Journal of Neuroscience Methods 197 (2011) 180–185

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Classification of selective attention to auditory stimuli: Toward vision-free brain–computer interfacing

Do-Won Kim [a], Han-Jeong Hwang [a], Jeong-Hwan Lim [a], Yong-Ho Lee [b], Ki-Young Jung [c], Chang-Hwan Im [d,*]

[a] Department of Biomedical Engineering, Yonsei University, Wonju, Republic of Korea
[b] Korea Research Institute of Standard and Science, Daejeon, Republic of Korea
[c] Department of Neurology, Korea University College of Medicine, Seoul, Republic of Korea
[d] Department of Biomedical Engineering, Hanyang University, 17 Haengdang-dong, Seongdong-gu, Seoul 133-791, Republic of Korea

ARTICLE INFO

Article history:
Received 31 August 2010
Received in revised form 6 February 2011
Accepted 8 February 2011

Keywords:
Brain–computer interface (BCI)
Auditory steady-state response (ASSR)
Selective attention
Amyotrophic lateral sclerosis (ALS)
Completely locked-in syndrome

ABSTRACT

Brain–computer interface (BCI) is a developing, novel mode of communication for individuals with severe motor impairments or those who have no other options for communication aside from their brain signals. However, the majority of current BCI systems are based on visual stimuli or visual feedback, which may not be applicable for severe locked-in patients that have lost their eyesight or the ability to control their eye movements. In the present study, we investigated the feasibility of using auditory steady-state responses (ASSRs), elicited by selective attention to a specific sound source, as an electroencephalography (EEG)-based BCI paradigm. In our experiment, two pure tone burst trains with different beat frequencies (37 and 43 Hz) were generated simultaneously from two speakers located at different positions (left and right). Six participants were instructed to close their eyes and concentrate their attention on either auditory stimulus according to the instructions provided randomly through the speakers during the inter-stimulus interval. EEG signals were recorded at multiple electrodes mounted over the temporal, occipital, and parietal cortices. We then extracted feature vectors by combining spectral power densities evaluated at the two beat frequencies. Our experimental results showed high classification accuracies (64.67%, 30 commands/min, information transfer rate (ITR) = 1.89 bits/min; 74.00%, 12 commands/min, ITR = 2.08 bits/min; 82.00%, 6 commands/min, ITR = 1.92 bits/min; 84.33%, 3 commands/min, ITR = 1.12 bits/min; without any artifact rejection, inter-trial interval = 6 s), enough to be used for a binary decision. Based on the suggested paradigm, we implemented a first online ASSR-based BCI system that demonstrated the possibility of materializing a totally vision-free BCI system.

1. Introduction

Brain–computer interface (BCI, sometimes referred to as brain–machine interface) is a technology that translates brain signals into simple commands that can control external devices or into messages with which one can communicate (Wolpaw et al., 2002). The major targets of BCI systems have been disabled individuals who cannot freely move or control specific parts of their body because of serious neurological disease or injury, such as amyotrophic lateral sclerosis (ALS, also referred to as Lou Gehrig's disease) or brainstem stroke. Since many of these patients do not have cognitive impairment, their brain activity can be used as a source for communication. Among the various types of human brain mapping techniques used for implementing BCI systems, such as functional magnetic resonance imaging (fMRI) and near infrared spectroscopy, electroencephalography (EEG) has been the most widely used modality because it is noninvasive, economical, harmless, and readily applicable (Wang et al., 2004; Hoffmann et al., 2008; Nijboer et al., 2008).

One of the most important factors necessary for materializing a successful EEG-based BCI system is the selection of appropriate mental tasks that can elicit distinct task-specific brain activity patterns. To translate the acquired neural signals into appropriate commands, various experimental paradigms and tasks have been introduced, including visual attention tasks such as the P300 speller (Farwell and Donchin, 1988; Krusienski et al., 2008; Sellers and Donchin, 2006); steady state neural responses elicited while one is gazing a certain visual stimulus flickering with a specific frequency (steady state visual evoked potential: SSVEP) (Lalor et al., 2005; Lin et al., 2006; Middendorf et al., 2000); mental tasks associated with motor imagery (Decety and Ingvar, 1990; Jeannerod and Frak, 1999; Pfurtscheller and Neuper, 1997) or mental calculation

* Corresponding author. Tel.: +82 2 2220 2322; fax: +82 2 2296 5943.
E-mail address: ich@hanyang.ac.kr (C.-H. Im).

0165-0270/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2011.02.007

(Keirn and Aunon, 1990; Penny et al., 2000), and so on. Most of the mental tasks and paradigms listed above use visual stimuli, visual feedback, or both and are thereby applicable only to patients whose visual function is not impaired. In practice, however, some patients with severe neurological disorders, such as ALS and completely locked-in state (CLIS), often have difficulty controlling their voluntary extraocular movements or fixing their gaze on specific visual stimuli. Even for those who have normal visual function, gazing at stimuli for a long time can easily cause fatigue or loss of concentration. Moreover, EEG signals recorded at frontal electrodes can be contaminated by electrooculogram (EOG) elicited by eye-blinking and eyeball movements. A recent experimental study demonstrated that the performance of the P300-based speller paradigm can be substantially influenced by eye gaze (Brunner et al., 2010), which strongly suggests that the use of visual stimuli or cues might not be appropriate for those who have difficulty in gazing at specific target stimuli. In other mental task paradigms that do not directly use visual stimuli, visual cues or feedbacks are generally provided to the participants so as to instruct or assist them in performing the given mental tasks (Hwang et al., 2009). Even in such cases, the recorded signals can be contaminated by unwanted visual evoked responses. Therefore, developing new BCI paradigms that are not dependent on visual stimuli remains one of the challenging issues in modern BCI research (Nijboer et al., 2008).

To overcome the limitations of conventional BCI paradigms, some researchers have turned to auditory stimuli (Hill et al., 2005; Kanoh et al., 2008; Lopez et al., 2009; Klobassa et al., 2009; Schreuder et al., 2010) as an alternative to visual stimuli. Most of the previous studies used auditory oddball paradigms, which share most of the basic concepts with conventional visual BCI paradigms. Two of the earliest studies (Hill et al., 2005; Kanoh et al., 2008) independently introduced an auditory BCI paradigm in which the authors attempted to discriminate "attended" brain responses from "unattended" ones when two simultaneous auditory oddball streams were presented to subjects. In a study by Hill et al. (2005), deviant sounds were generated alternatively at either a right or left sound source, and subjects were asked to concentrate on one of the two sound sources. They extracted the feature vectors from the changes in the amplitude of the averaged event-related potential (ERP). Kanoh et al. (2008) used a similar paradigm, where the subjects were instructed to concentrate their attention on one of two oddball audio streams with different frequencies presented alternately with a short inter-stimulus interval. They used the peak amplitudes of P300 and mismatch negativity (MMN) as the feature vectors to classify the subject's selective attention. Recently, Halder et al. (2010) refined the auditory oddball paradigm and evaluated various auditory stimuli with different volumes, pitches, or directions.

Another group of researchers attempted to modify the P300 speller paradigm, which is a well-established protocol in BCI research (Donchin et al., 2000), into an auditory version (Klobassa et al., 2009). Instead of presenting matrix-type visual stimuli, they presented different environmental sounds to participants and detected which sound the participants were attending to. Spatial hearing was also adopted as a new auditory BCI paradigm (Schreuder et al., 2010), which used eight speakers spatially distributed around a participant and detected a single sound source that the participant concentrated on.

Apart from the oddball paradigms or modified oddball paradigms (P300 speller), (Lopez et al., 2009) investigated whether the auditory steady-state response (ASSR) is modulated by auditory selective attention (ASA) to a specific sound stream and discussed the possibility of using the ASSR as a new BCI paradigm. They provided eight participants with two amplitude-modulated (AM) sound streams (1 kHz and 2.5 kHz) with different modulation frequencies (38 Hz and 42 Hz) to both ears simultaneously (1 kHz tone with a 38 Hz modulation frequency for the left ear and 2.5 kHz tone with a 42 Hz modulation frequency for the right ear). The participants were then asked to either concentrate their attention on the stimulus from the left ear or ignore both auditory stimuli according to the instructions appearing on a monitor. In six out of eight participants, the spectral density of alpha rhythm was inversely proportional to that of the modulation frequency for the left ear (38 Hz), providing evidence that selective attention can modulate ASSR. They also showed, using the self organizing map (SOM) method, that the attended and ignored conditions could be clearly classified into two clusters, demonstrating the possibility of using ASSR modulated by auditory selective attention as a new BCI paradigm.

In the present study, inspired by the pilot study of (Lopez et al., 2009), we further investigated whether ASSR can be a feasible feature for a practical BCI system by implementing a modified BCI paradigm to classify one's auditory selective attention and by evaluating the classification accuracy of the BCI system. Similarly to the previous study, six participants were presented with two pure tone burst trains (2.5 kHz and 1 kHz) with different beat frequencies (37 Hz and 43 Hz) to both sound fields simultaneously (2.5 kHz tone with a 37 Hz beat frequency for the left sound field and 1 kHz tone with a 43 Hz beat frequency for the right sound field). In our modified paradigm, the participants were asked to close their eyes and concentrate their attention on either auditory stimulus according to the instructions provided randomly through speakers during the inter-stimulus interval (ISI). Indeed, to the best of our knowledge, our paradigm is one of the first auditory BCI paradigms that did not use any visual information during the entire experiment. Our paradigm can be regarded as an auditory version of the conventional SSVEP-based BCI paradigms that use multiple spatially separated visual stimuli flickering at different frequencies (Bin et al., 2009; Luo and Sullivan, 2010). Similarly to the conventional SSVEP-based BCI systems, we extracted feature vectors from spectral power densities evaluated at the two beat frequencies. We then classified the participants' selective attention and evaluated the accuracy of the ASSR-based BCI system. Furthermore, we used the proposed paradigm and analysis methods to implement an online ASSR-based BCI system to further demonstrate whether our paradigm could be used as a successful BCI paradigm.

2. Methods

2.1. Participants

Six healthy volunteers (one female and five male, mean age 25.0 ± 5.0 years) were recruited among the graduate and undergraduate students in the Department of Biomedical Engineering of Yonsei University. Before the experiment, all participants were given a detailed, written summary of the experimental procedures. Participants signed a written consent and received adequate reimbursement for their participation. The study protocol was approved by the Institutional Review Board (IRB) of Yonsei University, Korea.

None of the participants reported neurological or psychiatric disorders or previous head injury that might affect the experiment. It was also confirmed that all subjects had normal or corrected-normal vision and normal hearing. Most participants except one (LH) were reported as right-handers. No subject had any previous experience or knowledge of BCI studies. All experiments were conducted in the Bioelectromagnetics and Neuroimaging Laboratory of Yonsei University.

2.2. Auditory stimuli

ASSR is a brain electrical response elicited when one is hearing periodic amplitude modulated sinusoidal tones or click sound trains (Picton et al., 2003; Ross et al., 2000, 2002, 2003). ASSR generally shows increased spectral density around the modulation frequency of the sound stream. The optimal modulation frequency has been reported as values ranging from 30 Hz to 50 Hz, peaking around 40 Hz (Engelien et al., 2000; Pastor et al., 2002; Picton et al., 1987). Therefore, to obtain a sufficient signal-to-noise ratio (SNR) of ASSR, we chose two frequencies around 40 Hz, 37 Hz and 43 Hz, as the modulation frequencies (beat frequencies in the present study). The carrier frequencies of the two auditory stimuli were set to 2.5 kHz and 1 kHz, respectively, so that the subjects could easily distinguish each sound stream (Lopez et al., 2009). We used pure tone burst trains; each generated using MATLAB (The MathWorks, Natick, MA, USA, Version 7.7.0) at a sampling rate of 44,100 Hz. The pulse widths of the 37 Hz and 43 Hz pure tone pulses were 13.5 ms and 11.6 ms, respectively. The duration of each trial was 20 s. The auditory stimulus for each trial was exported in the waveform audio file format (*.wav).

2.3. Experimental protocols

Participants sat in a comfortable armchair in front of a pair of speakers (BR-1000A, Britz International, Paju, Kyunggi-Do, Korea), which were positioned 80 cm apart. The participants were asked to adjust the position of the chair to a comfortable location while maintaining equal distance (less than 60 cm from the speakers) from the two speakers (see Fig. 1). In each trial, the participants were presented with 2.5 kHz tone burst trains with 37 Hz beat frequency for their left sound field and 1 kHz tone burst trains with 43 Hz beat frequency for their right sound field. Since the stimuli were similar to those frequently used to elicit ASSR in previous studies (Picton et al., 2003; Lopez et al., 2009), ASSRs peaking around 37 Hz and 43 Hz were expected. Subjects were asked to close their eyes and remain as still as possible, particularly during the acquisition intervals.

Fig. 2 shows the experimental paradigm used for the EEG recordings in the present study. One segment of the auditory stimulus lasted for 20 s and a random interval of 6–10 s was inserted between each trial. Two seconds before the stimulus onset, five pulses of pure tone sounds were generated randomly from either the left or right side, to indicate which sound source they were to concentrate on. The five pulses of pure tone sounds had the same carrier frequency and beat frequency as the main auditory stimulus (2.5 kHz carrier frequency and a 37 Hz beat frequency for the left sound field; 1 kHz carrier frequency and a 43 Hz beat frequency for the right sound field) to help the participants recognize the direction of the stimulus more accurately. Our paradigm was implemented with TeleScan 2.2 for Windows (Laxtha, Inc., Daejoen, Korea), which was also used for the EEG data acquisition.

Each session consisted of 25 trials and lasted for approximately 10 min. Before the recording, one training session was performed to familiarize participants with the paradigm. The main experiment was performed in two sessions with a 10-min inter-session rest. In total, we acquired EEG data sets for 50 trials: 25 for selective attention to the left-sided stimulus and the other 25 trials for selective attention to the right-sided stimulus.

2.4. Data acquisition and processing

Electrodes were attached on the participants' scalp according to the international 10–20 system. The EEG signals were acquired at four electrodes (Cz, Oz, T7, T8), which represent the motor, visual, and auditory cortical areas, using a multi-channel EEG acquisition system (WEEG-32, Laxtha Inc., Daejeon, Korea) in a dimly lit, soundproof room. The four electrodes were selected considering previous ASSR studies (Lopez et al., 2009; Skosnik et al., 2007; Ross et al., 2000). Two midline electrodes, Cz and Oz, were also selected in the

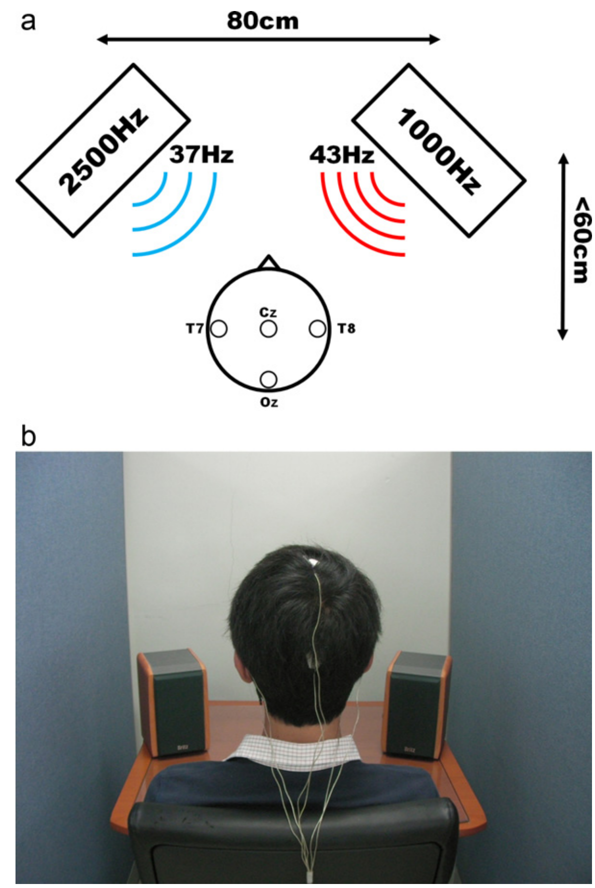

Fig. 1. Overall experimental environment: (a) a schematic diagram to elucidate the experimental environment. Two speakers were placed 80 cm apart. The participants were asked to adjust the position of the chair to a comfortable location while maintaining equal distance (less than 60 cm from the speakers) from the two speakers. The participants were presented with 2.5 kHz tone burst trains with a 37 Hz beat frequency for their left ear and 1 kHz tone burst trains with a 43 Hz beat frequency for their right ear; (b) a screenshot of the experiment showing one of the subjects (JP) sitting in front of two speakers with four EEG electrodes attached on his scalp.

Lopez et al.'s study (2009) that investigated whether ASSR could be modulated by selective attention. (Skosnik et al., 2007) investigated the effect of selective attention on the gamma-band ASSR using multiple electrodes, and they reported statistically meaningful ASSR value at Cz electrode. Two temporal lobe electrodes, T7 and T8, were selected according to a report by (Ross et al., 2000), in which their magnetoencephalography (MEG) data demonstrated that the ASSR with origin in primary auditory cortex could be enhanced during specific attention. The sampling rate was set at 512 Hz in all experiments. Most of the experiments were performed

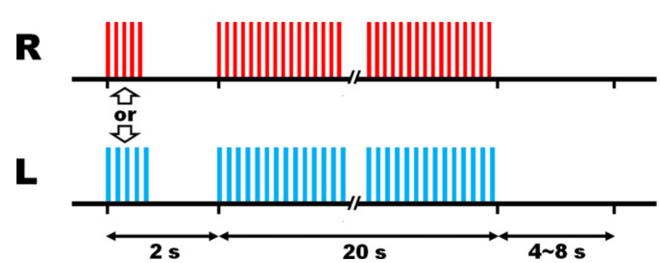

Fig. 2. The experimental paradigm used in the present study. One segment of the auditory stimulus lasted for 20 s and a random interval of 4–8 s was inserted between each trial. The left speaker generated 2.5 kHz pure tone burst trains with a 37 Hz beat frequency and the right speaker generates 1 kHz pure tone burst trains with a 43 Hz beat frequency. Two seconds before the stimulus onset, five pulses of pure tone sounds were generated randomly from either the left or right speaker.

at night to prevent unexpected noises that might occur during the experiments. The ground electrode was placed behind the left ear with the reference electrode on the opposite side.

The raw EEG data were segmented into 20-s epochs from the beginning of the main auditory stream. No preprocessing methods, such as re-referencing, band-pass filtering, or artifact rejection, were applied to the present analysis. The frequency spectrums of each epoch were calculated using the fast Fourier transform (FFT) algorithm with a 1 s long sliding window with a 50% overlap. The estimated frequency spectrums were accumulated and averaged over time for each epoch.

2.5. Feature selection and classification

As candidates of feature vectors, we first evaluated the EEG spectral densities of each electrode averaged over 37 ± 1 Hz (denoted as Cz_{37}, Oz_{37}, $T7_{37}$, $T8_{37}$) and 43 ± 1 Hz (Cz_{43}, Oz_{43}, $T7_{43}$, $T8_{43}$). We also evaluated the ratios between all possible pairs of spectral densities evaluated at the same modulation frequency ($Cz_{37}/T7_{37}$, $Cz_{37}/T8_{37}$, Cz_{37}/Oz_{37}, $T7_{37}/T8_{37}$, $T7_{37}/Oz_{37}$, $T8_{37}/Oz_{37}$, $Cz_{43}/T7_{43}$, $Cz_{43}/T8_{43}$, Cz_{43}/Oz_{43}, $T7_{43}/T8_{43}$, $T7_{43}/Oz_{43}$, $T8_{43}/Oz_{43}$) as well as the ratios between the spectral powers of each electrode evaluated at different modulation frequencies (Cz_{37}/Cz_{43}, $T7_{37}/T7_{43}$, $T8_{37}/T8_{43}$, Oz_{37}/Oz_{43}).

To investigate the changes in classification accuracy with respect to the number of feature vectors, we calculated the classification accuracy for all possible combinations of the 24 feature candidates listed above, assuming the number of selected features to be one, two, or three. To show the influence of the analysis window sizes (or analysis interval sizes) on the classification accuracy, we also tested different analysis window sizes (2–20 s from the main auditory stimulus onset with a step size of 1 s). Since the first second might contain unwanted components elicited by the transition of attention and preparation for performing the task, the results for window size of 1 s were not presented in the result section.

For the classification, we used a 10-fold cross-validation method considering the small number of trials. We first divided the 50 trials into 10 equal-size folds, and for each validation 45 trials were used as a reference data set and the other 5 trials were used as a test set. For each trial of the test set, Euclidean distances from the average feature vectors (each averaged to the left and right stimuli) computed on the reference data set were compared, and the trial was assigned to a class based on whichever had the shorter distance. The cross-validation was done separately for each of all possible feature sets.

3. Results

Fig. 3 shows the variations in classification accuracy averaged over the six participants with respect to the analysis window sizes and the number of feature vectors (please see the Supplementary Table file for values). We observed that higher classification accuracy could be obtained when larger numbers of feature vectors were used for the classification. The classification accuracy nearly monotonically increased with respect to the analysis window sizes, but after approximately 10 s the accuracy no longer increased. Since short analysis window size guarantees more possible commands per minute, the analysis window size of approximately 10 s was the most appropriate.

Fig. 4 shows variations in the classification accuracy evaluated for each participant with respect to the analysis window size when three feature vectors were selected (please see the Supplementary Table file for values). The six individual graphs show similar and consistent shape with those of Fig. 3 and show very small differences in the overall averaged classification accuracies, with a standard deviation of only 2.11%. The maximum classification accuracy of each participant was found at different analysis window sizes and varied from 80% to 92%. The average of the maximum classification accuracy of each subject was $86.33 \pm 3.54\%$, and the analysis window size that resulted in the highest accuracy was 14.00 ± 2.94 s. Subject SB showed the highest classification accuracy value (92%) among all of the participants, while subject JP showed the highest overall classification accuracy (81.26%). Although our cross-validation results using small number of trials might be somewhat biased for specific feature sets, the high classification accuracy consistently exceeding the chance level (50%) demonstrates the possibility of using ASSR for the binary decision of BCI.

In the supplementary tables, we listed all of the selected feature sets that resulted in the highest classification accuracy for each analysis window size when three feature vectors were selected (please find Supplementary Tables 1–6 that summarize the selected feature sets for each of six participants). The data revealed that the basis features evaluated from simple spectral density computations (Cz_{37}, Oz_{37}, $T7_{37}$, $T8_{37}$, Cz_{43}, Oz_{43}, $T7_{43}$, and $T8_{43}$) were selected more frequently than the other derived features, indicating that the derived features play an auxiliary role, helping to increase the overall classification accuracy. The most frequently selected features were Cz_{43} (18.14%), followed by Cz_{37} (14.06%), $T8_{37}$ (9.70%), Oz_{43} (8.01%), and Oz_{37} (7.45%), which were counted from all the selected feature set tables for analysis window sizes larger than 10 s. The frequently selected features were mostly extracted from

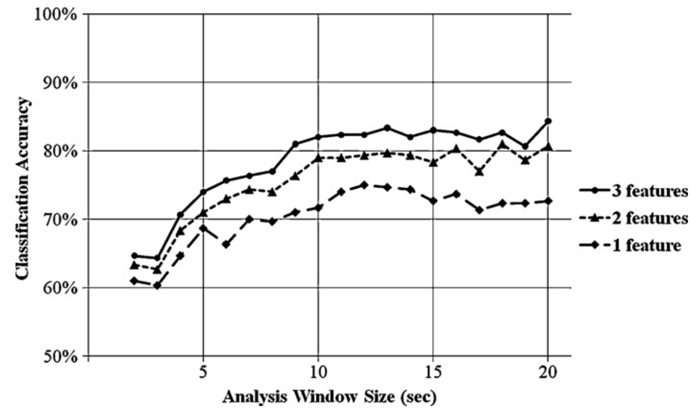

Fig. 3. Classification accuracy averaged over six participants with respect to different analysis window sizes and different numbers of feature vectors (1, 2, and 3). The exact values are summarized in Supplementary Table 7).

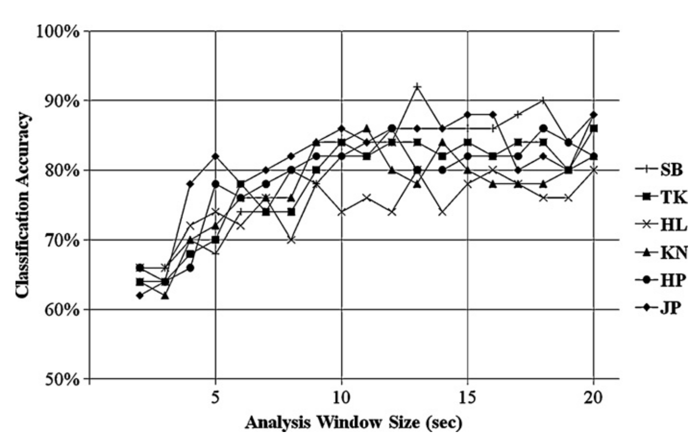

Fig. 4. Classification accuracy for each participant with respect to the analysis window sizes when three feature vectors were selected. The exact values are summarized in Supplementary Table 8).

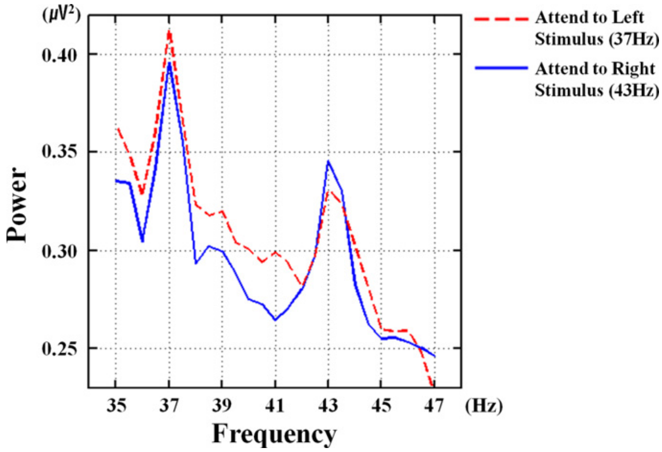

Fig. 5. An EEG spectral density plot at Cz electrode averaged across the participants with respect to two different conditions (red dashed line: attending to the left auditory stimulus condition; blue solid line: attending to the right auditory stimulus condition). The fast Fourier transform (FFT) algorithm with the same parameters used for the feature extraction was used for the spectral density calculation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

one of the midline electrodes, specifically in Cz electrode, which coincides well with the previous study that reported the most significant increment of ASSR in the Cz electrode (Skosnik et al., 2007).

To further verify that the power spectral density of each beat frequency would be modulated by selective attention, we calculated the power spectral density of EEG signals recorded at Cz electrode. Fig. 5 shows the power spectral density at Cz electrode averaged across the participants with respect to two different conditions (attending to left auditory stimulus and attending to right auditory stimulus). To evaluate the power spectral density, the FFT algorithm with the same parameters used for the feature extraction was applied. The EEG power spectral density plot depicted in Fig. 5 showed two clear ASSRs peaking at 37 Hz and 43 Hz. It could be observed from the figure that the EEG spectral density was modulated by auditory selective attention to a specific sound source, demonstrating that switching attentions between two different sounds would generate classifiable feature vectors.

We also implemented a pilot online ASSR-based BCI system and tested it to one of the participants (JP, female, 24 years old). Right before the online experiment, we selected an optimal feature set from a preliminary offline experiment. The experimental paradigm and analysis methods used for the feature selection were identical to those of previous offline experimental studies, except that the location of the reference electrode was moved from left ear to the participants' forehead. This change was made to avoid the potential influence of the reference electrode on the laterality between the electrodes T7 and T8. Since the participant was asked to close her eyes during the entire offline and online experiments, we confirmed that EOG artifact did not affect the recorded signals.

In our online experiment, the participant was instructed verbally to attend to either auditory stimulus, left stimulus or right stimulus, in a random order. After the instruction was made, the experimenter manually turned on a switch that starts generating two different tone burst trains from speakers located on the left and right sides of the participant. Then, the main computer system started recording the EEG signals, and at the same time calculated the values of 3 feature vectors. After 10 s from the beginning of the recording, our BCI system classified the participant's selective attention in real time and displayed the decision on the monitor screen so that the instructor can evaluate the result. All the analysis methods were identical to those used in the offline analyses. Since the participant was asked to close her eyes during the experiment, she could not have any information on whether the previous decision was right or wrong. In our pilot online experiment, we did not provide the participant with any feedbacks as they might affect her attention. The online experiment consisted of 14 continuous trials (7 for right stimulus and 7 for left stimulus) and showed a fair classification accuracy of 71.4%. The readers can watch the full video of our online experiment from the supplementary movie file attached in this article.

4. Discussion

In the present study, we investigated whether ASSR modulated by selective attention to a specific sound stream can be used to create a practical auditory BCI system, with the goal of classifying the intentions of individuals who have difficulty controlling their vision. Inspired by the conventional SSVEP-based BCI paradigms that use multiple spatially separated visual stimuli with different flickering frequencies, we presented the participants with multiple spatially separated auditory stimuli with different tones and modulation frequencies. In our experiments performed to six healthy volunteers, we were able to discriminate which sound source the participants were selectively attending to with high classification accuracy fairly exceeding the chance level of a binary decision (50%), demonstrating the feasibility of using ASSR modulated by selective attention as one of the promising BCI features.

Our paradigm has several advantages that are suitable for use in practical BCI systems. First, we did not apply any complex preprocessing procedures. In fact, we did not even use basic filtering or artifact rejection processes, which would be advantageous in realizing an efficient real-time BCI system. Second, since the paradigm was simple and intuitive, the participants could easily understand and get accustomed to the target tasks, for which they were only asked to concentrate their attention on either the left or right sound source. Therefore, the proposed paradigm overcomes one of the drawbacks of mental task-based BCI paradigms that require complex and time-consuming training processes. Most importantly, we did not use any visual information during the whole experiment, considering that the main targets of auditory BCI systems would be patients with advanced ALS or CLIS, who have difficulty controlling visual fixation. In the present study, we did not use any types of feedbacks; however, in a practical online ASSR-based BCI system, the participants would also be provided with auditory feedbacks (Nijboer et al., 2008). Then, a totally vision-free BCI system could be realized.

In the conventional visual stimuli-based BCI paradigms, one of the most important advantages of SSVEP-based BCI paradigms over the P300-based BCI paradigms is that the SSVEP-based BCI can generate command signals continuously by monitoring changes in the spectral power density without interruption (Bakardjian et al., 2010), while the P300-based BCI requires time intervals to wait until the next flashing or deviant stimulus is given. We expect that this advantage of SSVEP-based BCI would also be valid in our ASSR-based BCI paradigm, as our present results indirectly demonstrated that the classification accuracy exceeded the chance level even for analysis window sizes as brief as 2 s.

Although the present study discussed the possibility of ASSR in practical BCI applications, the current ASSR-based BCI paradigm still needs to be refined and optimized reflecting the physiological findings on ASSR. Unfortunately, however, the mechanisms of ASSR still stay unclear; especially on the reason why the 40 Hz component is most distinct. One of the hypotheses is that the spectral power change of ASSR at 40 Hz might be caused by temporal coherence of middle latency evoked responses (Stach, 2002). Another possibility is that since 40 Hz may be the most preferred resonant frequency of auditory neural circuits, the neural circuits might produce relatively large oscillations when a 40 Hz auditory stimulus is given (Picton et al., 1987). Yet another hypothesis is that top-down attention processing in general can generate auditory evoked gamma band activity (Debener et al., 2003), but a more recent study reported no significant increment in gamma band activity when a 20 Hz stimulus was provided (Skosnik et al., 2007). On the other hand, earlier studies on ASSR reported no significant relationship between ASSR and attention (Linden et al., 1987), whereas multiple recent studies have suggested that ASSR can be influenced by an individual's attention (Skosnik et al., 2007; Lopez et al., 2009). Further exploration of the mechanisms of ASSR should eventually contribute to the realization of an optimized ASSR-based BCI system.

The research we present here focused on the feasibility of a practical ASSR-based BCI system that might help patients who need an alternative mode of communication. In future studies, we will pursue further improvements on the current paradigm. First, more studies are needed to extend the current paradigm to one suitable for multi-class classification problems. For instance the feasibility of adding more sound streams with different beat frequencies positioned at other distinguishable locations could be investigated. Next, development of an asynchronous or self-paced BCI paradigm would potentially make the current paradigm more practical. All synchronous BCI paradigms need prior input that informs the system of when the participant will start to perform a task. However, synchronous BCI systems cannot take account of the participants' will to perform a specific task. Indeed, such a system usually enforces the participants to perform the task only when an execution cue is given. If a participant is asked to ignore both sound sources when they are unwilling to perform the task and if such an "ignore-all" condition is distinguishable from "attend-to-left" and "attend-to-right" conditions, it is expected that the system might detect the participant's selective attentions asynchronously, which is an exciting prospect that we will explore in future studies. Although the current results of our offline studies have shown relatively low information transfer rate (ITR), of which the maximum value was 1.20 bits/min, we are expecting that the ITR of our system should be enhanced in our future studies since we are trying to develop a new method to enhance the classification accuracy as well as to implement a multi-class BCI paradigm. In the present study, we have demonstrated only a single on-line test experiment; however, our future studies mentioned above would be fully based on online experiments. We are currently planning to apply our system to patients with ALS or CLIS.

Acknowledgements

This work was supported in part by the KRISS-WCL project (Development of Measurement Technology for Cognitive Process) and in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (no. 2010-0015957).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jneumeth.2011.02.007.

References

Bakardjian H, Tanaka T, Cichocki A. Optimization of SSVEP brain responses with application to eight-command brain–computer interface. Neurosci Lett 2010;469:34–8.

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 2009:046002.

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 (9pp).

Debener S, Herrmann CS, Kranczioch C, Gembris D, Engel AK. Top-down attentional processing enhances auditory evoked gamma band activity. Neuroreport 2003;14:683–6.

Decety J, Ingvar DH. Brain structures participating in mental simulation of motor behavior: a neuropsychological interpretation. Acta Psychol (Amst) 1990;73:13–34.

Donchin E, Spencer KM, Wijesinghe R. The mental prosthesis: assessing the speed of a P300-based brain–computer interface. IEEE Trans Neural Syst Rehabil Eng 2000;8(2):174–9.

Engelien A, Schulz M, Ross B, Arolt V, Pantev C. A combined functional in vivo measure for primary and secondary auditory cortices. Hear Res 2000;148:153–60.

Farwell LA, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalogr Clin Neurophysiol 1988;70:510–23.

Halder S, Rea M, Andreoni R, Nijboer F, Hammer EM, Kleih SC, Birbaumer N, Kübler A. An auditory oddball brain-computer interface for binary choices. Clinical Neurophysiology 2010;121:516–23.

Hill NJ, Lal TN, Bierig K, Birbaumer N, Schölkopf B. An auditory paradigm for brain–computer interfaces. In: Adv neural inf process syst. Cambridge: MIT Press; 2005. pp. 569–576.

Hoffmann U, Vesin JM, Ebrahimi T, Diserens K. An efficient P300-based brain–computer interface for disabled subjects. J Neurosci Methods 2008;167:115–25.

Hwang HJ, Kwon K, Im CH. Neurofeedback-based motor imagery training for brain–computer interface (BCI). J Neurosci Methods 2009;179:150–6.

Jeannerod M, Frak V. Mental imaging of motor activity in humans. Curr Opin Neurobiol 1999;9:735–9.

Kanoh S, Miyamoto K, Yoshinobu T. A brain–computer interface (BCI) system based on auditory stream segregation. Conf Proc IEEE Eng Med Biol Soc 2008:642–5.

Keirn ZA, Aunon JI. Man-machine communications through brain-wave processing. IEEE Eng Med Biol Mag 1990;9:55–7.

Klobassa DS, Vaughan TM, Brunner P, Schwartz NE, Wolpaw JR, Neuper C, Sellers EW. Toward a high-throughput auditory P300-based brain–computer interface. Clin Neurophysiol 2009;120:1252–61.

Krusienski DJ, Sellers EW, McFarland DJ, Vaughan TM, Wolpaw JR. Toward enhanced P300 speller performance. J Neurosci Methods 2008;167:15–21.

Lalor EC, Kelly SP, Finucane C, Burke R, Smith R, Reilly RB, McDarby G. Steady-state VEP-based brain–computer interface control in an immersive 3D gaming environment. Eurasip J Appl Sig P 2005;2005:3156–64.

Lin ZL, Zhang CS, Wu W, Gao XR. Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. IEEE Trans Biomed Eng 2006;53:2610–4.

Linden RD, Picton TW, Hamel G, Campbell KB. Human auditory steady-state evoked potentials during selective attention. Electroencephalogr Clin Neurophysiol 1987;66:145–59.

Lopez MA, Pomares H, Pelayo F, Urquiza J, Perez J. Evidences of cognitive effects over auditory steady-state responses by means of artificial neural networks and its use in brain–computer interfaces. Neurocomputing 2009;72:3617–23.

Luo A, Sullivan TJ. A user-friendly SSVEP-based brain–computer interface using a time-domain classifier. J Neural Eng 2010:026010.

Middendorf M, McMillan G, Calhoun G, Jones KS. Brain–computer interfaces based on the steady-state visual-evoked response. IEEE Trans Rehab Eng 2000;8:211–4.

Nijboer F, Furdea A, Gunst I, Mellinger J, McFarland DJ, Birbaumer N, Kubler A. An auditory brain–computer interface (BCI). J Neurosci Methods 2008;167:43–50.

Pastor MA, Artieda J, Arbizu J, Marti-Climent JM, Penuelas I, Masdeu JC. Activation of human cerebral and cerebellar cortex by auditory stimulation at 40 Hz. J Neurosci 2002;22:10501–6.

Penny WD, Roberts SJ, Curran EA, Stokes MJ. EEG-based communication: a pattern recognition approach. IEEE Trans Rehab Eng 2000;8:214–5.

Pfurtscheller G, Neuper C. Motor imagery activates primary sensorimotor area in humans. Neurosci Lett 1997;239:65–8.

Picton TW, John MS, Dimitrijevic A, Purcell D. Human auditory steady-state responses. Int J Audiol 2003;42:177–219.

Picton TW, Skinner CR, Champagne SC, Kellett AJC, Maiste AC. Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. J Acoust Soc Am 1987;82:165–78.

Ross B, Borgmann C, Draganova R, Roberts LE, Pantev C. A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. J Acoust Soc Am 2000;108:679–91.

Ross B, Draganova R, Picton TW, Pantev C. Frequency specificity of 40-Hz auditory steady-state responses. Hear Res 2003;186:57–68.

Ross B, Picton TW, Pantev C. Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. Hear Res 2002;165:68–84.

Schreuder M, Blankertz B, Tangermann M. A new auditory multi-class brain–computer interface paradigm: spatial hearing as an informative cue. Plos One 2010;5.

Sellers EW, Donchin E. A P300-based brain–computer interface: initial tests by ALS patients. Clin Neurophysiol 2006;117:538–48.

Skosnik PD, Krishnan GP, O'Donnell BF. The effect of selective attention on the gamma-band auditory steady-state response. Neurosci Lett 2007;420: 223–8.

Stach BA. The auditory steady-state response: a primer. Hear J 2002;55:10–8.

Wang T, Deng J, He B. Classifying EEG-based motor imagery tasks by means of time-frequency synthesized spatial patterns. Clin Neurophysiol 2004;115:2744–53.

Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain–computer interfaces for communication and control. Clin Neurophysiol 2002;113:767–91.