Annals of Biomedical Engineering, Vol. 34, No. 10, October 2006 (© 2006) pp. 1641–1654
DOI: 10.1007/s10439-006-9175-8

The Brain Computer Interface Using Flash Visual Evoked Potential and Independent Component Analysis

PO-LEI LEE,[1,2,3] JEN-CHUEN HSIEH,[2,3,4,5,6] CHI-HSUN WU,[1,2] KUO-KAI SHYU,[1] SHYAN-SHIOU CHEN,[2] TZU-CHEN YEH,[2,5] and YU-TE WU [2,3,7]

[1]Department of Electrical Engineering, National Central University, Taoyuan, Taiwan; [2]Integrated Brain Research Laboratory, Department of Medical Research and Education, Taipei Veterans General Hospital, Taipei, Taiwan; [3]Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan; [4]Center for Neuroscience, National Yang-Ming University, Taipei, Taiwan; [5]Faculty of Medicine, School of Medicine, Institute of Radiological Science, National Yang-Ming University, Taipei, Taiwan; [6]Institute of Neuroscience, School of Life Science, National Yang-Ming University, Taipei, Taiwan; and [7]Institute of Radiological Science, National Yang-Ming University, No. 155, Li-Nong Street, Section 2, Pei-Tou, Taipei 112, Taiwan, ROC

(Received 4 September 2005; accepted 9 August 2006; published online: 20 September 2006)

Abstract—In this study flashing stimuli, such as digits or letters, are displayed on a LCD screen to induce flash visual evoked potentials (FVEPs). The aim of the proposed interface is to generate desired strings while one stares at target stimulus one after one. To effectively extract visually-induced neural activities with superior signal-to-noise ratio, independent component analysis (ICA) is employed to decompose the measured EEG and task-related components are subsequently selected for data reconstruction. In addition, all the flickering sequences are designed to be mutually independent in order to remove the contamination induced by surrounding non-target stimuli from the ICA-recovered signals. Since FVEPs are time-locked and phase-locked to flash onsets of gazed stimulus, segmented epochs from ICA-recovered signals based on flash onsets of gazed stimulus will be sharpen after averaging whereas those based on flash onsets of non-gazed stimuli will be suppressed after averaging. The stimulus inducing the largest averaged FVEPs is identified as the gazed target and corresponding digit or letter is sent out. Five subjects were asked to gaze at each stimulus. The mean detection accuracy resulted from averaging 15 epochs was 99.7%. Another experiment was to generate a specified string '0287513694E'. The mean accuracy and information transfer rates were 83% and 23.06 bits/min, respectively.

Keywords—Flash visual evoked potential (FVEP), Electroencephalography (EEG), Independent component analysis (ICA), Brain computer interface (BCI).

INTRODUCTION

Patients suffering from severe motor disabilities, such as amyotrophic lateral scleroses (ALS), severe cerebral palsy, muscular dystrophies, etc., are incapable of communicating with external environment (locked-in syndrome). Over the past decades, several groups have attempted to develop novel communication techniques independent of peripheral nerves and muscles for neuromuscular-impairment patients.[29] One of the promising methods is the use of electroencephalography (EEG) induced from performing designed mental tasks to activate external devices or express user's intension.[29] Such a technique is referred to as 'brain computer interface (BCI)'.[29]

A reliable BCI system relies on discernible neuroelectric signals. Visual stimulation using flash light is a popular and easy means to induce event-related visual evoked potential (VEP). Flash VEP (FVEP) has been a popularly clinical index to monitor anesthesia level during surgical operation,[22,27] to diagnose prechiasmal and retrochiasmal lesions,[28] to indicate intracranial pressure induced by head injury,[15,30] and to alarm brain death.[23,26] According to the neural connections and interactions of the route from retina to primary visual cortex, visual stimuli at central visual field can generate the so-called 'cortical magnification' which makes the central FVEPs more prominent than any FVEPs evoked from peripheral visual fields.[17,19,20] In the proposed system, twelve flashing stimuli, i.e., digits '0–9', one Backspace 'B' and one Enter 'E', were created and displayed on a PC-controlled LCD monitor. Subjects were asked to stare at the target stimulus one by one such that the detected central FVEPs can be used to identify the gazed stimulus and output the corresponding digit or letter to form a string.

Since the measured EEG can be mixed with artifacts (e.g., ocular artifact) or task-unrelated components (e.g., occipital alpha rhythm),[8,11] the present study

Address correspondence to Yu-Te Wu PhD, Institute of Radiological Science, National Yang-Ming University, No. 155, Li-Nong Street, Section 2, Pei-Tou, Taipei 112, Taiwan, ROC. Electronic mail: ytwu@ym.edu.tw

employs an independent component analysis (ICA) to separate the FVEP component from EEG recordings. ICA is a data-driven method for multivariate data analysis, which has been used to recover temporally independent neuronal activities of EEG[5,10,11,14] and MEG measurements,[10,24] or to recover spatially independent maps in the fMRI[4,13] and perfusion MRI[9] studies. Specifically, brain signals across all EEG channels can be decomposed into mutually independent components (ICs) by means of solving an unmixing matrix in which each column represents a spatial map tailoring the weights of the corresponding temporal component at each EEG sensor. Task-related ICs are screened and identified by correlating their spatial maps with a pre-defined spatial template, which is created based on the spatial weight distribution of P2 peak in conventional FVEP obtained from each individual.[19] Selected task-related ICs are used to recover FVEPs which have significantly improved signal-to-noise ratio (SNR). It is noteworthy that the reconstructed FVEPs consist of both the central FVEPs induced by flashing target and peripheral FVEPs induced by surrounding non-target flickering sequences.

In order to remove the contamination of peripheral FVEPs from reconstructed FVEPs, special flickering sequences are designed for the multiple flashing stimuli. Each flickering sequence consists of alternative ON–OFF states in which each ON or OFF duration is random to make all the flickering sequences mutually independent. EEG recordings are segmented into 12 sets of FVEP epochs based on the flash onsets (OFF-to-ON flashes) of each stimulus. Since FVEPs are time-locked and phase-locked to visual stimulus,[21] the central FVEP epochs are synchronized to the flash onsets of gazed flickering stimulus and can be enhanced by simple averaging. On the contrary, the peripheral FVEP epochs are asynchronized to the flash onsets of gazed flickering stimulus and can be suppressed after averaging. By comparing the 12 sets of averaged FVEP, the stimulus producing the largest peak-to-valley FVEP amplitude was identified as the gazed target.

Various EEG-based BCI systems have been developed with elaborately designed paradigms to induce specific types of neuroelectric signals. Pfurtscheller *et al.* measured sensorimotor mu rhythms during subjects' imagery movements and achieved a high recognition rate of 90%.[18] Birbaumer *et al.* developed Thought Translation Device (TTD) to measure slow cortical potentials (SCPs) which has been successfully implemented in a binary selection task.[1] Donchin *et al.* adopted odd-ball paradigm to implement an alphabet typing system based on visually-induced P300.[3] Mason and Birch designed an asynchronous detector to detect motor-related potentials (MRPs) within 1–4 Hz for controlling binary switches.[12]

Two other VEP-based BCI systems have been proposed. One was based on the fast multifocal visual evoked potential (FMFVEP)[21] and the other on the steady-state visual evoked potential (SSVEP).[2] The flickering stimuli of FMFVEP-based system were generated by a pseudo-random binary sequence with a fixed time lag between any two adjacent channels. Each entire pseudo-random sequence was convoluted with a standard VEP response to create a so-called "expected response". By finding the maximum correlation between the measured EEG signals and the expected response of each flickering stimulus, the gazed stimulus was recognized. Instead of using a binary sequence with fixed flickering frequency, each stimulus in the SSVEP-based system was designed to have its own flickering frequency. The gazed target was identified by finding the stimulus which contributes maximum power of SSVEP at Fourier spectrum. However, these are limitations in these two methods. The FMFVEP-based method presumes identical response of VEP across all trials and uses it as template in correlation computation.[21] Such a stringent assumption is irreconcilable with the truth[8,24] and the resultant correlation may not be optimal in detecting the gazed target. In the SSVEP-based method, the flickering frequencies were confined to be lower than 14 Hz due to the frame rate of PC monitor, and flickering frequencies around alpha band should also be excluded to avoid the interference of spontaneous alpha rhythm. These two constrains may reduce the available flickering channels and communication bandwidth. Our method employs mutually independent flickering sequences generated by random ON–OFF durations and uses the timing of flash onsets to segment EEG data followed by simple averaging. Only peak-to-valley amplitudes, rather than correlation values or power spectrum, are computed and compared to easily determine the gazed target.

MATERIALS AND METHODS

Subjects and EEG Recording

Five volunteers (Four males and one female), ages from 25 to 32 years, were recruited to sit on a comfortable armchair in a dimly illuminated room. Each subject had corrected Snellen visual acuity of 6/6 or better, with no history of clinical visual disease. Among the five subjects, two of them (WCH and LSK, both were male) had 1-h experience in performing this visual stimulation task while the other three were naive subjects. All subjects were requested to complete a sequence of numbers and letter, which was '0287513694E'. Scalp neuroelectric activities were measured by a whole-head forty-channel EEG system (bandpass, 0.05–100 Hz; digitized at 250 Hz; Nu

Amplifier; Neuroscan Ltd., USA). Two out of the forty EEG channels were, respectively, used as bipolar horizontal and vertical electrooculograms (EOG), one was placed below and above the left eye and the other at the bilateral outer canthi. Only 38 EEG channels were used in the subsequent ICA processing. The signals recorded from two additional electrodes placed on left and right mastoids were averaged and used as the reference to all EEG channels. Inter-electrode impedances were kept below 5 kΩ during signal recordings. The acquired EEG signals were bandpass filtered, within 1–50 Hz, to remove 60 Hz electricity noise and low-frequency drifts which were then real-time transmitted (Client/Sever Operation, Scan 4.3; Neuroscan Ltd., USA) to a PC for gazed-FC detections.

Visual Stimuli

The visual stimuli of the proposed system consisted of 12 flickering channels, namely 10 digits ('0–9'), one Backspace ('B') and one Enter ('E'). Those flickering stimuli were arranged as a 3×4 tessellation array and presented on a 17-inch ViewSonic LCD monitor (model VG724; reaction time <3 ms; 60 frames/s) which was 40 cm apart from the viewer. Each flickering channel was designed in square-shape (subtended angle = 3) inlaid with a functional indicator, i.e., a digit or a letter, and a small cross and driven by a flickering sequence consisted of alternative ON–OFF states. The small crosses were used to draw subject's attention so that his/her gaze was fixated at the center of the crosses. Duration of each ON or OFF state consists of a fixed length of 116.7 ms and a variable length which is uniformly distributed with values between 0 ms and 233.3 ms. (see Fig. 1). Accordingly, the overall duration of each ON or OFF state is between 116.7 ms and 350 ms, with mean duration of 233.3 ms. It is worthy to note that the design of fixed duration is to prevent the overlap of two consecutive FVEPs since the major FVEPs usually occur within 0–100 ms.[19] The luminance of ON and OFF sates in each FC were 158.7 candelas (cd)/m^2 and 8.1 cd/m^2, respectively, measured by a luminance meter (LS-110; Konica Minolta Photo Imaging Inc., USA) resulting in Michelson contrast of 90.3%. Fig. 2 shows the schematic diagram of the proposed system.

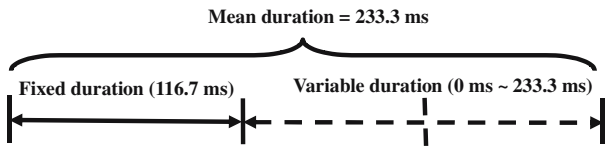

FIGURE 1. The random duration of ON or OFF state. Each ON or OFF duration in a flickering sequence is generated by a fixed duration of 116.7 ms elongated by a uniformly random duration with values between 0 ms and 233.3 ms. The overall duration of each ON or OFF state is between 116.7 ms and 350 ms, with mean duration of 233.3 ms.

Extraction of FVEP using ICA and Spatial Template Created from Flash Visual Evoked Potential

To extract the task-related components using ICA, EEG signals across m channels ($m = 38$) and n points ($n = 5000$), from −20 s to 0 s, were first arranged into a $m \times n$ matrix $\mathbf{B}$. The ith row was the measured signal from ith EEG channel, and the jth column was the observed samples at the jth time point across all EEG channels. In this study, FastICA was adopted because of its fast convergence.[6,7] The FastICA first removed means of the row vectors in the $\mathbf{B}$ matrix. A whitening process based on principal component analysis was then used to sphere the covariance matrix of the zero-mean matrix into an identity matrix. In FastICA algorithm, we can optionally select the first p ($p \leq m$) most significant eigenvectors for the whitening process. Since the electro-neural activities can be smaller than internal/external noise in this study, all eigenvectors were retained (i.e., $p = m$) so that all the electro-neural information were decomposed in the subsequent ICA processing. Finally, FastICA searched for a matrix to transform the whitened data into a set of components which were as mutually independent as possible. Assume that the matrix $\mathbf{B}$ is zero-mean, in combination with previous whitening process, the matrix $\mathbf{B}$ can be transformed into p independent components (ICs) in a matrix $\mathbf{S}$ via an un-mixing matrix $\mathbf{W}$, i.e.,

$$\mathbf{S}_{p\times n} = \mathbf{W}_{p\times m} \cdot \mathbf{B}_{m\times n} \tag{1}$$

where each row of matrix $\mathbf{S}$ represents samples of an independent component (IC). The pseudoinverse matrix $\mathbf{W}^{-1}$ of matrix $\mathbf{W}$ is a mixing matrix, referred to as U, which combines the p ICs to reconstruct signal matrix $\mathbf{B}$. The ith column in mixing matrix $\mathbf{U}$ represents weight distribution values of the corresponding ith IC across all EEG channels. The column vectors of mixing matrix $\mathbf{U}$ are hereinafter referred to as spatial maps.

Because the spatial distribution characterizes physiological meaning of specific neural activities,[10,16] the spatial maps of all ICs were correlated with a spatial template which was created beforehand for each subject based on the topographical distribution of a conventional flash visual evoked potential (FVEP). The conventional FVEP was induced by a flickering stimulus with 2 Hz flickering frequency with a square-shape (subtend angle = 3°) inlaid with a small cross. Subjects were asked to focus binocularly on the center of the visual stimulus for 1-min recording. One hundred epochs segmented from −0.1 s to 0.45 s relative to flash onsets were averaged to obtain the conventional

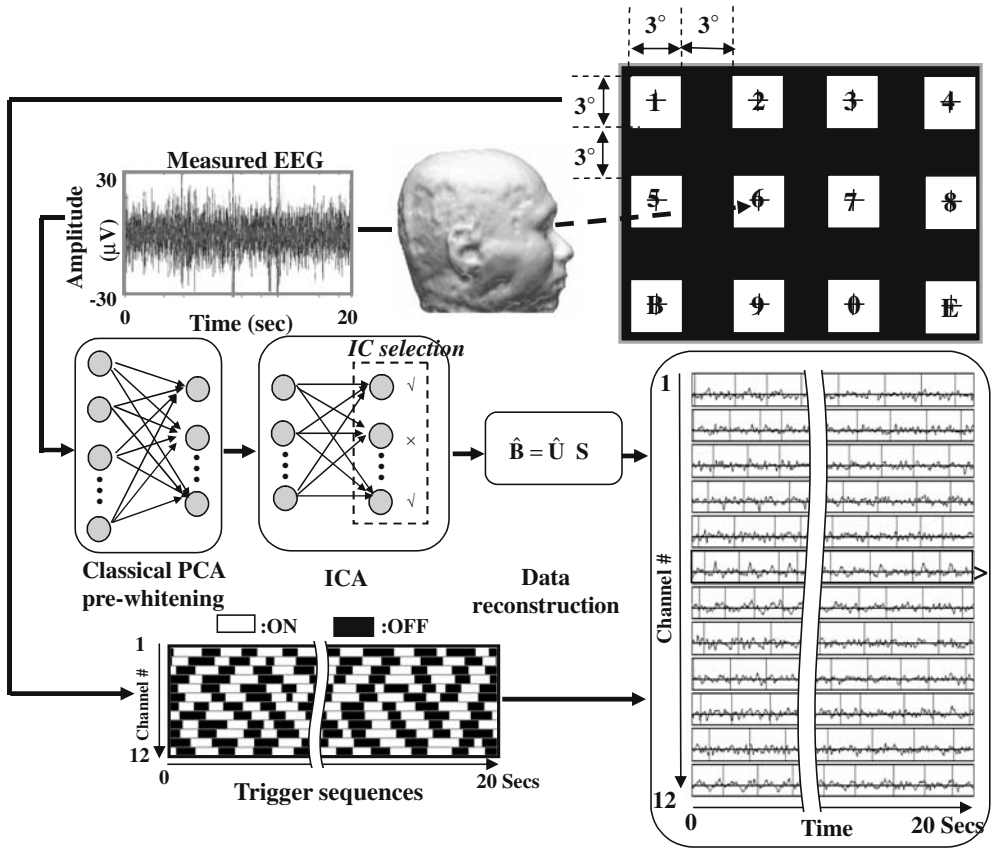

FIGURE 2. The schematic diagram of ICA-based FVEP-actuated BCI system.

FVEP (see Fig. 3). The P2 peak at each channel, which represents the most prominent amplitude in FVEPs, was used to constitute the spatial template.

ICs with correlation coefficients higher than 95% ranking over all ICs were selected (i.e., $5\% \cdot 38 = 1.9 \approx 2$ ICs) as task-related components and used in subsequent data reconstruction. Unselected columns, i.e., spatial maps of task-unrelated ICs, of mixing matrix **U** were zeroed to produce a matrix $\hat{U}$ such that task-related signals were reconstructed by multiplying $\hat{U}$ and **S**, which can be represented as

$$\hat{B} = \hat{U} \cdot S \tag{2}$$

where $\hat{B}$ is the reconstructed EEG data, $\hat{U}$ contains the spatial maps of task-related ICs, and **S** stores the ICs. The row $\hat{B}$ corresponds to Oz channel is used to generate the FVEP for detecting peak-to-valley amplitude in the followings.

Peak-to-valley Amplitude ($Amp_{p\text{-}v}$) in FVEP

Human FVEP has four major peaks, two negative and positive peaks, i.e., N1, P1, N2, and P2, within 200 ms after flash onset[19] (see Fig. 3). In normal subjects, the N2 and P2 peaks usually present the most prominent responses.[17,19,20] The amplitude difference between the N2 and P2 peaks, $Amp_{p\text{-}v}$, was chosen as the feature for detecting gazed stimulus. Because the presence of the latencies of N2 and P2 peaks could vary from trial-to-trial, the N2 and P2 peaks were searched in a time window by extending ± 15 ms around the timing of N2 (t_{n2}) and P2 (t_{p2}) peaks in individual's conventional FVEP. The values of $Amp_{p\text{-}v}$ obtained from all stimuli were compared to determine the gazed target.

Determination of Gazed Target by Detecting the Largest $Amp_{p\text{-}v}$ Among All Flickering Channels

The ICA-reconstructed signals were contaminated by peripheral FVEPs induced from non-target visual stimuli. ON and OFF durations were designed to be random which made all the flickering pattern sequences mutually independent so that interferences from non-target stimuli can be suppressed via simple averaging. Based on the timings of flash onset in the ith flickering sequence, $i = 1...12$, ICA-reconstructed signals were segmented into epochs, from -0.1 s to 0.45 s,

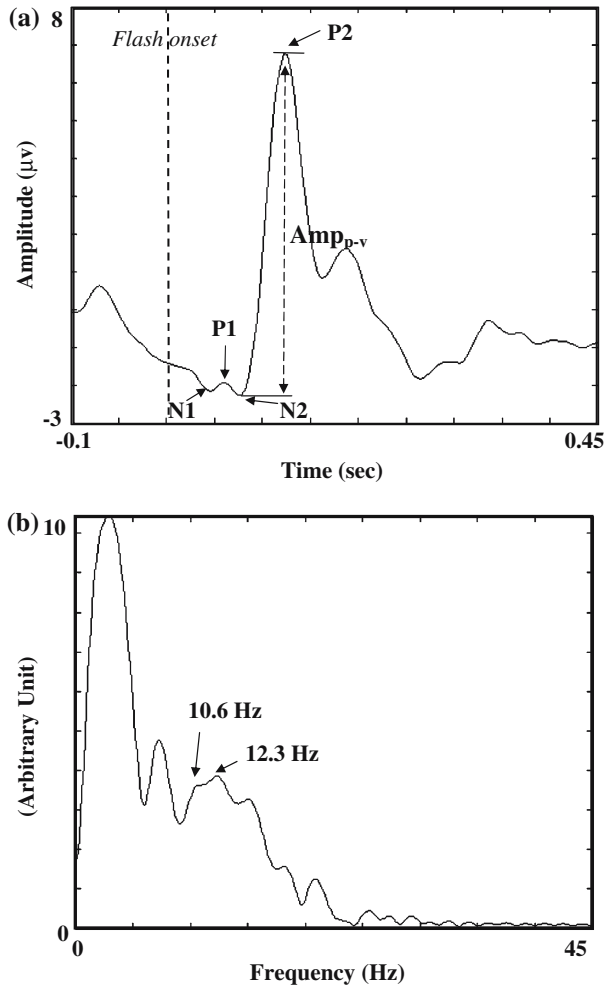

FIGURE 3. Flash visual evoked potential. (a) Human FVEP has two major negative and two major positive peaks within 200 ms after flash onset, namely, N1, P1, N2, and P2. In normal subjects, the N2 and P2 peaks usually exhibit the most prominent responses and their differences were chosen as the feature for identifying the gazed stimulus. (b) The Fourier spectrum of the conventional FVEP.

and stored in an assigned register, namely the *i*th epoch register, in the computer memory. The procedure of detecting gazed target was as follows. First, every N epochs ($N = 5$ in our implementation) in each *i*th epoch register, $i = 1...12$, were averaged. Second, the averaged epoch was lowpass filtered (< 30 Hz) to produce a noise-suppressed $FVEP_i$. Third, 12 Amp_{p-v} were computed and the stimulus inducing the largest Amp_{p-v} was recognized. Finally, the screen letter or digit defined by the identified stimulus was sent out with a concurrent auditory bio-feedback presented to the subject. In our current design, a gazed stimulus was detected in every second and was confirmed as the target after three consecutively successful detections, i.e., a letter or digit was sent out in every 3 s. The overall processing flowchart is summarized in Fig. 4.

RESULTS

ICA has been utilized in this study to extract FVEP activities from a variety of artifacts in EEG signals under the assumption that brain activities are independent of artifacts.[8,10,11,14] Fig. 5 is a typical example (subject WCH) to demonstrate the advantage of ICA. In Fig. 5(a), the first column displays the spatial template and the measure EEG at Oz channel. The second column displays six IC components which are the occipital alpha rhythm (IC3), FVEP activities (IC8 and IC16), right sensorimotor mu rhythm (IC21), ocular artifact (IC28), and left sensorimotor mu rhythm (IC36), respectively, with resultant correlation coefficients 0.17, 0.89, 0.83, −0.35, −0.34, and −0.36 after fitting with the spatial template. The third and fourth columns are the corresponding temporal waveforms and their Fourier spectra, respectively. The IC8 and IC16, which have correlation coefficients within the highest 5% ranking were selected to reconstruct the whole-head FVEP activities. The spatial maps of all ICs were displayed in Fig. 5(b) with decreasing correlation coefficients (from left to right, top to bottom). It is worthy to note that, in addition to IC8 and IC16, the IC3 also presented prominent weightings at the occipital region. To investigate the independence between the temporal waveform of IC3 and that reconstructed from IC8 and IC16, we used the onsets of target flickering sequence to segment these two waveforms into epochs followed by averaging. Fig. 6 shows that the averaged neural activity of the signal reconstructed from IC8 and IC16 exhibited evident FVEP at Oz, whereas that from IC3 did not reveal any relevant FVEP features, i.e., N1, P1, N2, P2, etc. Besides, their Fourier spectral patterns were clearly distinct. These suggest that the temporal waveform of IC3 would be the occipital alpha rhythm originated from occipital-parietal region and irrelevant to FVEP activities.

Fig. 7 demonstrates the detection of largest Amp_{p-v} when a subject was focusing binocularly on the stimulus '6'. The first panel shows the flickering sequence of '6', where vertical marks denote the flash onsets, and the reconstructed EEG signals at Oz from ICA. The temporal waveform in the panel of stimulus '6' was the averaged result of every 5 consecutive epochs based on the flash onsets in the flickering sequence of stimulus '6'. Signals in the remaining panels were generated in the same manner based on their own flash onsets. Since central FVEPs were synchronized to the flash onsets of target flickering sequence and the peripheral FVEP epochs, on the contrary, were asynchronized to such flash onsets, the averaged FVEPs induced from stimulus '6' resulted in the greatest Amp_{p-v} value and has been successfully segregated from the surrounding

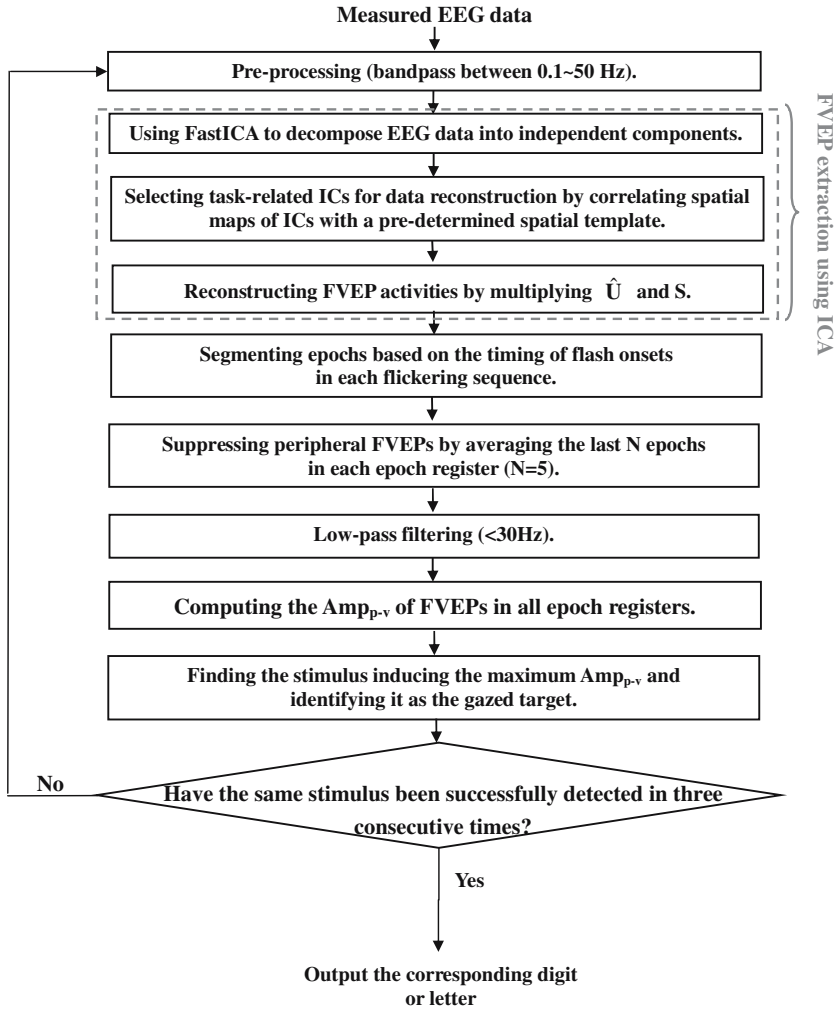

FIGURE 4. The overall data processing flowchart.

flickering sequences, i.e., stimuli '2', '5', '7', and '9'. The whole-head channel plot in Fig. 8 shows that FVEPs resulted from the flickering sequence of stimulus '6' can only be recognized at O1, O2, and Oz channels in the occipital area, which validates the feasibility of using random duration sequences in inducing human FVEPs.[29]

To further demonstrate the effectiveness of ICA, each subject was asked to gaze binocularly at the center of each flickering channel for 1-min recording. The detection of gazed-FC was performed continuously until all of them were processed. Different numbers of epochs were averaged to compare the with-ICA results to without-ICA results. Fig. 9 displays the mean accuracies, i.e., $N_{\text{correct}}/N_{\text{total}}$, and standard deviations of five subjects with 1, 5, 10, 15, 20, 25, 30, and 35 epochs being averaged. The mean accuracies of with-ICA results were 66.8%, 90.4%, 98.1%, and 99.7%, 100%, 100%, and 100%, respectively, compared to that of without-ICA results which were 27.8, 67.9, 91, 95, 96, 99, 99.6, and 100%, respectively. It is obvious that the accuracies increased dramatically with the aid of ICA and reached 100% accuracy after averaging 15 or more epochs. Accordingly, five-epoch average has been adopted to compromise the computation and accuracy in current system.

In the second experiment, each subject was requested to produce a string '0287513694E'. The letter 'B' was not used in this experiment since it will be reserved for the purpose of correcting the erroneous spelling. The ICA decomposition, reconstruction and gazed-FC detection, were executed in every second at a personal computer (CPU 3.0 GHz/1 GB RAM). Whenever each single input digit was confirmed for three consecutive times by the system, the subject was prompted by a voice feedback to proceed with next digit/letter. Table 1 lists the results from five subjects. All the five participants completed the string with minor errors, which were marked underlined. The accuracy was defined as the number of correct digits

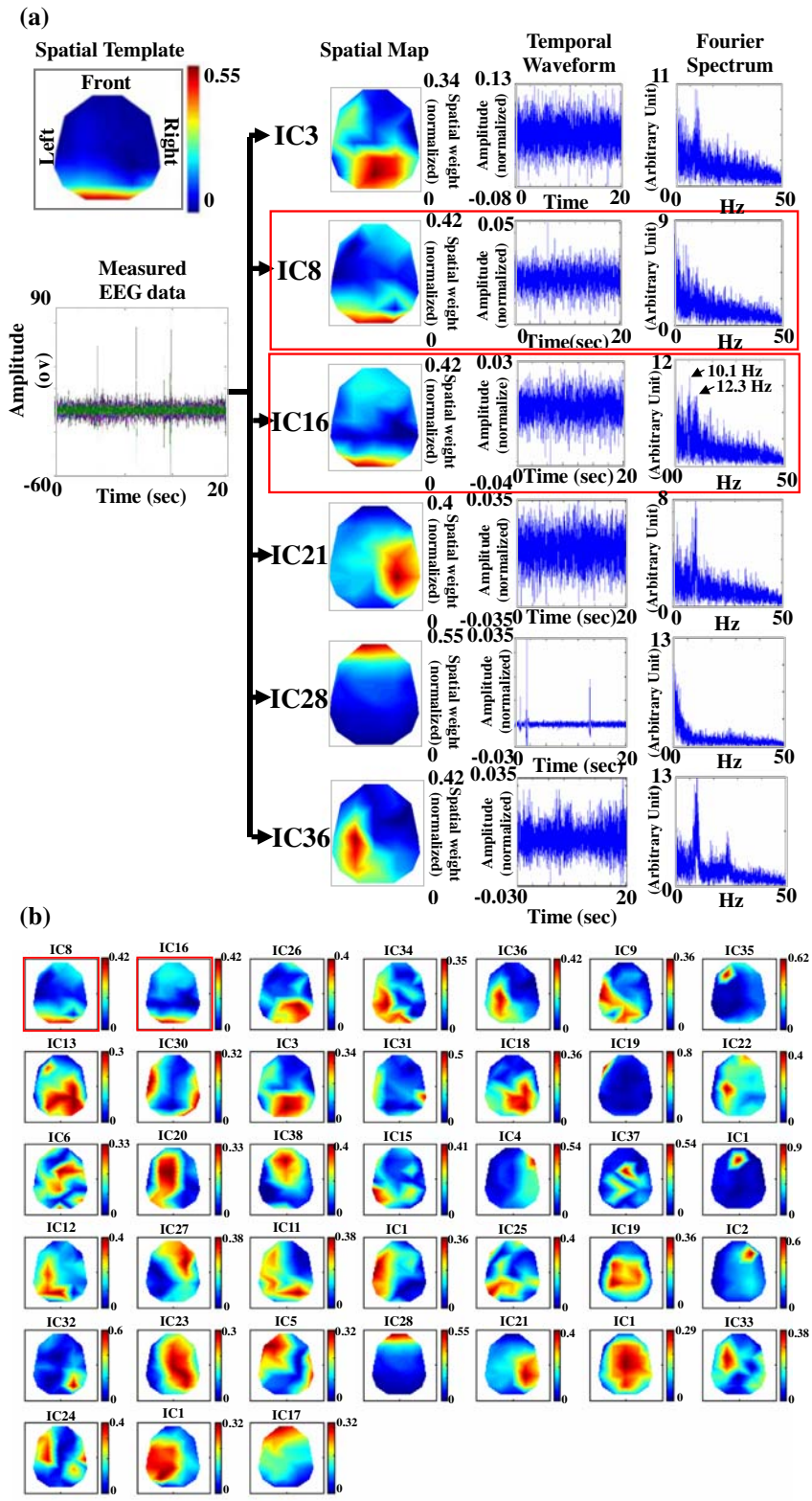

FIGURE 5. Extraction of FVEP activities using ICA. (a) The spatial maps, temporal waveforms, and Fourier spectra of six independent components (ICs) are presented. It can be observed that the occipital alpha rhythm (IC3), FVEP activities (IC8 and IC16), right sensorimotor mu rhythm (IC21), ocular artifact (IC28), and left sensorimotor mu rhythm (IC36) were decomposed into distinct ICs, respectively. IC8 and IC16 meet the component selection criteria and are used for data reconstruction. (b) The spatial maps of all ICs were displayed with decreasing correlation coefficients (from left to right, top to bottom).

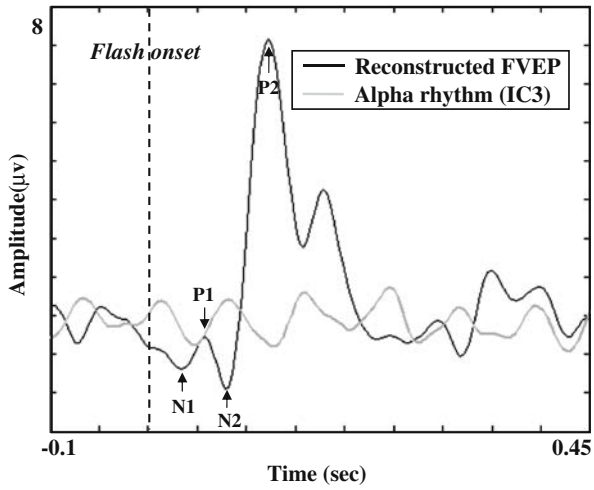

FIGURE 6. Averaged neural activity at Oz reconstructed from IC8 and IC16 and that from IC3 in Fig. 5. Compared the neural activity reconstructed from IC8 and IC16 with that from IC3, the former one presents evident FVEP. Accordingly, the temporal waveform of IC3 would be the occipital alpha rhythm originated from different neural sources irrelevant to FVEP activities.

and letter (N_{correct}) divided by the number of total digits and letters (N_{total}), i.e., $N_{\text{correct}}/N_{\text{total}}$. The command transfer interval for each subject, CTI, was defined as the total experimental time (T_{command}) divided by the number of total output digits and letters, i.e., $T_{\text{command}}/N_{\text{total}}$. The mean accuracy and the mean CTI over the five subjects were 83.3% and 6.14 s/command, respectively. The information transfer rate (ITR) was computed by

$$ITR = \log_2 N + P\log_2 P + (1 - P)\log_2[(1 - P)/(N - 1)],$$

where N is the total number of stimuli and P is the accuracy.[12] The mean ITR of the five subjects was 23.06 bits/min.

DISSCUSSION

When a subject performs the selection task by sequentially shifting his/her gaze on target flickering stimuli, the use of $\text{Amp}_{\text{p\_v}}$ from FVEP as a discernible neuro-electric feature in BCI has several advantages. First, the FVEP can be easily and stably obtained from most people.[19] Second, $\text{Amp}_{\text{p\_v}}$ in FVEP is time-locked and phase-locked to the stimuli and thus is a robust feature. Third, the quick response of $\text{Amp}_{\text{p\_v}}$ (around 100 ms–140 ms) is essential to accomplish an efficient BCI system. These merits enable us to implement a reliable and robust BCI system with high ITR.

In our FVEP studies, we experienced that the number of ICs, which were highly correlated with the spatial template (i.e., correlation coefficient > 0.8), was less than or equal to two for all five subjects, as illustrated by the IC8 and IC16 in Fig. 5(a). From the selected ICs and their Fourier spectrum, we did find that one of them is the most task-related IC, which consists of major VEP features. This is exemplified by the IC16 in Fig. 5(a) whose Fourier spectrum exhibits two clear peaks at 10.1 and 12.3 Hz (marked by arrows) and resembles the Fourier spectrum of the conventional FVEP which is an average of 100 trials (see Fig. 3(b)). In addition, the reconstructed signal based on ICA also allowed us to analyze inter-trial variability in latency and amplitude due to significantly improved SNR. Fig. 10 shows the raster plot of 118 single-trial FVEPs sorted by the latencies of P2 peaks. Values of the latencies and amplitudes of P2 peaks were 122.1 ± 4.3 ms and 9.43 ± 1.05 μv, respectively, and in line with other ICA groups.[8,24] This suggests that the cross-correlation technique in FMFVEP-based system[21] using a fixed-form VEP template with constant latency might be too stringent to optimally detect target stimulus.

Comparing the FVEP-based system with the SSVEP or FMFVEP systems, both the flickering sequence design and the translation algorithm in three systems are different. In our system, we used mutually independent flickering sequences and measured the peak-to-valley amplitude between N2 and P2 peaks as an index to identify the gazed visual stimulus. Simple averaging was applied to suppress the peripheral FVEP epochs asynchronized to the flash onsets of gazed flickering stimulus. The mean ITR of our system was 23.06 bits/min (Table 1). In the SSVEP-based system, each visual stimulus was designed to have its own flickering frequency. The designed flickering frequency should exclude the frequency range of 9–14 Hz in order to avoid the interference of occipital alpha rhythm.[2] Besides, higher flickering frequencies (> 20 Hz) were seldom used because of their low evoked amplitudes.[2] The gazed target was identified by finding the stimulus which contributed maximum power of SSVEP at Fourier spectrum. The reported ITR has achieved 27.15 bits/min.[2]

To compare the performance of FMFVEP method with the proposed FVEP-based system, we conducted two studies, one without noise and one with noise. The FVEP at Oz induced by gazing the flickering stimulus '6' was de-noised by ICA and was assumed as the noise-free data. The noise-free study is illustrated in Fig. 11 where (a) exhibits the "expected response" (lower panel) by convoluting the "FVEP template" (upper left panel), which is the conventional FVEP shown in Fig. 3(a), with the flickering sequence of stimulus '6' (upper right panel), and (b) displays the noise-free FVEP at Oz. One data segment of the

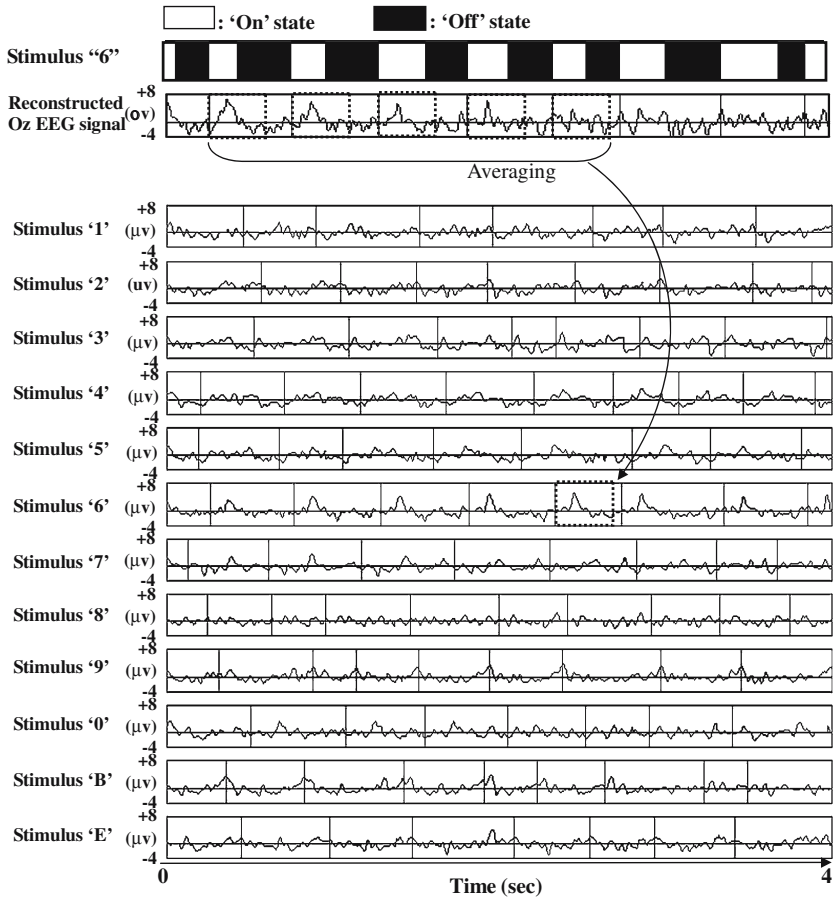

FIGURE 7. Extraction of central FVEP while a subject stared at stimulus '6'. The first panel shows the flickering sequence of '6', where vertical red marks denote the flash onsets, and the reconstructed EEG signals at Oz from ICA. The temporal waveform in the panel of stimulus '6' was the averaged result of every five consecutive epochs based on the flash onsets in the flickering sequence of stimulus '6'. Signals in the remaining panels were generated in the same manner based on their own flash onsets. It is obvious that the averaged FVEPs induced from stimulus '6' resulted in the greatest $Amp_{p\text{-}v}$ value compared with that in other panels.

"expected response", dark gray curve with 4000 samples (16 s) centered at the 5000th sample position in lower panel of Fig. 11(a), was used to cross-correlate with the curve in (b) and the result is shown in Fig. 11(c), where the position resulting in maximum cross-correlation was located at the 5002nd sample (light gray line), i.e., two samples (8 ms at 250 Hz sampling rate) shifted from the true position (dark gray line). To investigate the reasoning of incorrect result, the 4000-sample data segment of "expected response" centered at the true position (dark gray curve) and at the maximum-correlation position (light gray curve) were plotted, respectively, and superimposed on the noise-free FVEP (black curve) (see the upper panels of Fig. 11(d)). To see the jittering between the FVEP and "expected responses", the curves within window A and B were enlarged. The middle panel of Fig. 11(d) displays that the P2 peaks of the black curve lag behind the P2 peaks of dark and light gray curves, whereas the lower panel shows that the P2 peaks of black curve are between those of dark and light gray curves. This demonstrates that the cross-correlation method can be affected by the inter-trial variations, such as latencies, waveforms, peak amplitudes, etc., and results in incorrect detection. In addition, the cross-correlation method relies on lengthy data segment of "expected responses" to suppress internal/external noise or flickering sequences of non-gazed flickering stimuli. Fig. 11(e) shows the mean accuracies of the FMFVEP's cross-correlation method with respect to different data lengths of "expected response" in this study. To achieve 90% accuracy, the FMFVEP system required data length longer than 5.8 s while our simple averaging method merely needs five ON–OFF stimulus cycles, i.e., mean duration of 2.33 s. In other words, shorter reaction time can be achieved by the proposed FVEP-based system.

In the second study, randomly-generated Gaussian noise were added to the ICA-extract FVEP. Eleven levels of signal-to-noise ratio (SNR), ranging from 0,

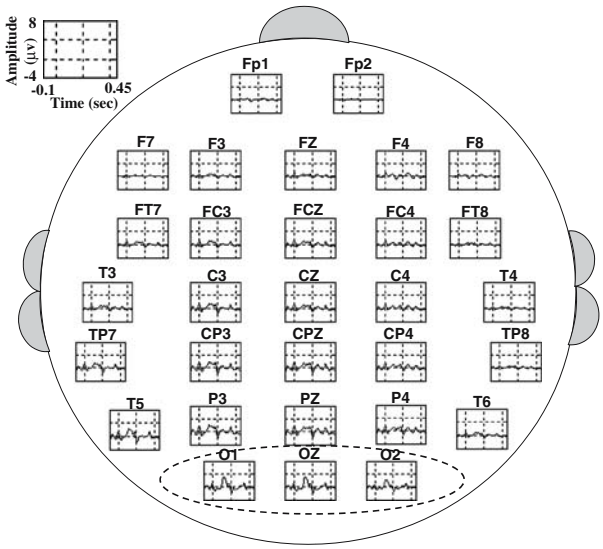

FIGURE 8. Whole-head channel plot of five-trial averaged FVEPs. The whole-head channel plot shows that FVEPs induced from the flickering sequence of stimulus '6' can only be recognized at O1, O2, and Oz channels in the occipital area, which validates the feasibility of using random duration sequences in inducing human FVEPs.

−2, −4, −6, −8, −10, −12, −14, −16, −18, and −20 dB, were tested. The mean hit rates over five subjects are shown in Fig. 12(a) where SNR levels of 0, −2, −4, −6, −8, −10, −12, −14, −16, −18, and −20 dB require 5, 5, 5, 5, 7, 8, 9, 13, 15, 22, and 25 trials, respectively, to achieve 90% accuracies, which correspond to 2.33, 2.33, 2.33, 2.33, 3.23, 3.72, 4.20, 6.06, 7.00, 10.26, and 11.66 s command transfer interval (CTI), respectively. The same noise levels were also tested in the FMFVEP system (Fig. 12(b)). For the SNR levels of 0, −2, −4, −6, −8, −10, −12, and −14 dB, the system requires 5.8, 6.6, 7.9, 8.4, 11.2, 13, 16.1, and 18.5 s, respectively, to achieve 90% accuracies, which were inferior to the FVEP-based system. When the SNR level was lower than −16 dB, 90% accuracy was not achievable if the data length of the "expected response" was shorter than 20 s.

It should be noted that the time window (550 ms) used in the proposed system can be shortened. The proposed method focused on detecting the amplitude difference between N2 and P2 peaks of FVEP, whose mean latencies were 87.2 ms and 125.8 ms, respectively, for five subjects. The late response, which occurred between 233.3 ms and 450 ms, and the early response induced by the next stimulus appear in one FVEP epoch were longer than necessary. Therefore, the response later than 200 ms was not taken into account, which has little effect after averaging. In fact, the purpose of introducing a fixed duration of 116.7 ms in the design of flickering sequence was to prevent the overlap of N2 and P2 peaks in two consecutive FVEPs. According to our data, the length of window can be chosen as short as 125.8 ms–87.2 ms = 38.6 ms in theory, and any longer window will be suitable to produce comparable results. The purpose of using of 550 ms time window was to display the overall picture of FVEP data as in the literature[8,17,20] so that the features of FVEPS can be better explained.

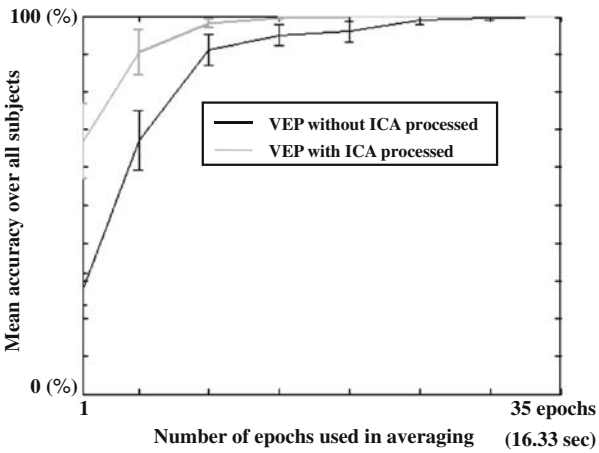

FIGURE 9. Comparison of with-ICA and without-ICA accuracies. Five subjects with 1, 5, 10, 15, 20, 25, 30, and 35 averaged epochs were used for comparison. The mean accuracies of with-ICA results were 66.8%, 90.4%, 98.1%, and 99.7%, 100%, 100%, and 100%, respectively, compared to that of without-ICA results which were 27.8%, 67.9%, 91%, 95%, 96%, 99%, 99.6%, and 100%, respectively. The accuracies increased dramatically with the use of ICA and reached 100% accuracy after averaging 15 or more epochs.

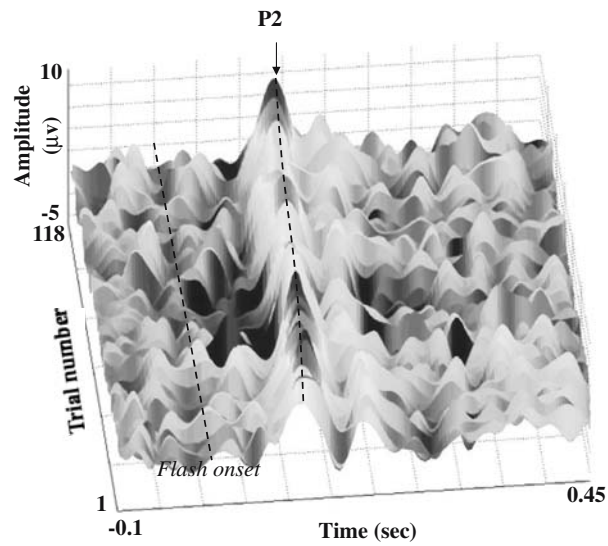

FIGURE 10. Raster plot of single-trial FVEPs extracted by ICA. 118 single-trial FVEPs were sorted by the latencies of P2 peaks. Values of the latencies and amplitudes of P2 peaks were 122.1 ± 4.3 ms and 9.43 ± 1.05 μv, respectively.

TABLE 1. The results of producing the string '0287513694E' from five subjects.

	Subjects	Output results (incorrect ones are underlined)	Total time (s)	CTI (s/command)	ITR (bits/min)	Mean detection rate ($N_{correct}/N_{total}$)
Skillful group	WCH	0287513694E	57	5.18	41.52	11/11 (100%)
	LSK	02875103694E	71	5.91	29.29	11/12 (91.7%)
	Average		64	5.55	35.40	95.95%
Naïve group	DMN	03280752139694E	93	6.20	17.65	11/15 (73.3%)
	SSC	02875137659E4E	91	6.50	19.3	11/14 (78.5%)
	CCC	02E8751436B984E	104	6.93	15.8	11/15 (73.3%)
	Average		96	6.33	17.58	75.03%
Average of two groups			83.2	6.14	23.06	83.36%

The familiarity of experiment and attention of subjects may affect the detection rates. Since only two subjects were participated in this task before (1-h training) while the other three were naïve it, we observed that experienced subjects have better concentration on the target stimulus than the naïve

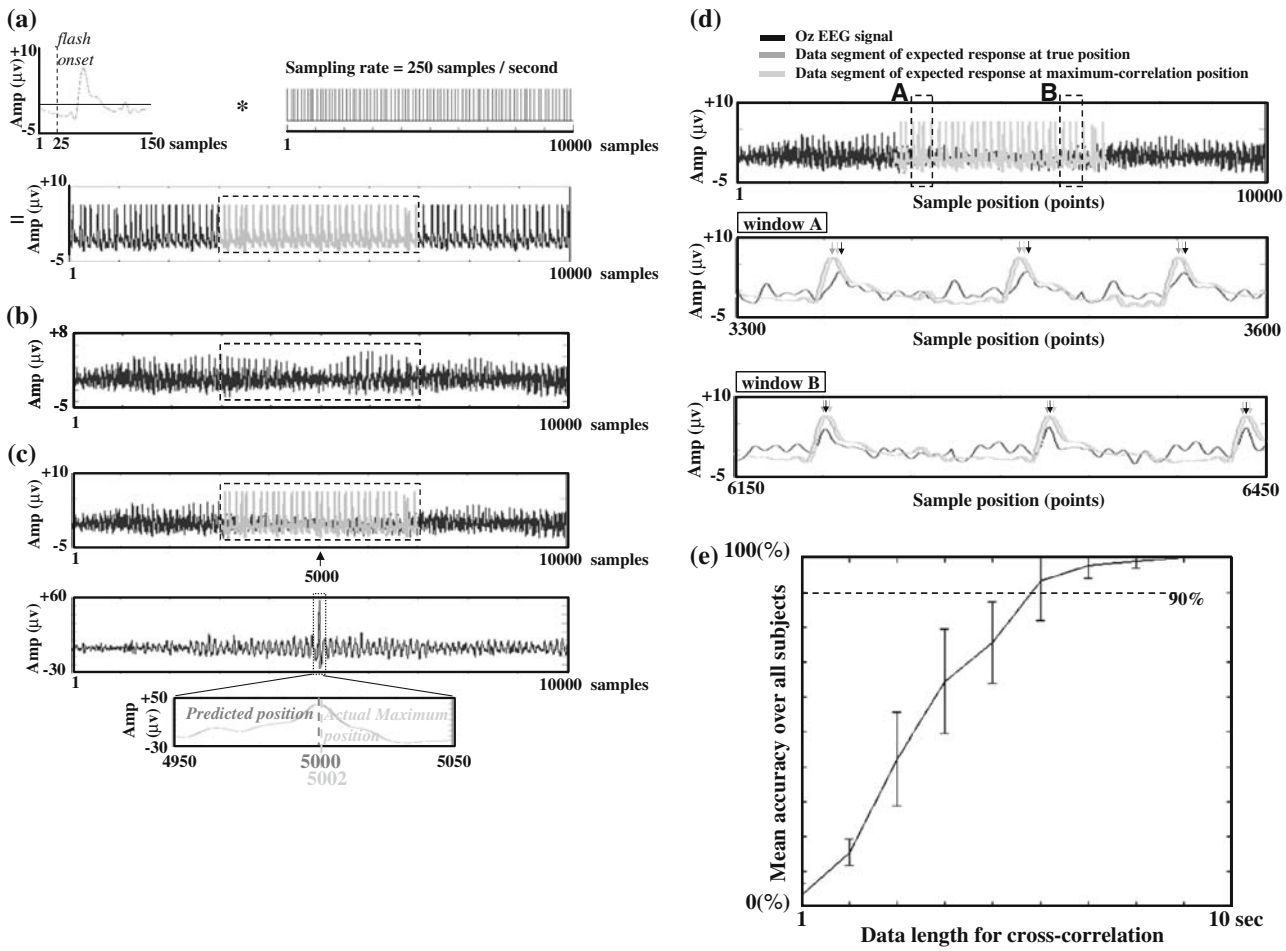

FIGURE 11. Application of the FMFVEP method to noise-free data induced by gazing at stimulus '6'. (a) The ''expected response'' (lower panel) created from convoluting the ''FVEP template'' (upper left panel) with the flickering sequence of stimulus '6' (upper right panel). (b) The noise-free (ICA-extracted) FVEP activities at Oz channel. (c) Upper panel: the noise-free FVEP activities (black curve) superimposed with the data segment of ''expected response'' at true position (dark gray curve). Middle panel: the cross-correlation result of the noise-free FVEP activities with the ''expected response''. Lower panel: enlarged display of the cross-correlation result where the position producing maximum cross-correlation was at the 5002nd sample (light gray line) and the true position was at 5000th sample (dark gray line). (d) Upper panel: the noise-free FVEP activities (black curve) superimposed with the data segment of ''expected response'' at true position (dark gray curve) and the data segment of ''expected response'' at maximum-correlation position (light gray curve). Middle panel: enlarged display of window A at upper panel. Lower panel: enlarged display of window B at upper panel. (e) Mean accuracies of the FMFVEP's cross-correlation method with respect to different data lengths of ''expected response''.

subjects who were distracted occasionally by surrounding non-target stimuli. For example, subject SSC has incautiously shifted his gaze on the wrong stimulus '7' after selected '3' and made another two mistakes in '5' and 'E' at (see Table 1). Overall results in Table 1 show that the skillful group (i.e., WCH and LSK) has faster mean CTI (5.55 s/command vs. 6.54 s/command) and better mean detection rate (95.95% vs. 75.03%) than the naïve group (i.e., DMN, SSC, CCC).

CONCLUSIONS

A FVEP-actuated BCI system has been proposed to generate strings for communication purpose. Subjects shift their gazes at flashing digits or letters on a monitor from which prominent FVEPs are induced as control signals. The salient features of proposed system are: (1) FVEP is a very stable and reliable neuro-electric signal to be detected; (2) task-unrelated components can be

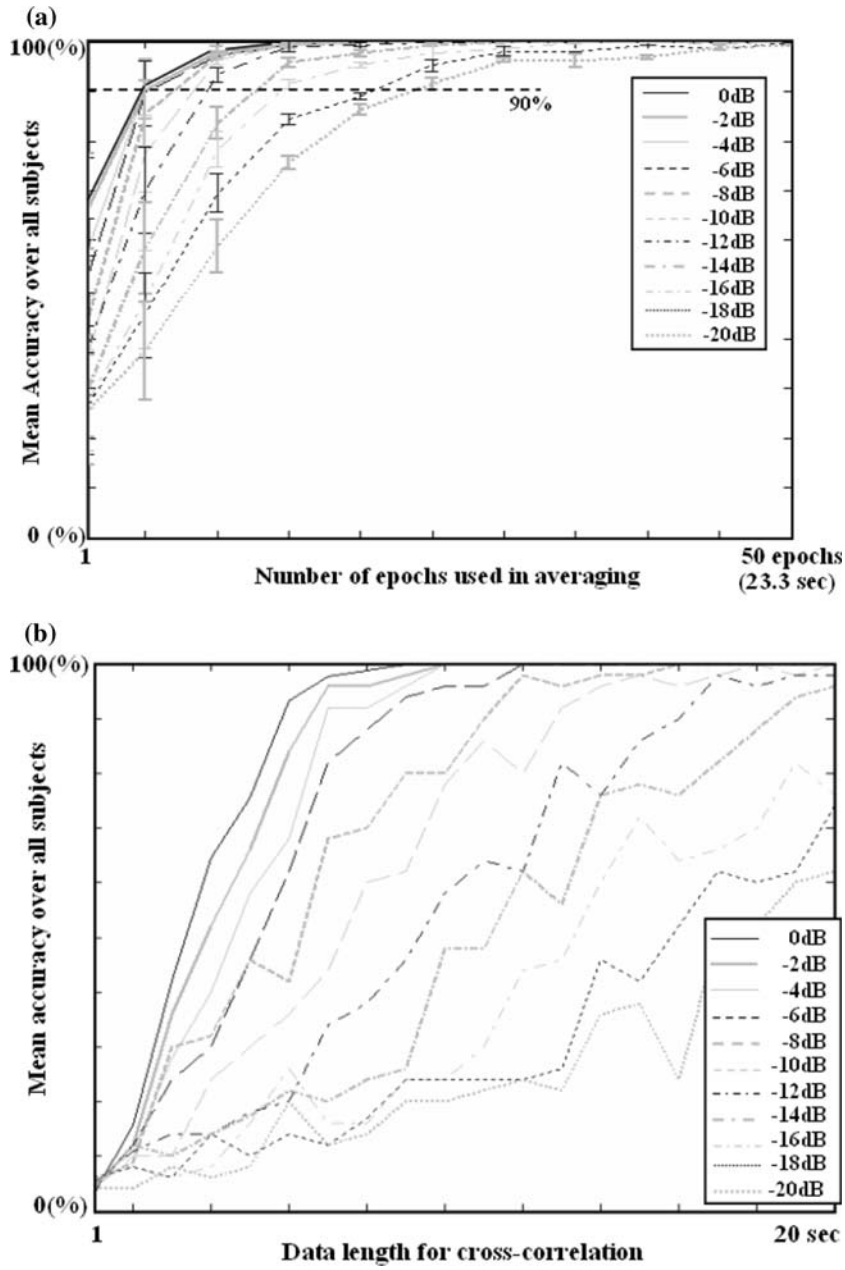

FIGURE 12. Comparison between the FVEP-based system and FMFVEP-based system under various SNR levels. (a) The mean accuracies of the FVEP-based system over the five subjects under various SNR levels. (b) The mean accuracies of the FMFVEP-based system over the five subjects under various SNR levels.

removed by ICA to improve the signal-to-noise ratio and thus accuracy; (3) Mutually independent flickering sequences have been designed to suppress the contamination induced from non-target stimuli; (4) satisfactory accuracy and information transformation rate have been achieved based on simple detection of the largest peak-to-valley FVEP amplitude. The proposed system can be used by disabled people as an efficient and reliable tool to communicate with external environments.

ACKNOWLEDGMENTS

This study was funded by the National Central University, National Science Council (94-2218-E-008-013, 94-2218-E-010-004) and National Health Research Institutes (NHRI-EX94-CD9401) of Taiwan.

REFERENCES

[1]Birbaumer, N., H. Flor, N. Ghanayim, T. Hinterberger, I. Iverson, E. Taub, B. Kotchoubey, A. Kubler and J. Perelmouter. A spelling device for the paralyzed. *Nature* 398:297–298, 1999.

[2]Cheng, M., X. Gao, S. Gao and D. Xu. Design and implementation of a brain-computer interface with high transfer rates. *IEEE T. Bio-Med Eng*. 10:1181–1186, 2002.

[3]Donchin, E, K. M. Spencer and R. Wilesinghe. The mental prosthesis: Assessing the speed of a P300-based brain-computer interface. *IEEE T. Rehabil. Eng*. 8:174–179, 2000.

[4]Duann, J. R., T. P. Jung, W. J. Kuo, T. C. Yeh, S. Makeig, J. C. Hsieh and T. J. Sejnowski. Single-trial variability in event-related BOLD signals. *Neuroimage* 15:825–835, 2002.

[5]Hung, C. I., P. L. Lee, Y. T. Wu, L. F. Chen, T. C. Yeh and J. C. Hsieh. Recognition of motor imagery electroencephalography using independent component analysis and machine classifiers. *Ann. Biomed. Eng*. 33:1053–1070, 2005.

[6]Hyvarinen, A., J. Karhunen, and E. Oja. Independent component analysis. New York: Wiley, New York, 2001.

[7]Hyvarinen, A. and E. Oja. A fast fixed-point algorithm for independent component analysis. *Neural Comput*. 9:1483–1492, 1997.

[8]Jung, T. P., S. Makeig, M. Westerfield, J. Townsend, E. Courchesne and J. T. Sejnowski. Analysis and visualization of single-trial event-related potentials. *Hum. Brain Mapp*. 14:166–185, 2001.

[9]Kao, Y. H., W. Y. Guo, Y. T. Wu, K. C. Liu, W. Y. Chai, C. Y. Lin, Y. H. Hwang, A. J. K. Liou, H. C. Cheng, T. C. Yeh, J. C. Hsieh and M. M. H. Teng. Hemodynamic segmentation of MR brain perfusion images using independent component, thresholding and Bayesian estimation. *Magn. Reson. Med*. 49:885–894, 2003.

[10]Lee, P. L., Y. T. Wu, L. F. Chen, Y. S. Chen, C. M. Cheng, T. C. Yeh, L. T. Ho, M. S. Chang and J. C. Hsieh. ICA-based spatiotemporal approach for signle-trial analysis of post-movement MEG beta synchronization. *Neuroimage* 20:2010–2030, 2003.

[11]Makeig, S., T. P. Jung, A. J. Bell, D. Ghahremani and T. Sejnowski. Blind separation of auditory event-related brain responses into independent components. *Proc. Natl. Acad. Sci. USA*. 94:10979–10984, 1997.

[12]Mason, S. G. and G. E. Birch. A brain-controlled switch for asynchronous control applications. *IEEE T. Bio-Med Eng*. 47:1297–1307, 2000.

[13]McKeown, M. J., S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell and T. J. Sejnowski. Analysis of fMRI data by blind separation into independent spatial components. *Hum. Brain Mapp*. 6:160–188, 1998.

[14]McKeown, M. and R. Radtke. Phasic and tonic coupling between EEG and EMG demonstrated with independent component analysis. *J. Clin. Neurophysiol*. 18:45–47, 2001.

[15]McSherry, J. W., C. L. Walters and J. D. Horbar. Acute visual evoked potential changes in hydrocephalus. *Electroen. Clin. Neuro*. 53:331–333, 1982.

[16]Muller-Gerking, J., G. Pfurtscheller and H. Flyvbjerg. Designing optimal spatial filters for single-trial EEG classification in a movement task. *Clin. Neurophysiol*. 110:787–798, 1999.

[17]Odom, J. V., M. Bach, C. Barber, M. Brigell, M. F. Marmor, A. P. Tormene and G. E. Holder. Vaegan: Visual evoked potentials standard (2004). *Doc. Ophthalmol*. 108:115–123, 2004.

[18]Pfurtscheller, G., C. Neuper, C. Guger, W. Harkam, H. Ramoser, A. Schlogl, B. Obermaier and M. Pregenzer. Current trends in Graz brain-computer interface (BCI) research. *IEEE T. Rehabil. Eng*. 8:216–219, 2000.

[19]Spehlmann, R. "Evoked potential primer." In: The transient VEP to diffuse light simuli, edited byMisulis, K. E., and T. Fakhoury. Stoneham: Butterworth publishers, 1985, pp. 135–142.

[20]Spehlmann, R. "Evoked potential primer." In: VEPs to other stimuli, edited by Misulis, K. E., andT. Fakhoury. Stoneham: Butterworth publishers, 1985, pp. 144–158.

[21]Sutter, E. E.. The brain response interface: Communication through visually-induced electrical brain responses. *J. Microcomput. Appl*. 15:31–45, 1992.

[22]Raitta, C., U. Karhunene, A. M. Seppalainen and M. Naukkarinen. Changes in the electroretinogram and visual evoked potentials during general anaesthesia. Albrecht von Graefes Arch. *Klin. Exp. Ophthalmol*. 211:139–144, 1979.

[23]Reilly, E. L., C. Kondo, J. A. Brunberg and D. B. Doty. Visual evoked potentials during hypothermia and prolonged circulatory arrest. *Electroen. Clin. Neuro*. 45:100–106, 1978.

[24]Tang, A. C., B. A. Pearlmutter, N. A. Malaszenko and D. B. Phung. Independent components of magnetoencephalography: Single-trial response onset times. *Neuroimage* 17:1773–1789, 2002.

[25]Tesche, C. D., M. A. Unsitalo, R. J. Ilmoniemi, M. Huotilainen, M. Kajola and O. Salonen. Signal-space projections of MEG data characterize both distributed and well-localized neuroal sources. *Electroencephalogr. Clin. Neurophysiol*. 95:189–200, 1995.

[26]Trojaborg, W. and E. O. Jorgensen. Evoked cortical potentials in patients with "isoelectric" EEGs. *Electroen. Clin. Neuro*. 35:301–309, 1973.

[27]Uhl, R. R., K. C. Squires, D. L. Bruce and A. Starr. Effect of halothane anesthesia on the human cortical visual evoked response. *Anesthesiology* 53:273–276, 1980.

[28]Wilson, W. B.. Visual-evoked response differentiation of ischemic optic neuritis from the optic neuritis of multiple sclerosis. *Am. J. Ophthalmol*. 86:530–535, 1978.
[29]Wolpaw, J. R., N. Birbaumer, D. J. McFarland, G. Pfurtscheller and T. M. Vaughan. Brain-computer interfaces for communication and control. *Clin. Neurophysiol*. 113:767–791, 2002.
[30]York, D. H., M. W. Pulliam, J. G. Rosenfeld and C. Watts. Relationship between visual evoked potentials and intracranial pressure. *J. Neurosurg*. 55:909–916, 1981.