

A Practical VEP-Based Brain-Computer Interface

Yijun Wang, Ruiping Wang, Xiaorong Gao, Bo Hong, and Shangkai Gao

Abstract—This paper introduces the development of a practical brain-computer interface at Tsinghua University. The system uses frequency-coded steady-state visual evoked potentials to determine the gaze direction of the user. To ensure more universal applicability of the system, approaches for reducing user variation on system performance have been proposed. The information transfer rate (ITR) has been evaluated both in the laboratory and at the Rehabilitation Center of China, respectively. The system has been proved to be applicable to >90% of people with a high ITR in living environments.

Index Terms—Brain-computer interface (BCI), information transfer rate (ITR), steady-state visual evoked potential (SSVEP).

I. INTRODUCTION

In recent years, brain-computer interfaces (BCI) based on noninvasive scalp electroencephalography (EEG) have become an increasingly active research area. Event-related potentials, mu and beta rhythms, event-related synchronization and desynchronization, slow cortical potentials, and visual evoked potentials (VEP) are commonly used signals in EEG-based BCIs [1]–[3]. Different from other systems, the VEP-based BCI is considered a dependent BCI because the generation of the VEP depends on the control of eye movements via the output pathways of cranial nerves and extraocular muscles. Therefore, for the few people with severe neuromuscular disabilities, who may even lack the output channel of extraocular muscle control, this BCI is inapplicable. However, for most people, the VEP-based BCI is more feasible than other systems. It has the advantages of a high information transfer rate (ITR), convenient system preparation, and little user training.

In 1992, Sutter developed a BCI based on transient VEP (TVEP), which was called the brain response interface (BRI) [4]. A subject with implanted electrodes reached communication rates of 10 to 12 words/minute. In other dependent VEP-based systems, two different methods utilizing steady-state VEP (SSVEP) as the communication medium were employed. One approach was to train the user to self-regulate the amplitude of SSVEP through slight shift of gaze [5]. The other approach was to use frequency detection of frequency-coded SSVEP to determine the gaze direction of the user [5]–[7]. We employ the latter approach to implement our system. Different from the above approaches, Kelly *et al.* recently presented an independent VEP-based BCI controlled by left/right visual spatial attention [8].

As a common problem in BCI research, a high user variation also exists in the SSVEP-based system [1], [6]–[8]. In our previous study, the average ITR was about 25 bits/min; whereas, five of the 13 subjects had an average ITR of only 3.05 bits/min [6]. Realizing this problem, we consciously focused our study on the problem of user applicability. Our investigation has demonstrated that user variation can be significantly reduced by carefully selecting the *channel location*, the *stimulus frequency*, and the *speed of selection*.

Manuscript received July 23, 2005; revised March 18, 2006; accepted March 21, 2006. This work was supported in part by the National Natural Science Foundation of China under Grant 60318001 and Grant 60205003, in part by the Key Project of the Chinese Ministry of Health under Grant 2004BA706B13, and in part by the Tsinghua-Yu-Yuan Medical Sciences Fund.

The authors are with the Department of Biomedical Engineering, Tsinghua University, Beijing 100084, China (e-mail: wyj97@mails.tsinghua.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2006.875576

SSVEP-Based BCI

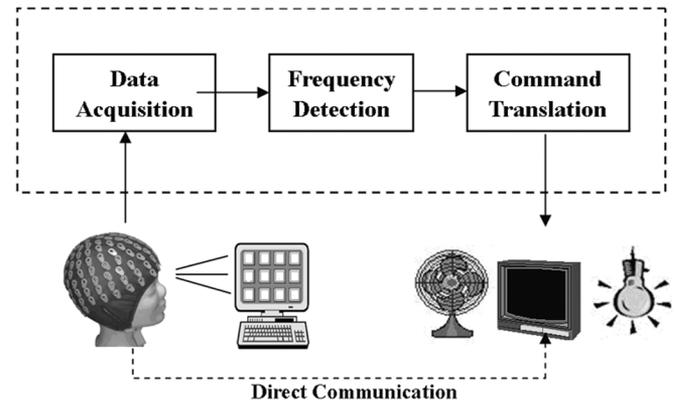


Fig. 1. Block diagram of system.

II. METHODS

A. Physiological Background

VEPs are derived from the brain's response to visual stimulation. They reflect the visual information processing mechanism in the brain. SSVEP is a response to a visual stimulus modulated at a frequency higher than 6 Hz. SSVEP can be recorded from the scalp over the visual cortex, with maximum amplitude at the occipital region. Photoc driving response, which is characterized by an increase in amplitude at the stimulus frequency, results in significant fundamental and second harmonics. Therefore, it is possible to detect the stimulus frequency based on measurement of SSVEP.

Large areas of the visual cortex are allocated to process the center of our field of vision, so the acuity is greatest when the stimulus is located in the center of the visual field [4], [16]. This effect is called central magnification, i.e., the amplitude of SSVEP increases enormously as the stimulus is moved closer to the central visual field. Different SSVEP can be produced by directly looking at one of a number of frequency-coded stimuli. This is the basic principle of an SSVEP-based BCI.

B. System Configuration

Fig. 1 is the block diagram of the system, which is similar to other BCI designs [1]. It includes a visual stimulator, EEG acquisition equipment, signal processing algorithms, and device control methods. A virtual keypad consisting of several buttons on a CRT monitor was designed as the stimulator [6]. The output device was controlled by alteration of the user's gaze direction.

A number of buttons flash at different repetition rates, composing a frequency-coded flashing matrix. The subject was asked to gaze at the target attentively, ignoring the other buttons in the operation. At this time, the target will be located in the center of the visual field, only causing an increased response associated with the target. Then, signal processing technology can be used to determine the stimulus frequency. The button, which matches the detected frequency, is the target the user wants to select.

The task of signal processing is to detect the existence of the SSVEP and determine its frequency. The target will induce a peak in the amplitude spectrum at the stimulus frequency, which is larger than the mean amplitudes of the lower and higher frequency bands. The first step is to search for the peak value and determine the corresponding frequency. To avoid false positives caused by spontaneous EEG signals, the next step is to make sure that the normalized signal amplitude at the determined frequency is above a predefined threshold. This step is effective

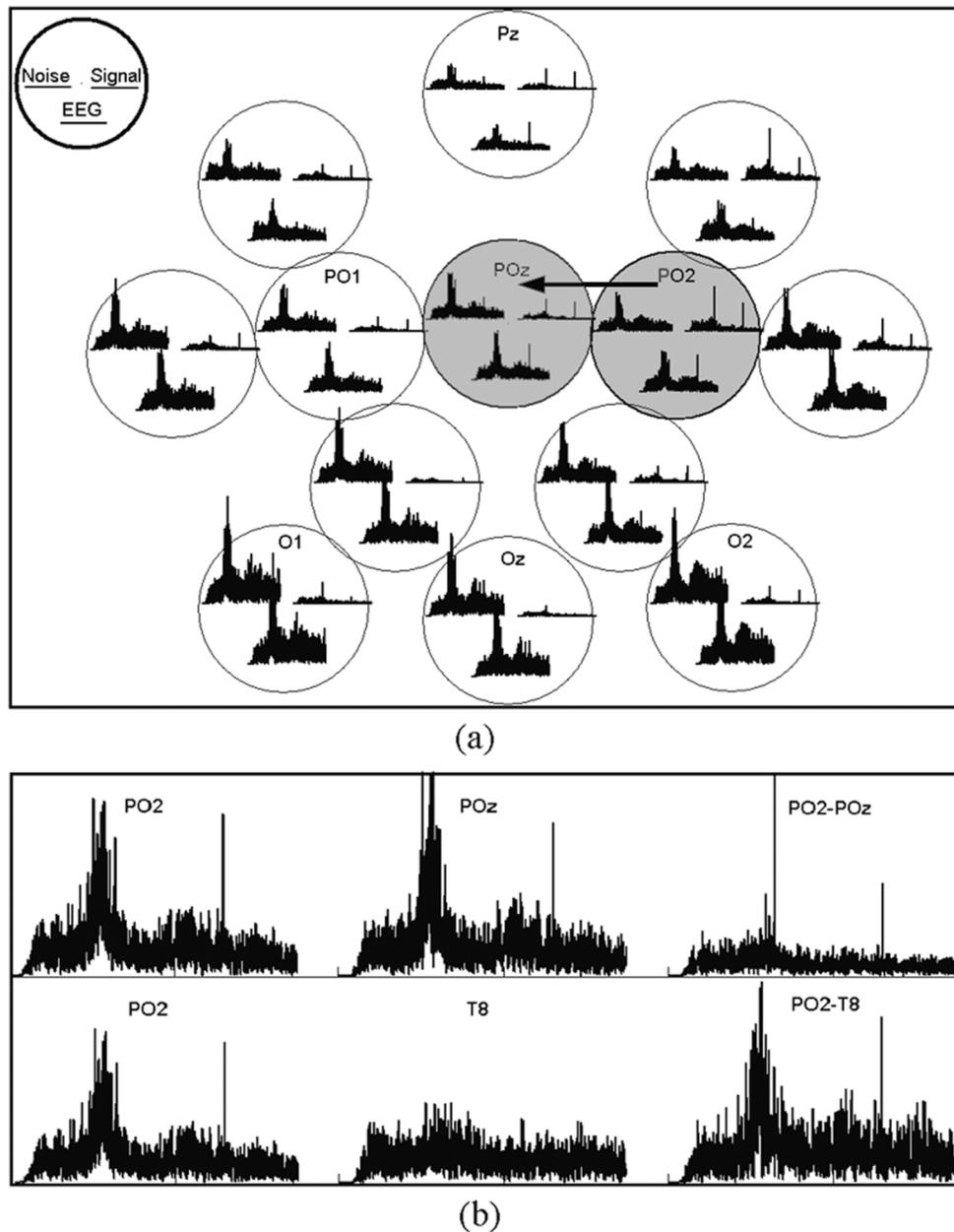


Fig. 2. (a) Spatial distributions of amplitude spectra on scalp over visual cortex. (b) Amplitude spectra on two pairs of channels, i.e., PO2-POz and PO2-T8.

for discriminating SSVEP and spontaneous EEG signals. A predefined command is executed if these two criteria can be satisfied.

C. System Optimization

To overcome the problem of individual diversity, a preparation experiment was implemented before a real practical application. The purpose of the experiment is to find three optimal parameters for each specific subject. The parameters will be used in his/her further applications.

The setup of the preparation experiment was as follows. Multichannel EEG signals were recorded with a BioSemi ActiveTwo system using 32 channels. Thirteen channels were located between Pz and Oz to record signals over the visual cortex with a high spatial resolution [Fig. 2(a)]. A blinking light-emitting diode (LED) modulated by square wave was used as the stimulator. The integer stimulus frequency covered the bandwidth from 9 to 17 Hz. Sixty-second-long

data were acquired in each test with different stimulation frequencies. Signals were sampled at 256 Hz and preprocessed by a 50-Hz notch filter and a 4–35-Hz band-pass filter. Three important parameters (including *channel location*, *stimulus frequency*, and *speed of selection*) were obtained from offline analyses of the recorded data.

1) *Channel Location*: To detect the frequency of SSVEP accurately and conveniently, a proper bipolar lead should be selected for real-life application. In practice, the channel with the most significant amplitude of SSVEP can be considered as the signal channel that commonly locates the visual cortex. The precise position can be determined with the study of EEG power mapping at the stimulus frequency. The difficulty is to determine the reference channel, which should enhance the signal-to-noise ratio (SNR). Here, two factors for reference channel selection are under consideration: the amplitude of SSVEP and the distance from the signal channel. To retain the signal, the reference channel must have lower SSVEP amplitude. To reduce the noise,

it should have similar background activities with the signal channel. Therefore, some channels close to the signal channel could be good candidates.

The channel selection of one representative subject is displayed in Fig. 2, which significantly increases user performance when compared with the conventional ear reference channel. Independent component analysis (ICA) is used for decomposition of signal and noise from single channel EEG [9], [10]. The detailed procedures are described as follows.

- 1) Independent component analysis. Thirteen-channel EEGs \mathbf{X} (with embedded SSVEP at 13 Hz) between Pz and Oz are selected as the input. Thirteen independent components (ICs) are calculated as sources \mathbf{S} through ICA, i.e., $\mathbf{S} = \mathbf{W} \cdot \mathbf{X}$, where \mathbf{W} is the demixing matrix.
- 2) Decomposition and reconstruction of signal and noise. The ICs with highest normalized amplitude [defined in Part C(2)] at the stimulus frequency are supposed to be signal activities and the remaining are considered as noise activities. They are denoted as $\mathbf{S}^{\text{Signal}}$ and $\mathbf{S}^{\text{Noise}}$. Then, \mathbf{S} can be expressed as

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}^{\text{Signal}} \\ \mathbf{S}^{\text{Noise}} \end{bmatrix}.$$

With the equation $\mathbf{X} = \mathbf{W}^{-1} \cdot \mathbf{S}$, \mathbf{X} can be divided into two components: $\mathbf{X}^{\text{Signal}} = \mathbf{W}^{-1} \cdot \mathbf{S}^{\text{Signal}}$ and $\mathbf{X}^{\text{Noise}} = \mathbf{W}^{-1} \cdot \mathbf{S}^{\text{Noise}}$. $\mathbf{X}^{\text{Signal}}$ and $\mathbf{X}^{\text{Noise}}$ are the reconstructions of signal and noise activities over the scalp. The number of signal ICs is predetermined empirically (e.g., 4).

- 3) Amplitude spectra analyses of \mathbf{X} , $\mathbf{X}^{\text{Signal}}$ and $\mathbf{X}^{\text{Noise}}$. Fig. 2(a) displays the results of the thirteen channels over the scalp. For each group, the amplitude spectrum of \mathbf{X} is in the bottom, $\mathbf{X}^{\text{Signal}}$ on the top right corner, and $\mathbf{X}^{\text{Noise}}$ on the top left corner.

The signal correlation and noise correlation between different channels are the basis of channel selection. Besides calculating the correlation coefficients between different channels, they can also be obtained directly by observing the spatial distributions of amplitude spectra over the scalp [10]. As shown in Fig. 2(a), the SSVEP of this subject is contaminated by strong spontaneous EEG signals. From the original EEG data, we are unable to choose the signal channel. The key point is to reduce the background noise. Through signal decomposition by ICA, the spatial distribution of signal activities shows that PO2 has the most significant SSVEP. Therefore, the channels close to PO2 are preferable to be selected as a reference channel. As shown in Fig. 2(b), PO2-POz is a good choice to weaken the background activities; in contrast, the SSVEP is drowned out by the spontaneous EEG signal with the conventional channel, e.g., PO2-T8.

2) *Stimulus Frequency*: The second parameter is *the stimulus frequency*. In order to implement a practical system, two problems related to stimulus frequency have to be solved. The first one concerns false positives. If the stimulus frequency band overlaps with alpha rhythms (8–13 Hz), the spontaneous EEG may likely satisfy the criteria of peak detection, without the user performing any intentional action. To implement an asynchronous BCI that allows the user to operate the system at any moment, avoidance of false positive is absolutely necessary. The second problem that needs to be solved concerns the efficiency of frequency detection. For most subjects, background components in the SSVEP are depressed, while the signal amplitude at the stimulus frequency increases enormously. However, for some subjects (see Fig. 3), a majority of signal energy lies within the alpha region. Fig. 3 displays the normalized amplitude spectra corresponding to three different stimulus frequencies (i.e., 13, 15, and 17 Hz) for a subject. Normalized amplitude spectrum is calculated by

$$P = \frac{|\text{FFT}(x)|}{\sum |\text{FFT}(x)|}$$

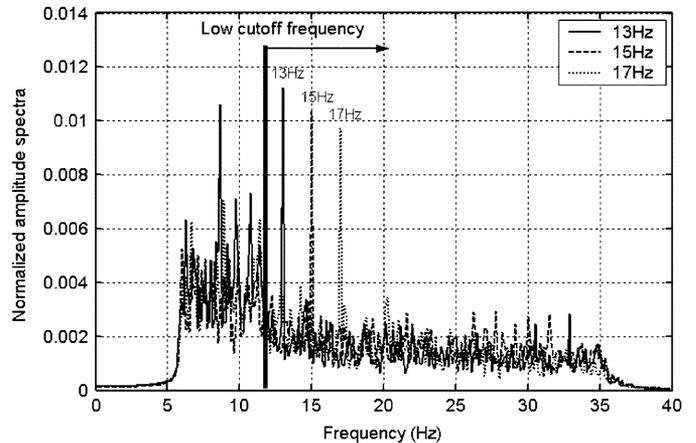


Fig. 3. EEG amplitude spectra corresponding to three different stimulus frequencies.

where x is the preprocessed EEG data and $\text{FFT}(x)$ is the fast Fourier transform of x . $\sum |\text{FFT}(x)|$ denotes summing over the total frequency points of the spectrum, thereby sum of the amplitude spectrum is normalized to one. The normalized amplitude at the stimulus frequency reflects the SNR of the SSVEP. Although the peak at stimulus frequency can be clearly identified, for this subject, the SNR of the peak value cannot reach the predefined threshold. This shows that the lower stimulus frequency is improper for this user. To exclude the frequency components in the alpha region, a high-pass filter is adopted in the system. For this subject, 12 Hz is a good choice of low cutoff frequency for the filter, which can effectively avoid the interference of alpha rhythms.

3) *Speed of Selection*: The third parameter is *the speed of selection (seconds per selection)*. Self-regulation of epoch length is an effective approach to increase the information transfer rate of BCIs [8], [11]. Due to physiological individual variation, the SNRs of the SSVEP have great inter-user differences. Suppose the background noise is a random signal and the SSVEP a deterministic one; then, the SNRs can be enhanced through increasing the length of the data for FFT. In order to reach the signal-to-noise threshold of operation, subject-specific speed, i.e., the length of EEG data per selection, needs to be considered. Fig. 4 displays normalized amplitude spectra corresponding to different data lengths (1, 2, 4, and 8 s) for two subjects. The data x was padded with zeros if it is shorter than 8 s (2048 points). With the same paces, subject (b) has a much better SNR than subject (a). If the threshold is set as 0.02, the data length for subject (a) has to be above 4 s, while that of subject (b) is less than 2 s. The SSVEP of most subjects has already reached a good SNR when the data length is 8 s in our experiments. In order to obtain the best system performance, an adaptive approach of data length regulation has been provided. A buffer with a length of 2048 points is prepared for FFT. During the operation, the buffer is filled with continuous EEG data in chronological order. If the buffer is filled up, the data will be shifted to the left, discarding the beginning points. Real-time calculation of frequency detection is executed every several hundred milliseconds using the total 2048 points data in the buffer. A decision will be made if the same frequency is detected in several consecutive calculations. Then, the buffer will be cleared out and will start to receive new data accompanied by the next operation. Empirically, the self-regulating data length for most users is less than 8 s with the aid of channel location and stimulus frequency optimization. This approach reduces the inter-user variation and makes the system more flexible and reliable.

D. Experiment Conditions and Performance Measure

After the preparation experiment, subject-specific parameters can be determined. Only one bipolar lead is needed for the operation task. The

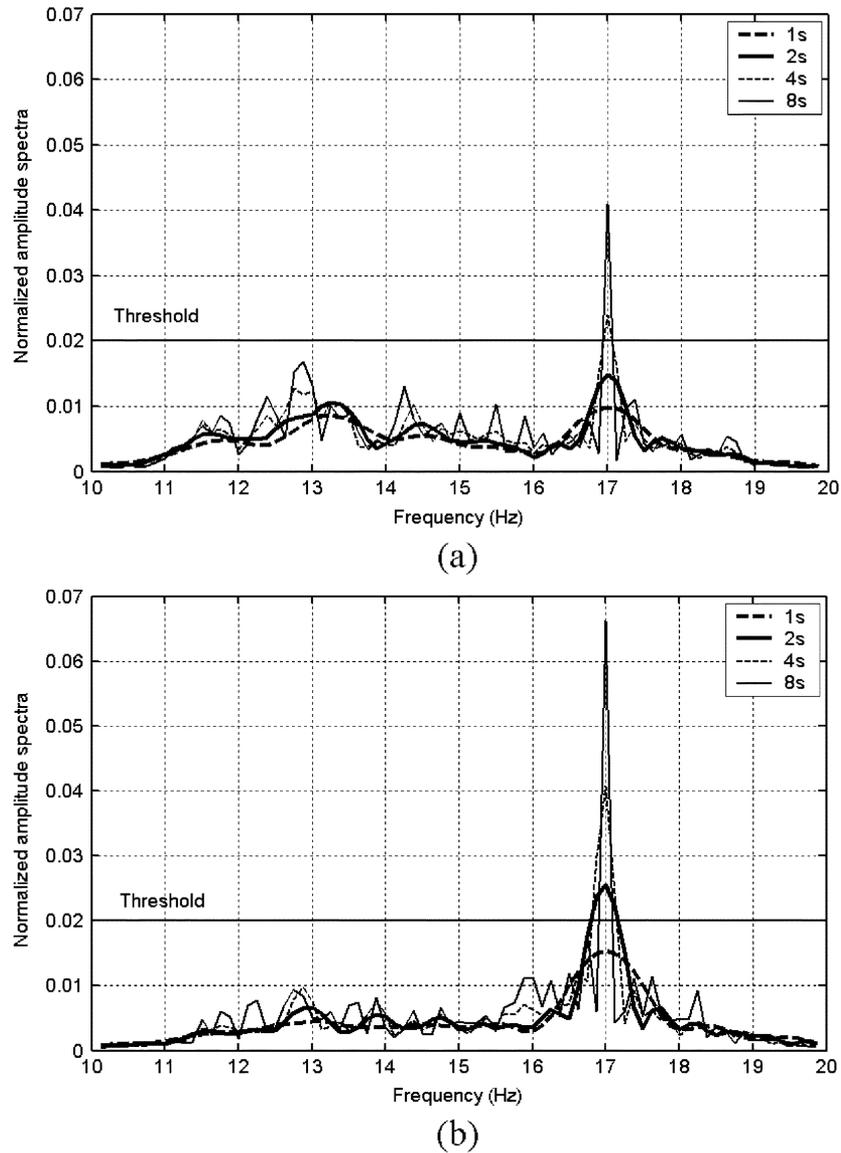


Fig. 4. Normalized amplitude spectra corresponding to different data lengths for two subjects (a) and (b).

TABLE I
ONLINE OPERATION RESULTS OF THREE SUBJECTS IN LABORATORY (←: BACKSPACE, ✓: ENTER)

Subject	Stimulus Frequency (Hz)	Input Results	Seconds per selection (max/min/average) (s)	ITR (bits/min)
QK	13-16	5917667775 ✓ ✓	3.77/3.21/3.40	63.10
GL	11-14	24527951 ← 411	5.96/4.70/4.87	44.17
MJ	11-14	363 ← 4119 ← (✓ ✓ ←) 885	10.46/4.98/5.68	31.35

stimulator is a virtual number keypad consisting of 13 buttons [6]. The frequency resolution is set to 0.25 Hz, so the required stimulus frequency band is $13 \times 0.25 = 3.25$ Hz, which can be easily selected away from the alpha band. Frequency detection is calculated by a 1024-point ($1/0.25 = 4$ s) FFT. The buffer used for adaptive pace regulation is divided into two 4-s sections. The amplitude spectrum is obtained by summing the two spectra derived from two sections, respectively.

System performance is evaluated by ITR. The measurement of ITR is obtained through the task of completing 12 commands consecutively

with the virtual keypad. The bit rate of each selection can be calculated with the method presented in [1]. The subjects were asked to input randomly arranged commands using a rectangular frame surrounding the instructed target as a cue. When a correct decision was made, the cue disappeared and the next target would appear. If a wrong decision was made, a new cue would appear around the wrong target, instructing the subject to clear it by gazing at it.

Implementation and evaluation of the demonstration system should be carried out with a large number of subjects. Sixteen volunteers (six

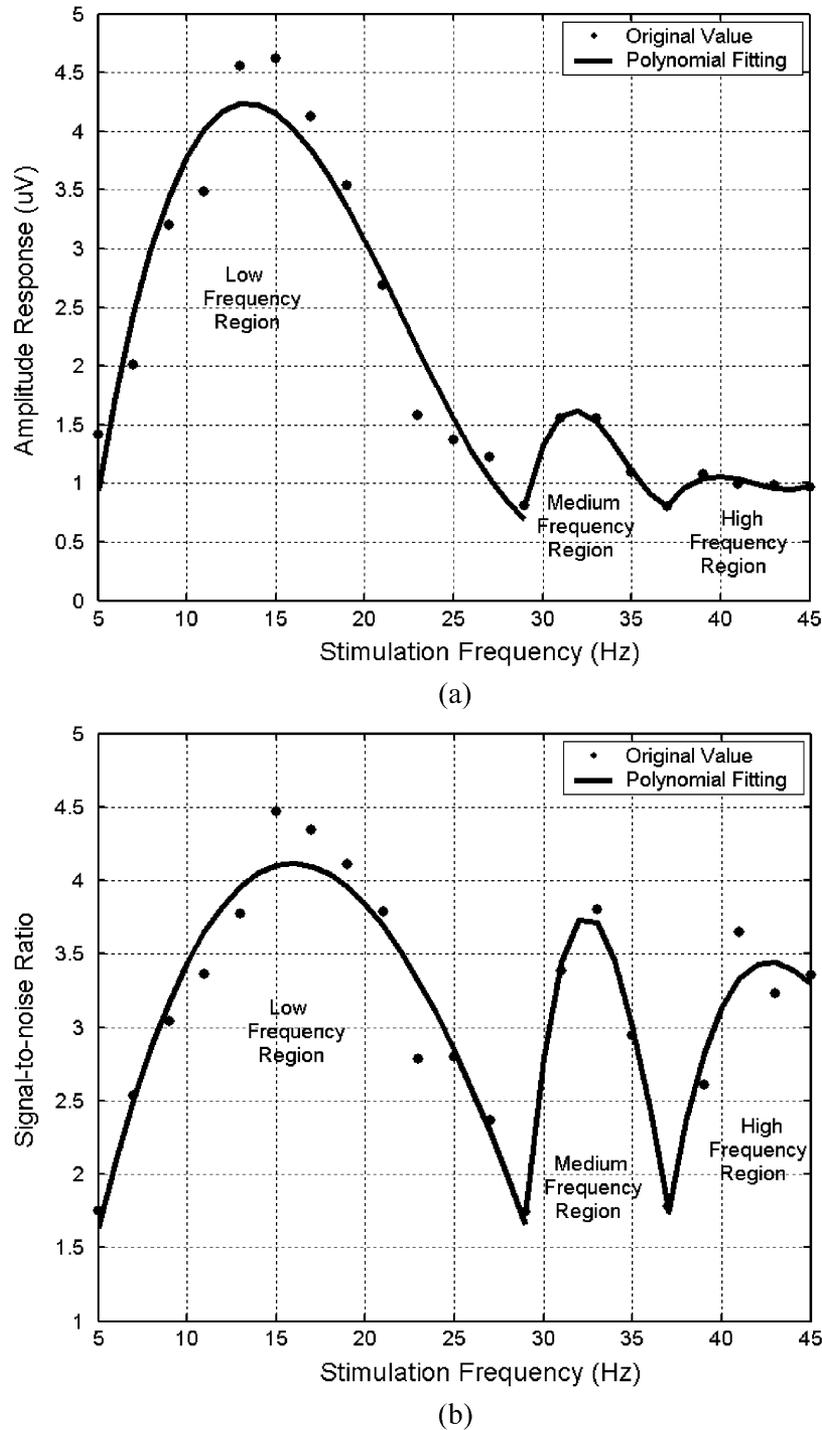


Fig. 5. (a) Three SSVEP subsystems of subject. (b) SNR curves corresponding to subsystems.

female and ten male, 22–42 years old, randomly selected from the students in the department) with normal vision participated in the experiments in the laboratory. All participants were seated in a comfortable armchair in an unshielded room, which is dim and quiet without any distractions. The subjects varied in their prior experience of this BCI. Five subjects (two of them from our previous study [6]) had experienced many online experiments of digit number input, and the remaining subjects had some knowledge of this BCI but without any experience. Furthermore, to investigate system applicability under different conditions, the demonstration system had also been tested in the

Rehabilitation Center of China. A BCI-based environmental controller [12] was used to help the motion-disabled to control home appliances. A trainable infrared remote controller was used to control a TV in a doctor's office. Eleven volunteers (two female and nine male, 22–53 years old) with spinal cord injury (range C4–C7 injured) served as subjects.

III. RESULTS AND DISCUSSION

After subject-specific parameters were derived from the preparation experiment, all 16 subjects in the laboratory fulfilled the 12-target tasks

successfully. The average ITR was 43 bits/min (range 29–63 bits/min). Table I lists the detailed results of three representative subjects corresponding to different performance levels. Two of them completed the task without errors. Subject MJ made a wrong input “↙” and then cleared it while the task was to input a “←”. The mistake is denoted as “(↙/←)” in the table. In the Rehabilitation Center, ten subjects could operate the system without any training and the average ITR was 21 bits/min (range 8–36 bits/min). The decrease in ITR may be caused by the following factors: environmental brightness and noise, lack of training practice, large distance between the stimulator and the subject “locked” in the wheelchair, physical discomfort caused by the duration of the pose, and the temporary inability to adapt to the flickering stimulation. The results indicate that the SSVEP-based BCI can be applicable to >90% of people with a high ITR in living environments.

The ITR depends on the number of possible selections, accuracy, and speed of operation. For SSVEP-based BCI, lower speed can enhance the SNRs of SSVEPs, leading to a higher accuracy. The tradeoff between these two factors should be considered carefully to obtain the highest intra-user ITR. A possible way to enhance the SNR is optimization of the stimulus parameters such as flicker modulation depth, modulation wave, and alternation color. Our investigation shows that the white/black alternation modulated by a square wave with a modulation depth of about 40% is good for eliciting the SSVEP.

An increase in the number of possible selections is another effective approach to improve ITR. Here are two possible methods. The first is to decrease the frequency resolution, and the second is to extend the stimulus frequency bandwidth. The former will increase the difficulty of frequency detection thereby decreasing the accuracy, and the latter has more potential to improve the system performance. For this reason, we have tried to employ high-frequency SSVEP (>20 Hz) in the system. This is also because the participants’ subjective feelings show that the high-frequency stimulation has the advantages of reducing the visual fatigue caused by flicker and making them more comfortable. Since flicker is believed to be a contributing factor to eyestrain, to some extent, this phenomenon can be explained by flicker fusion effect [13] under high-frequency stimulations.

According to research done in the area of human brain electrophysiology, the amplitude of the SSVEP varies in a complex manner with the frequency of stimulation [12], [14]–[16]. The amplitude response has several peaks within three regions, often referred to as subsystems: low-frequency region, medium-frequency region, and high-frequency region [16]. The amplitudes in the subsystems depend on many factors including electrode position, luminance, and flicker modulation depth. The amplitude versus frequency response curves for the three SSVEP subsystems of a subject are shown in Fig. 5(a). The stimulus frequency varied from 5 to 45 Hz (2 Hz each step). Filled dots indicate the original results and solid lines denote the polynomial fitting results. The three subsystems are centered at 15-Hz low frequency, 31-Hz medium frequency, and 41-Hz high frequency, respectively. The lower frequency region has the larger amplitude response.

The lower amplitude response may decrease the SNR if the background noise is unchanged. Fortunately, the noise, i.e., spontaneous EEG, also decreases in higher frequency bands. In fact, the SNRs of the three subsystems have almost identical levels [as shown in Fig. 5(b)]. Here, SNR is defined as the ratio of $y(f)$ to the mean value of the n adjacent points

$$\text{SNR} = \frac{n \times y(f)}{\sum_{k=1}^{n/2} [y(f + 0.25 \times k) + y(f - 0.25 \times k)]}$$

where y is the amplitude spectrum calculated by a 1024-point FFT and f is the stimulus frequency. Offline frequency detection has been done in medium- and high-frequency regions. The detection accuracy

is above 95% when the stimulation frequencies are in the medium-frequency region (31–35 Hz) and the high-frequency region (39–45 Hz). For this subject, the additional bandwidth in these two regions is about 10 Hz and the total usable bandwidth in the three subsystems can allow more than 60 targets.

Current BCIs are mostly at the stage of laboratory demonstrations due to problems associated with high user variation, cumbersome recording preparation, and low ITR. By solving these problems through subject-specific system optimization, the SSVEP-BCI has become a practical system which can be easily implemented for most people. This is truly a good preparation for the future development of a commercial product.

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.*, vol. 113, no. 6, pp. 767–791, 2002.
- [2] T. M. Vaughan, W. J. Heetderks, L. J. Trejo, W. Z. Rymer, M. Weinrich, M. M. Moore, A. Kubler, B. H. Dobkin, N. Birbaumer, E. Donchin, E. W. Wolpaw, and J. R. Wolpaw, “Brain-computer interface technology: A review of the second international meeting,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 94–109, Jun. 2003.
- [3] B. Blankertz, K. R. Müller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schloegl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schroeder, and N. Birbaumer, “The BCI competition 2003: Progress and perspectives in detection and discrimination of EEG single trials,” *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1044–1051, Jun. 2004.
- [4] E. E. Sutter, “The brain response interface: Communication through visually-induced electrical brain response,” *J. Microcomput. Applicat.*, vol. 15, pp. 31–45, 1992.
- [5] M. Middendorf, G. McMillan, G. Calhoun, and K. S. Jones, “Brain-computer interfaces based on the steady-state visual-evoked response,” *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 211–214, Jun. 2000.
- [6] 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.*, vol. 49, no. 10, pp. 1181–1186, Oct. 2002.
- [7] G. R. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller, “Steady-state visual evoked potential (SSVEP)—Based communication: Impact of harmonic frequency components,” *J. Neural Eng.*, vol. 2, pp. 123–130, 2005.
- [8] 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.*, vol. 52, no. 9, pp. 1588–1596, Sep. 2005.
- [9] B. Hong, F. S. Yang, Y. F. Pan, Q. Y. Tang, K. Chen, and Y. M. Tie, “Single-trial estimation and analysis of PVEP based on independent component analysis,” *Tsinghua Sci. Technol.*, vol. 6, no. 5, pp. 503–508, 2001.
- [10] Y. Wang, Z. Zhang, X. Gao, and S. Gao, “Lead selection for SSVEP-based brain-computer interface,” in *Proc. 26th Int. IEEE EMBS Conf.*, 2004, pp. 4507–4510.
- [11] J. D. Bayliss, “A flexible brain-computer interface,” Ph.D. dissertation, Univ. Rochester, Rochester, NY, Aug. 2001.
- [12] X. Gao, D. Xu, M. Cheng, and S. Gao, “A BCI-based environmental controller for the motion-disabled,” *IEEE Trans. Rehab. Eng.*, vol. 11, no. 2, pp. 137–140, Jun. 2003.
- [13] T. N. Cornsweet, *Visual Perception*. New York: Academic, 1970.
- [14] C. S. Herrmann, “Human EEG responses to 1–100 Hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena,” *Exp. Brain Res.*, vol. 137, pp. 346–353, 2001.
- [15] M. M. Müller, T. W. Picton, P. Valdes-Sosa, J. Riera, W. A. Teder-Sälejärvi, and S. A. Hillyard, “Effects of spatial selective attention on the steady-state visual evoked potential in the 20–28 Hz range,” *Cog. Brain Res.*, vol. 6, pp. 249–261, 1998.
- [16] D. Regan, *Human Brain Electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine*. North Holland, The Netherlands: Elsevier, 1987.