figure 4 shows the frequency spectra of one representative subject for each target at the POz electrode site with a frequency resolution of 0.2 Hz. Visual stimulation resulted in a clear peak on each of the stimulus frequencies as well as on the harmonics. These results indicate that the brain synchronized precisely with the visual stimulation [41] and further suggest that each stimulus frequency can be successfully generated by the sampled sinusoidal stimulation.

In SSVEP-based BCI systems, the classification accuracy is mainly influenced by the SSVEP strength, the SNR, and the difference in the properties of the stimuli. High SSVEP SNR can enhance the classification accuracy.[42] Therefore, to further investigate whether the SSVEP responses could be applied to practical BCI systems, we calculated the average SSVEP SNR across the above nine electrode sites for each stimulus frequency. Figure 5 shows the mean SNR of the SSVEP responses averaged across all 10 subjects. Due to the large value of SNR at all stimulus frequencies, the SSVEP responses were high enough for practical BCI use.

In order to choose an optimal time window, we studied the influence of the length of time window on the classification accuracy. Figure 6 shows the classification accuracies for each subject with respect to different lengths of time window. It can be seen that classification

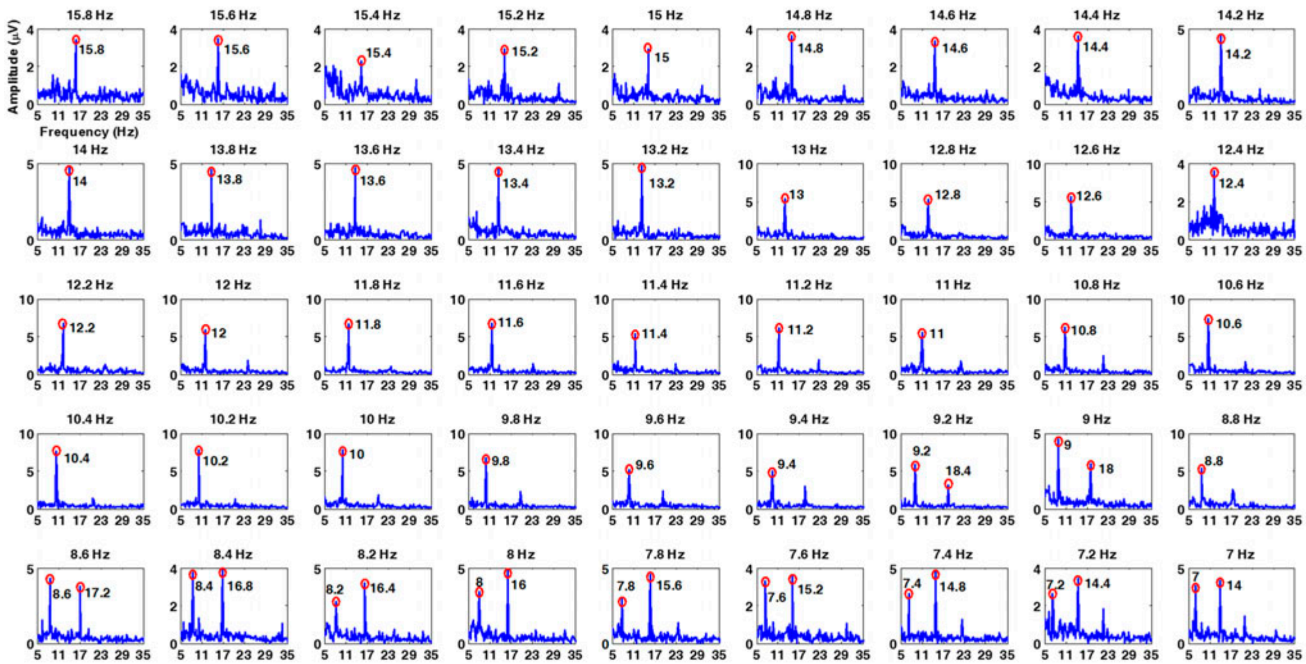

Figure 4. (Color online) Frequency spectra of each single stimulus frequency from the POz electrode site. The red circles represented the corresponding main frequency components.

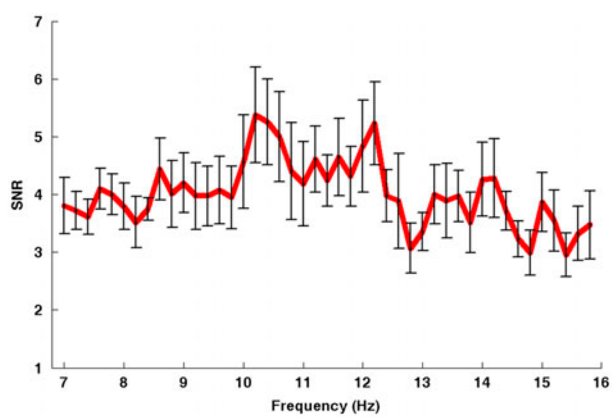

Figure 5. (Color online) Mean SNR of low-frequency SSVEP responses averaged across all subjects. Error bars represented standard errors of the means (SEM).

accuracy increased as the time window grew until it reached a plateau. The average accuracies for all subjects were 80.2%, 90.7%, 94.7% and 96.7% for time-window lengths of 2, 3, 4 and 5 s, respectively. Although longer time windows improved the SNR of the SSVEP, and thus enhanced the overall classification accuracy, there should be a trade-off between time-window length/accuracy and ITR.[43] Given this point, both 2 s and 3 s time windows were chosen for the following online studies.

2.2.2. Online results

Our online BCI system only made use of the nine electrode sites mentioned above (Figure 3) with time windows of 2 and 3 s. Table 1 lists the online results for each subject. The average accuracy for the 2 s time window was 84.1% and the corresponding ITR was 105 bits/min, while the vales for the 3 s time window were 90.2% and 82 bits/min. As we can see, the 2 s time window resulted in higher average ITR while the 3 s time window resulted in higher average accuracy, which suggested that the trade-off between the time-window length/accuracy and ITR should be carefully considered. These results revealed that this BCI speller's performance was good enough to be used in practice.

3. Experiment 2: SSVEP speller with high-frequency stimulation

3. 1. Materials and methods

3.1.1. Subjects

Ten healthy individuals (eight males and two females, one left-handed, aged 20–28 years) volunteered to participate in the online experiments. Only five of them took part in the offline experiments too. All had normal or corrected to normal vision. The subjects gave written informed consent and were paid for their participation. Six subjects had no prior experience with BCI, while four had previous experience with an SSVEP-based BCI.

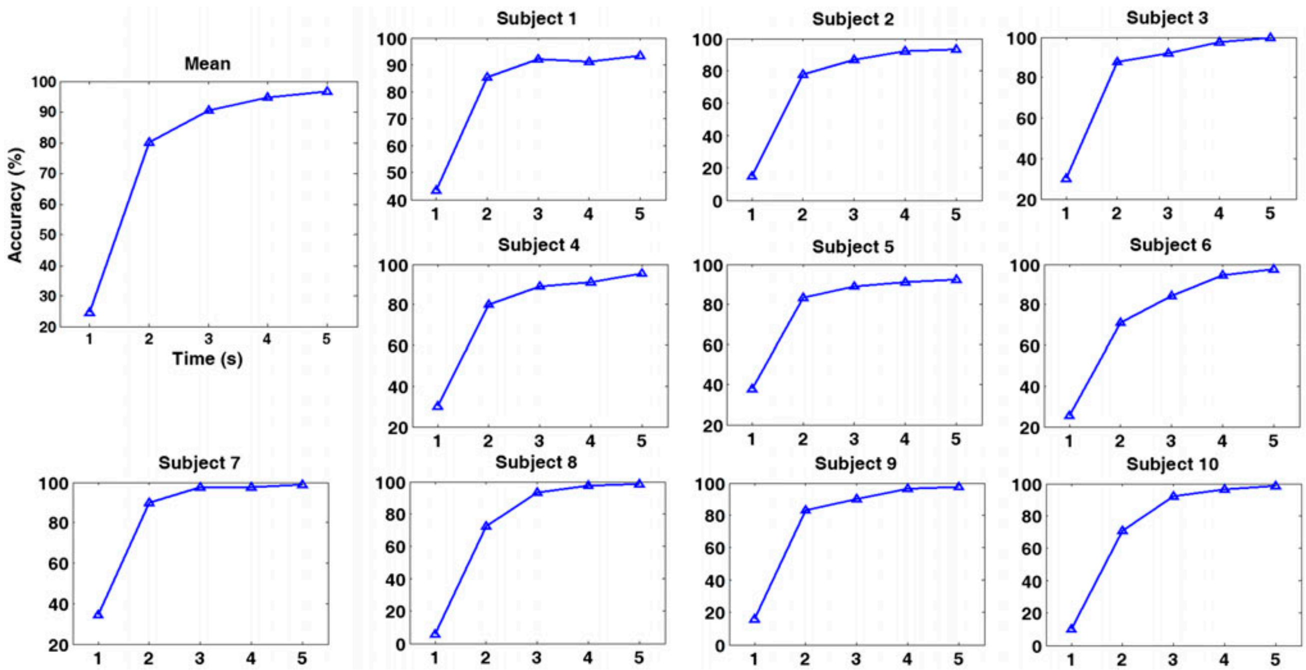

Figure 6. (Color online) Relationship between classification accuracy and the length of time window during low-frequency stimulation.

Table 1. Online results of the SSVEP speller with low-frequency stimulation.

	2 s		3 s	
Subject	Accuracy (%)	ITR (bits/min)	Accuracy (%)	ITR (bits/min)
S1	86.7	110	92.2	85
S2	78.9	94	87.8	78
S3	91.1	119	88.9	80
S4	81.1	98	90.0	81
S5	84.4	105	83.3	72
S6	76.7	90	87.8	78
S7	95.6	130	98.9	97
S8	81.1	98	92.2	85
S9	85.6	107	95.6	91
S10	80.0	96	85.6	75
Average (mean ± SEM)	84.1 ± 1.8	105 ± 3.8	90.2 ± 1.5	82 ± 2.5

3.1.2. Visual stimulation and procedures

All of the visual stimulation and procedures were identical to those of Experiment 1, except that the screen refresh rate was 120 Hz and the frequencies of the 45 targets ranged from 35.6 to 44.4 Hz with a frequency resolution of 0.2 Hz. Some previous work suggested that the frequency range around 40 Hz was a good choice due to good classification accuracy and barely visible flickering.[27]

3.1.3. Data acquisition and analysis

EEG acquisition and analysis were the same as in Experiment 1.

3. 2. Results

3.2.1. Offline results

Figure 7 shows the mean SNR of the high-frequency SSVEP responses averaged across all five subjects. The SNR values of SSVEP responses in the low- and high-frequency ranges are almost on the same level (compare Figure 5 and Figure 7). These results demonstrated that the SNR values of the high-frequency stimulation are high enough to be used in a practical SSVEP-based BCI system.

The accuracies with respect to different lengths of the time window are shown in Figure 8. It can be seen that the accuracy increased as the length of the time window increased. The results showed tendencies similar to

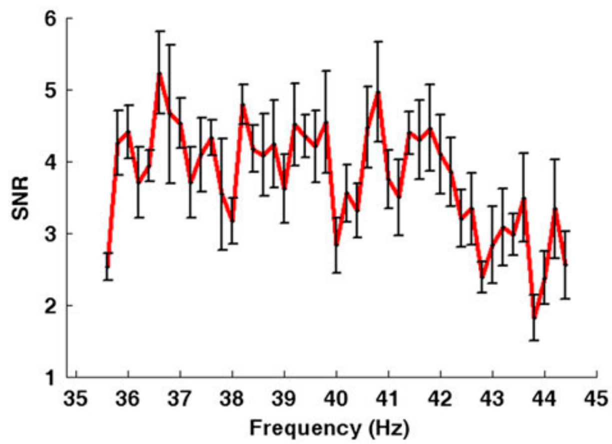

Figure 7. (Color online) Mean SNR of high-frequency SSVEP responses averaged across all subjects. Error bars represented standard errors of the means (SEM).

the low-frequency stimulation. Namely, the longer time window facilitated the frequency detection and improved the classification accuracy. Furthermore, this speller can achieve an average accuracy of 87.6% when the time window reaches 4 s. These results also supported the feasibility of a high-frequency SSVEP speller.

3.2.2. Online results

To confirm whether the high-frequency stimulation could be utilized for a multi-target SSVEP speller, we conducted online experiments with 10 participants. The online experiments also used the aforementioned nine electrode sites with a 4 s time-window length. The results of the online experiments are summarized in Table 2. The average ITR, over 10 participants, was 61 bits/min with an average accuracy of 88.7%. Notably, the classification accuracies of participants S2′ and S3′ were up to 98.9%. The experimental results demonstrated that this speller with high-frequency stimulation could be used as a practical BCI speller.

4. Discussion

In this study, we proposed and implemented a high-ITR SSVEP-based BCI speller which can be practical with both low-frequency stimulus and high-frequency stimulus. From the results of the offline and online experiments, we demonstrated the feasibility of this proposed BCI speller.

4.1. Information transform rate

ITR is a representative metric that has been widely used to assess the overall performance of BCI systems. In order to make the BCI system practical in reality, a high ITR is essential. According to formula (6), the value of ITR depends on the number of targets, classification accuracy, and speed. Thus, there are three ways to obtain a high ITR: (1) by increasing the number of targets,

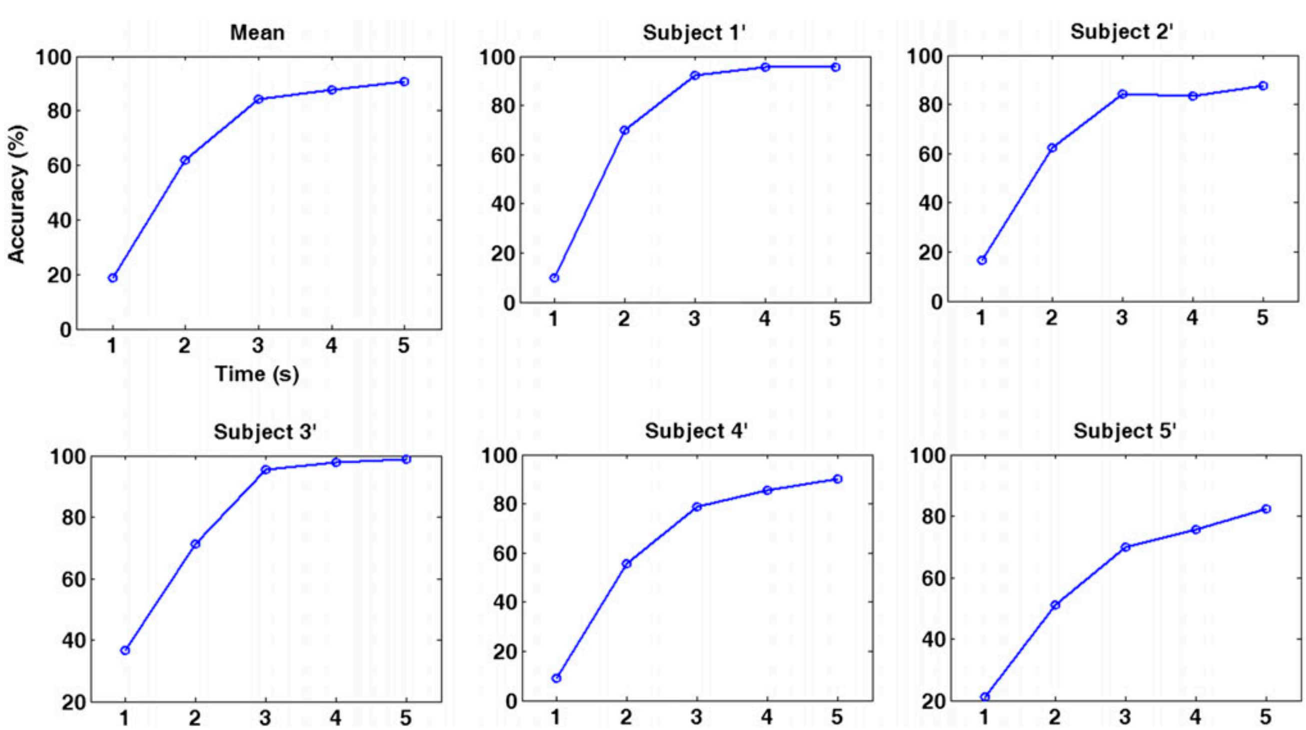

Figure 8. (Color online) Relationship between classification accuracy and the length of the time window during high-frequency stimulation.

Table 2. Online results of the SSVEP speller with high-frequency stimulation.

	Classification accuracy (%)	ITR (bits/min)
S1′	90.0	62
S2′	98.9	75
S3′	98.9	75
S4′	87.8	60
S5′	73.3	45
S6′	88.9	61
S7′	87.8	60
S8′	87.8	60
S9′	87.8	60
S10′	85.6	57
Average (mean ± SEM)	88.7 ± 2.3	61 ± 2.7

(2) by improving the accuracy of target selection, and (3) by decreasing the time needed to recognize each target. Additionally, a large number of targets makes sense in some particular situations such as a practical speller application, since more targets can provide more information. However, when using the 'on/off' visual stimulation to realize target presentation, the number of targets is usually limited by the monitor refresh rate. The top number of SSVEP-BCI targets displayed on a computer screen was 16 in previously published systems. The highest average ITR of an SSVEP-based BCI systems reported so far has reached 75.4 bits/min.[18] Both offline and online results indicated that the sampled sinusoidal stimulation could efficiently mitigate the problem of frequency limitation caused by the traditional 'on/off' visual stimulation. Furthermore, it is worth mentioning that the average ITR in this study reached 105 bits/min (low-frequency condition). Therefore, this proposed BCI speller is capable of providing users with high ITR due to its large number of target characters and thus allows users to select each character in only one step, even in the high-frequency condition.

4.2. *Stimulus frequency*

Despite the fact that the neural mechanisms behind SSVEP are not yet fully understood, the dependence of SSVEP amplitude on stimulus frequency has been investigated in a number of studies,[26,41,44] and is usually used to guide the selection of stimulus frequencies. It is widely acknowledged that the stimulus frequencies used in SSVEP studies can be divided into three main frequency ranges: low- (up to 12 Hz), medium- (12–30 Hz) and high-frequency ranges (>30 Hz).[41] Generally, the SSVEP responses show a global amplitude maximum in the low-frequency range and additional local maxima in the medium- and high-frequency ranges.[41,44] Additionally, the amplitude of the SSVEP significantly decreases in the high-frequency range; the higher the SSVEP amplitude, the more easily it is detected. For this reason, the majority of SSVEP-based BCI systems have focused on low- and medium-frequency ranges for easier SSVEP detection.[6,18,20,23] However, no consensus exists on the optimal stimulus frequencies for SSVEP-based BCI systems. In contrast to previous low-frequency SSVEP BCI systems, the average ITR reported here was 105 bits/min, which confirmed that our system was a high-performance speller.

A stimulus of low or medium frequency can evoke a strong SSVEP response but is very annoying to the user and carries an increased risk of triggering a photosensitive epileptic seizure. A probable solution addressing this problem is to adopt high-frequency stimulation; however, this has received little attention so far.[26–29, 31] Researchers might be concerned that high-frequency SSVEP would be too weak to detect. However, in this study we found that the SNR curves of SSVEP were similar in the low- and high-frequency ranges, which was consistent with a previous study.[26] Additionally, our proposed SSVEP speller, using high-frequency stimulation, also resulted in high performance. In comparison with previous studies on high-frequency stimulation, [27,29,31,45] the classification accuracy and ITR of our SSVEP speller were much better. Furthermore, the frequency resolution with the high-frequency stimulus provided by our speller was 0.2 Hz, the highest SSVEP frequency resolution provided as far. It is important to investigate the detectable frequency resolution of SSVEP responses, which leaves us more choices in the range of available stimulus frequencies.

4.3. *Comparison with other BCI spellers*

As mentioned above, a BCI speller can be also based on the P300 potential or motor imagery. However, such BCI systems generally require a training session. Additionally, to produce sufficiently strong signals, these systems require significant mental effort, which means that users may easily become fatigued. Additionally, they cannot achieve high ITR. By contrast, an

SSVEP speller has the advantage of little training and high ITR.

Low- and medium-frequency stimulations are often utilized in SSVEP spellers, and the number of targets on the computer screen is dramatically limited due to the screen refresh rate. For example, Cecotti realized an SSVEP speller with a three-step decision tree design, in which a target character was selected in at least three steps.[15] Volosyak developed an SSVEP speller with a virtual keyboard and five-command stimuli that were flickered with different frequencies and were used to control the cursor movement. The target character was selected by moving the cursor to the corresponding position to confirm the desired character.[17] Although these stratagems have several potential advantages, including a smaller number of stimulus frequencies and a high classification accuracy, the time for a target character selection will be longer than the directly coded systems and therefore result in low ITR. In our study, the number of stimulus frequencies is equal to the total number of displayed characters so that users can spell each target character in a single step. More recently, Hwang et al. [19] proposed an SSVEP-based BCI spelling system using custom-built LED-based hardware. There was a one-to-one mapping between stimuli and characters, which allowed users to spell one target character per target selection. The system achieved an average accuracy of 87.58% and an ITR of 40.72 bits/min for six healthy participants. In contrast, our proposed speller does not require an additional hardware device (e.g. an LED keyboard) for the stimuli presentation, which makes it easy to realize different frequency ranges of SSVEP stimulations and even different stimulus shapes. Additionally, our system contained a much larger number of targets (45 targets) than their system (30 targets) and also identified the user's intended targets in shorter periods (2 or 3 s) compared to their system (4–7 s). Furthermore, our speller had good performance not only with low-frequency stimuli but also with high-frequency stimuli, which primarily results from the sampled sinusoidal stimulation and results in a larger number of targets and a smaller frequency interval.

4.4. Future work

This study developed a high-ITR BCI speller based on the sampled sinusoidal stimulation method. Currently, the majority of the existing SSVEP-based BCI systems are synchronous, in which users have to follow the predefined pace of the BCI system.[6,18,20,23] To further improve the practicability of our BCI system, in future work a multi-command, real-time and asynchronous BCI system will be built. For asynchronous SSVEP-based BCI systems, a critical problem is to discriminate the control and idle states. The ON/OFF switch design could be used to provide a good solution for this problem.[21] According to the SNR difference between the spontaneous EEG and the evoked EEG, the optimal threshold for each stimulus frequency can be obtained.

Additionally, this current study evaluated the performance of our speller using cued target sequences, which could facilitate a visual search of the target location. Another future research direction will be to assess the performance of our speller with free-spelling in a real world scenario.[46] In the free-spelling condition, the external cue is eliminated. The subjects could spell whatever they want and they are only instructed to delete the misspelled character by choosing the 'backspace' key.

Finally, it is important to reduce interference caused by adjacent stimuli during practical use. Furthermore, the familiar keyboard layouts could facilitate visually finding target characters.[19] Future work will focus on the layout of the character matrix and stimulation frequencies.

5. Conclusion

In this study we adopted the sampled sinusoidal stimulation method to build an SSVEP speller that was capable of spelling each target character in only one step. The proposed SSVEP speller comprised 45 targets with a frequency resolution of 0.2 Hz. Both offline and online experiments were conducted to verify the feasibility of the proposed spelling system. The results of the online experiments in the low-frequency condition achieved an average ITR of 105 bits/min. Additionally, the online results of the SSVEP speller in the high-frequency stimulus condition revealed that an average classification accuracy of 88.7% and an ITR of 61 bits/min could be achieved.

Acknowledgements

This work was supported by the National Basic Research Program (973) of China (No. 2011CB933204), National Natural Science Foundation of China under Grant 90820304, 91120007, Chinese 863 Project: 2012AA011601. We would like to thank Ian Daly of the University of Reading for his help in language editing.

Notes on contributors

Xiaogang Chen received the B.E. degree in biomedical engineering from Xianning College, Xianning, China, in 2008 and the M.E. degree in biomedical engineering from Hebei University of Technology, Tianjin, China, in 2011. He is currently working toward the Ph.D. degree in the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing, China. His current research interests include brain-computer interfaces.

Zhikai Chen received the B.E. degree in biomedical engineering from Beijing Institute of Technology, Beijing, China, in

2008 and the M.E. degree in biomedical engineering from Tsinghua University, Beijing, China, in 2013. His current research interests include brain-computer interfaces and binocular rivalry.

Shangkai Gao received the B.S. degree in electrical engineering and the M.E. degree in biomedical engineering from the Department of Electrical Engineering, Tsinghua University, Beijing, China, in 1970 and 1982, respectively. She is currently a Professor in the Department of Biomedical Engineering, Tsinghua University. Her research interests include biomedical signal processing and medical imaging, especially the study of brain-computer interfaces.

Xiaorong Gao received the B.S. degree in biomedical engineering from Zhejiang University, Hangzhou, China, in 1986, the M.S. degree in biomedical engineering from Peking Union Medical College, China, in 1989, and the Ph.D. degree in biomedical engineering from Tsinghua University, Beijing, China, in 1992. He is currently a Professor in the Department of Biomedical Engineering, Tsinghua University. His research interests include biomedical signal processing.

References

[1] Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain-computer interfaces for communication and control. Clin Neurophysiol. 2002; 113:767–791.

[2] Fabiani GE, McFarland DJ, Wolpaw JR, Pfurtscheller G. Conversion of EEG activity into cursor movement by a brain-computer interface (BCI). IEEE Trans Neural Syst Rehabil Eng. 2004; 12:331–338.

[3] Hong B, Guo F, Liu T, Gao X, Gao S. N200-speller using motion-onset visual response. Clin Neurophysiol. 2009; 120:1658–1666.

[4] Brunner P, Joshi S, Briskin S, Wolpaw JR, Bischof H, Schalk G. Does the ‘P300’ speller depend on eye gaze? J Neural Eng. 2010; 7:056013.

[5] Leeb R, Friedman D, Müller-Putz GR, Scherer R, Slater M, Pfurtscheller G. Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with tetraplegics. Comput Intell Neurosci. 2007; 79642:1–8.

[6] Gao X, Xu D, Cheng M, Gao S. A BCI-based environmental controller for the motion-disabled. IEEE Trans Neural Syst Rehabil Eng. 2003; 11:137–140.

[7] Lotte F, Congedo M, Lécuyer A, Lamarche F, Arnaldi B. A review of classification algorithms for EEG-based brain-computer interfaces. J Neural Eng. 2007; 4:R1–R13.

[8] Donchin E, Spencer KM, Wijesinghe R. The mental prosthesis: assessing the speed of a P300-based brain-computer interface. IEEE Trans Rehabil Eng. 2000; 8:174–179.

[9] Scherer R, Muller GR, Neuper C, Graimann B, Pfurtscheller G. An asynchronously controlled EEG-based virtual keyboard: improvement of the spelling rate. IEEE Trans Biomed Eng. 2004; 51:979–984.

[10] Blankertz B, Dornhege G, Krauledat M, Schröder M, Williamson J, Murray-Smith R, Müller KR. The Berlin brain-computer interface presents the novel mental typewriter Hex-O-Spell. Proc 3rd International Brain-Computer Interface Workshop and Training Course. Austria: Graz; 2006; p. 108–109.

[11] Neuper C, Müller-Putz GR, Scherer R, Pfurtscheller G. Motor imagery and EEG-based control of spelling devices and neuroprostheses. In: Neuper C, Klimesch W, editors. Event-related dynamics of brain oscillations. Amsterdam: Elsevier; 2006. p. 393–409.

[12] Pfurtscheller G, Neuper C. Motor imagery and direct brain-computer communication. Proc IEEE. 2001; 89:1123–1134.

[13] Pfurtscheller G, Brunner C, Schlogl A and Lopes da Silva FH. Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. NeuroImage. 2006; 31:153–159.

[14] Hillyard S, Hinrichs H, Tempelmann C, Morgan S, Hansen J, Scheich H, Heinze H-J. Combining steady-state visual evoked potentials and fMRI to localize brain activity during selective attention. Hum Brain Mapp. 1997; 5:287–292.

[15] Cecotti H. A self-paced and calibration-less SSVEP-based brain-computer interface speller. IEEE Trans Neural Syst Rehabi. Eng. 2010; 18:127–133.

[16] Allison B, Lüth T, Valbuena D, Teymourian A, Volosyak I, Gräser A. BCI Demographics: How many (and what kinds of) people can use an SSVEP BCI? IEEE Trans Neural Syst Rehabil Eng. 2010; 18:107–116.

[17] Volosyak I. SSVEP-based Bremen–BCI interface—boosting information transfer rates. J Neural Eng. 2011; 8:036020.

[18] Wang Y, Wang YT, Jung TP. Visual stimulus design for high-rate SSVEP BCI. Electron Lett. 2010;46:1057–1058.

[19] Hwang HJ, Lim JH, Jung YJ, Choi H, Lee SW, Im CH. Development of an SSVEP-based BCI spelling system adopting a QWERTY-style LED keyboard. J Neurosci Methods. 2012; 208:59–65.

[20] Chen X, Chen Z, Gao S, Gao X. Brain–computer interface based on intermodulation frequency. J Neural Eng. 2013;10:066009.

[21] Cheng M, Gao X, Gao S. Design and implementation of a brain–computer interface with high transfer rates. IEEE Trans Biomed Eng. 2002;49:1181–1186.

[22] Wang Y, Gao X, Hong B, Jia C, Gao S. Brain-computer interfaces based on visual evoked potentials. IEEE Eng Med Biol Mag. 2008; 27:64–71.

[23] Bin G, Gao X, Yan Z, Hong B, Gao S. An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method. J Neural Eng. 2009; 6:046002.

[24] Jia C, Gao X, Hong B, Gao S. Frequency and phase mixed coding in SSVEP-based brain–computer interface. IEEE Trans Biomed Eng. 2011;58:200–206.

[25] Yan Z, Gao X, Gao S. Right-and-left visual field stimulation: a frequency and space mixed coding method for SSVEP based brain-computer interface. Sci China Inf Sci. 2011; 54:2492–2498.

[26] Wang Y, Wang R, Gao X, Gao S. Brain-computer interface based on the high-frequency steady-state visual evoked potential. Proc. 1st Int. Conf. on Neural Interface and Control. China: Wuhan; 2005. p. 37–39.

[27] Hoffmann U, Fimbel EJ, Keller T. Brain-computer interface based on high frequency steady-state visual evoked potentials: a feasibility study. Proc 4th Int IEEE/EMBS Conf Neural Eng. Turkey: Antalya; 2009; p. 466–469.

[28] Molina GG, Ibanez D, Mihajlovic V, Chestakov D. Detection of high frequency steady state visual evoked potentials for brain–computer interfaces. Proc 17th European Signal Processing Conference Scotland: Glasgow; Aug 24–28, 2009. p. 646–650.

[29] Volosyak I, Valbuena D, Luth T, Malechka T, Graser A. BCI demographics II: how many (and what kinds of) people can use a high-frequency SSVEP BCI? IEEE Trans Neural Syst Rehabil Eng. 2011; 19:232–239.

[30] Wu CH, Chang HC, Lee PL, Li KS, Sie JJ, Sun CW, Yang CY, Li PH, Deng HT, Shyu KK. Frequency recognition in an SSVEP-based brain computer interface using empirical mode decomposition and refined generalized zero-crossing. J Neurosci Methods. 2011; 196:170–181.
[31] Diez PF, Mut VA. Avila Perona EM, Laciar Leber E. Asynchronous BCI control using high-frequency SSVEP. J Neuroeng Rehabil. 2011;8:1–8.
[32] Gao S, Wang Y, Gao X, Hong B. Visual and auditory brain-computer interface. IEEE Trans Biomed Eng. 2014; 61:1436–1447.
[33] Vialatte FB, Maurice M, Dauwels J, Cichocki A. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. Prog Neurobiol. 2010; 90:418–438.
[34] Wu Z, Lai Y, Wu D, Yao D. Stimulator selection in SSVEP-based BCI. Med Eng Phys. 2008; 30:1079–1088.
[35] Manyakov NV, Chumerin N, Robben A, Combaz A, van Vliet, Van Hulle MM. Sampled sinusoidal stimulation profile and multichannel fuzzy logic classification for monitor-based phase-coded SSVEP brain–computer interfacing. J Neural Eng. 2013; 10:036011.
[36] Brainard DH. The psychophysics toolbox. Spatial Vis. 1997; 10:433–436.
[37] Lin Z, Zhang C, Wu W, Gao X. Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. IEEE Trans Biomed Eng. 2006; 53:2610–2614.
[38] Wolpaw JR, Ramoser H, McFarland DJ, Pfurtscheller G. EEG-Based communication: improved accuracy by response verification. IEEE Trans Rehabil Eng. 1998; 6:326–333.
[39] Billinger M, Daly I, Kaiser V, Jin J, Allison BZ, Müller-Putz GR, Brunner R. Is it significant? Guidelines for reporting BCI Performance. In: Allison BZ, Dunne S, Leeb R, Millán JDR, Nijholt A, editors. Toward practical brain-computer interfaces: Biological and medical physics. Springer: Berlin; 2013. p. 333–354.
[40] Yuan P, Gao X, Allison B, Wang Y, Bin G, Gao S. A study of the existing problems of estimating the information transfer rate in online brain–computer interfaces. J Neural Eng. 2013; 10:026014.
[41] Regan D. Human Brain Electrophysiology: Evoked Potentials and evoked magnetic fields in science and medicine. New York: Elsevier; 1989.
[42] Zhu D, Bieger J, Molina GG, Aarts RM. A survey of stimulation methods used in SSVEP-based BCIs. Comput Intell Neurosci. 2010; 1:702357.
[43] Xu M, Chen L, Zhang L, Qi H, Ma L, Tang J, Wan B, Ming D. A visual parallel-BCI speller based on the time-frequency coding strategy. J Neural Eng. 2014; 11:026014.
[44] Herrmann CS. Human EEG responses to 1-100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. Exp Brain Res. 2001; 137:346–353.
[45] Diez PF, Torres Müller SM, Mut VA, Laciar E, Avila E, Bastos-Filho T, Sarcinelli-Filho M. Commanding a robotic wheelchair with a high-frequency steady-state visual evoked potential based brain-computer interface. Med Eng Phys. 2013; 5:1155–1164.
[46] Spüler M, Rosenstiel W, Bogdan M. Online adaptation of a c-VEP brain-computer interface (BCI) based on error-related potentials and unsupervised learning. PloS One. 2012; 7:e51077.