Computers in Biology and Medicine 84 (2017) 45–52

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Music and natural sounds in an auditory steady-state response based brain–computer interface to increase user acceptance

Jeong Heo[a], Hyun Jae Baek[b], Seunghyeok Hong[a], Min Hye Chang[c], Jeong Su Lee[b], Kwang Suk Park[d,*]

[a] *Interdisciplinary Program of Bioengineering, Seoul National University, Seoul, Republic of Korea*
[b] *Mobile Communication Business, Samsung Electronics Co., Ltd., Suwon, Republic of Korea*
[c] *Advanced Medical Device Research Division, Korea Electro-Technology Research Institute, Ansan, Republic of Korea*
[d] *Department of Biomedical Engineering, College of Medicine, Seoul National University, Seoul, Republic of Korea*

ARTICLE INFO

Keywords:
Brain–computer interface (BCI)
Auditory steady-state response (ASSR)
Auditory stimulation
Music
Natural sounds
Ergonomics

ABSTRACT

Patients with total locked-in syndrome are conscious; however, they cannot express themselves because most of their voluntary muscles are paralyzed, and many of these patients have lost their eyesight. To improve the quality of life of these patients, there is an increasing need for communication-supporting technologies that leverage the remaining senses of the patient along with physiological signals. The auditory steady-state response (ASSR) is an electro-physiologic response to auditory stimulation that is amplitude-modulated by a specific frequency. By leveraging the phenomenon whereby ASSR is modulated by mind concentration, a brain–computer interface paradigm was proposed to classify the selective attention of the patient. In this paper, we propose an auditory stimulation method to minimize auditory stress by replacing the monotone carrier with familiar music and natural sounds for an ergonomic system. Piano and violin instrumentals were employed in the music sessions; the sounds of water streaming and cicadas singing were used in the natural sound sessions. Six healthy subjects participated in the experiment. Electroencephalograms were recorded using four electrodes (Cz, Oz, T7 and T8). Seven sessions were performed using different stimuli. The spectral power at 38 and 42 Hz and their ratio for each electrode were extracted as features. Linear discriminant analysis was utilized to classify the selections for each subject. In offline analysis, the average classification accuracies with a modulation index of 1.0 were 89.67% and 87.67% using music and natural sounds, respectively. In online experiments, the average classification accuracies were 88.3% and 80.0% using music and natural sounds, respectively. Using the proposed method, we obtained significantly higher user-acceptance scores, while maintaining a high average classification accuracy.

1. Introduction

A brain–computer interface (BCI) is a technology that controls computers and other devices using electrical brain activities [38]. Currently, several research groups are studying BCIs as new communication tools. Since the development of BCIs, various applications, including artificial prosthetics [6], mind spellers [8], and mental games [18], have been introduced. Several modalities exist for measuring brain activity, including the electroencephalogram (EEG), electrocorticogram, single unit recording, magnetoencephalography, functional magnetic resonance imaging, and near infrared spectroscopy. Although EEGs have a low spatial resolution, EEG-based BCIs are extensively studied because they do not require surgery and the devices used to obtain EEGs are portable and cost effective. Various paradigms involving different EEG responses have been studied. These include event-related potential (ERP), steady-state evoked potential, and motor imagery.

Locked-in syndrome (LIS) is a condition in which the patient is awake and conscious; however, almost all of the patient's voluntary muscles are paralyzed [3]. LIS is divided into three categories [36]: classic LIS, incomplete LIS, and total LIS. In classic LIS, the patient experiences complete paralysis, except for vertical eye movements and blinking. In incomplete LIS, the patient partially recovers muscle control and eye movements. In total LIS, the patient is completely paralyzed and cannot move their eyes; however, they are completely conscious. Various communication-supporting devices have been developed by rehabilitation engineers. For patients who can control their eye movements, supporting systems, such as eye trackers [5] or visual

* Corresponding author.
E-mail address: pks@bmsil.snu.ac.kr (K.S. Park).

http://dx.doi.org/10.1016/j.compbiomed.2017.03.011
Received 15 March 2017; Accepted 15 March 2017

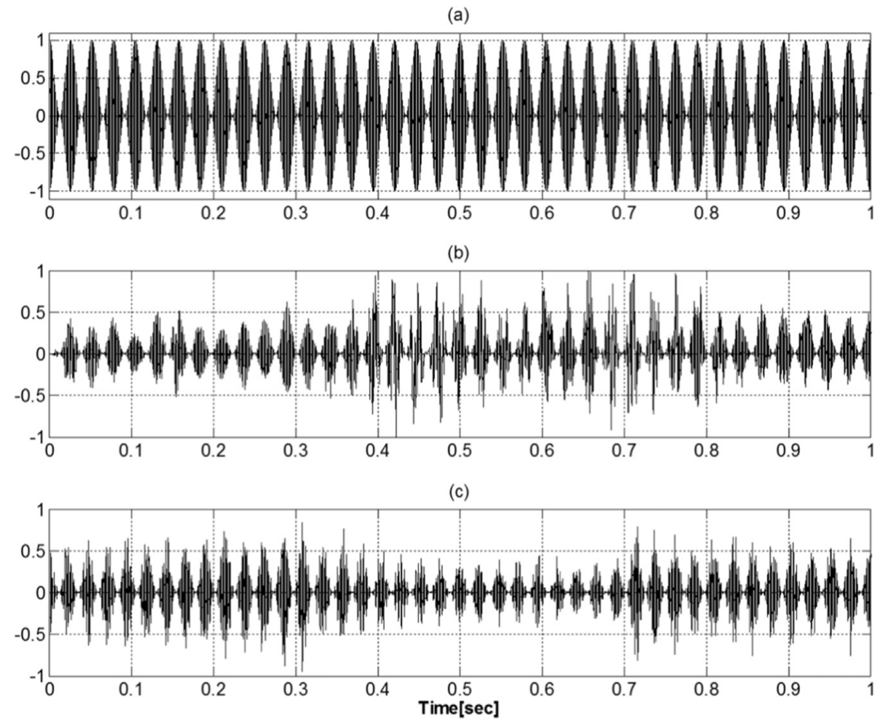

Fig. 1. Experimental auditory stimuli (modulation index=1.0) (a) Stimulus: 1 kHz monotone carrier (b) Stimulus: music (violin instrumental) carrier (c) Stimulus: natural sound (cicada singing) carrier.

stimulation-based BCIs, which can utilize eye movements, are suitable. However, total LIS patients require communication devices that utilize senses, such as hearing or touch.

For total LIS patients, a BCI that employs a motor-imagery, vibrotactile, or auditory paradigm can be an option for supporting communication. In a motor-imagery-based BCI, the user controls the device using the synchronization/desynchronization of the EEG that occurs in the sensory motor cortex when imaging a moving part of the body [28]. A vibrotactile P300-based BCI uses ERP elicited by a 'deviant' spot or pulse among 'standard' spots or pulses [23]. Auditory ERP-based and auditory steady-state response (ASSR)-based BCIs are representative auditory paradigms. ASSR is a response of the brain to periodic amplitude modulation of a continuous tone. An auditory ERP-based BCI is based on EEG responses elicited by a 'deviant' sound among 'standard' sounds [10,11]. Hill et al. [11] used two concurrent auditory stimulus sequences on both sides of the subject. The subject concentrated on the deviant sound on one side to make a binary decision.

In ASSR, when an auditory stimulus is applied, an increase in the power of the frequency components that are harmonics of the modulation frequency is detected. In clinics, ASSR has been used to predict hearing sensitivity and to create an estimated audiogram [4]. Evidence of ASSR modulation by cognitive effects was reported by Lopez et al. [22]. Based on this phenomenon, an ASSR-based BCI paradigm that uses auditory selective attention was proposed by Kim et al. [17]. They used a monotone sound with frequencies of 2.5 kHz and 1 kHz as a carrier for auditory stimulation.

We hypothesize that stimulation sounds that are more familiar to a user can enhance the performance of an ASSR-based BCI. We replace the monotone carrier used for auditory stimulation with other sounds that are more comforting for a user, such as music and natural sounds. The aim of this study is to propose an ergonomic BCI paradigm that can be comfortably experienced by the patient.

2. Methods

2.1. Stimuli

In previous studies, the auditory stimuli inducing an ASSR were periodic waves that were amplitude-modulated by message waves [4]. If *c(t)* denotes the carrier wave and *m(t)* denotes the message wave, then the amplitude modulation is defined as:

$$y(t) = [1 + m(t)] \times c(t) \tag{1}$$

Six different carriers were used in this study. Two sinusoidal waves with frequencies of 1 kHz and 2.5 kHz were utilized as monotone carriers. Violin and piano instrumentals were used as music carriers. The sounds of water flowing and cicadas singing were employed as natural sound carriers. The violin instrumental was extracted from "B Rossette," and the piano instrumental was extracted from "Rainbow Bridge." Both were downloaded from web-based music-streaming services.

Thirty-eight Hz and 42 Hz were selected as message frequencies. Previous studies reported that ASSR has large amplitudes for message frequencies of approximately 40 Hz [29,31]. The 1-kHz monotone wave, the violin instrumental, and the sound of water flowing were amplitude-modulated with a message frequency of 38 Hz. The 2.5-kHz monotone wave, the piano instrumental, and the sound of cicadas singing were amplitude-modulated with a message frequency of 42 Hz.

The AM modulation index is a measure of amplitude variation. This index represents variation in modulation around the original amplitude. It is defined as:

$$h = \frac{\text{peak value of } m(t)}{\text{peak value of } c(t)} \tag{2}$$

Under music and natural sound carrier conditions, three sessions were performed with a different modulation index for each condition.