Verification of 4-class classification accuracy based on auditory steady state response for Brain-Computer Interface

Kazusa Shintani
Graduate School of Science and Engineering for Education
University of Toyama, Toyama, Japan
Email: m1671112@ems.u-toyama.ac.jp

Tadanobu Misawa †
Graduate School of Science and Engineering
University of Toyama, Toyama, Japan
Tel: (+81) 76-445-6753, Email: misawa@eng.u-toyama.ac.jp

Shigeki Hirobayashi
Graduate School of Science and Engineering
University of Toyama, Toyama, Japan
Tel: (+81) 76-445-6889, Email: hirobays@eng.u-toyama.ac.jp

Abstract. Brain-computer interface (BCI) research has grown considerably in recent years. In particular, BCIs based on the principle of auditory steady-state response (ASSR) have been attracting attention. However, most ASSR-based BCIs are insufficiently versatile because they use only two class of classification. In this study, we investigate four-class classification accuracy in order to develop an improved ASSR-based BCI. In our experiment, four types of machine voice with different modulation frequencies (32, 36, 40, and 44 Hz) were adjusted so as to be heard from different directions (front, back, right, and left) and then played simultaneously through earphones. Five subjects were each instructed to concentrate on a particular voice determined randomly by on-screen instructions displayed during the interstimulus intervals. Features were extracted by analyzing the resulting electroencephalography. Learning and classification were then carried out using a support vector machine. Our experimental results show an overall average classification accuracy of ~40%, which is above the 25% (one in four) statistical baseline. This result suggests the possibility of a four-class ASSR-based BCI. However, the current classification accuracy is too low for actual use, and further improvement is deemed necessary.

Keywords: Brain-Computer interface (BCI), Auditory steady-state response (ASSR), Electroencephalography (EEG)

1. INTRODUCTION

Research on brain-computer interfaces (BCIs) has grown significantly in recent years (Wolpaw *et al.* 2002). A BCI is a control system that works without using body movements but rather by measuring changes in brain activity using near-infrared spectroscopy (NIRS), electroencephalography (EEG), and functional magnetic resonance imaging (fMRI). BCIs are particularly suited as communication tools for people who are incapacitated by the likes of amyotrophic lateral sclerosis (ALS) or cerebral palsy. ALS (also known as Lou Gehrig's disease) is a motor neuron disease that causes muscular atrophy and the eventual loss of voluntary movement control, although sensory perception is usually unaffected. As such, a BCI would be effective given that only brain activity, not bodily motion, is required to operate it (Wolpaw *et al.* 2002). In most BCI systems that have been studied, non-invasive EEG signals are used as the primary input. Such a system is categorized typical as being based on one of four possible command types: motor imagery (MI), P300 brain waves, steady-state visually evoked potential (SSVEP), or auditory steady-state response (ASSR).

At present, much of the technology for BCIs is based on MI, in which mental simulations of a physical action cause changes in brain activity. Often in MI-based BCI, the hands, feet, and tongue are used mentally to initiate commands. During MI, the amplitudes of the α and β bands of brain waves obtained via electrodes placed over the motor cortex are found to decrease; this is known as event-related desynchronization (ERD). The amplitudes return to their original levels once the subject ceases MI; this is known as event-related synchronization (ERS). ERD and ERS are the bases for an MI-based BCI, which has the advantages of relatively high classification accuracy and the fact that external stimuli are unnecessary. However, this form of BCI is not suitable for those people for whom actual physical motion is impossible and hence its visualization (even with MI training) is extremely difficult (Curran *et al*. 2004). Hence, other forms of BCI have been developed that do not rely on MI, in particular P300, SSVEP, and ASSR.

The P300 wave is a component of an event-related potential (ERP) that represents brain activity that is time locked to the eliciting event. P300 arises most frequently in association with the so-called oddball paradigm (Nijboer *et al*. 2008). However, a consequence of this is that command input based on P300 can suffer from a considerable time delay.

SSVEPs are brain responses to repetitive visual stimuli. SSVEP-based BCI involves the use of black-and-white checkerboard patterns and flashing light-emitting diodes (Hwang *et al*. 2012). Although this form of BCI has the advantages of relatively high information transfer rates and the fact that it does not require any initial training, there is a serious possibility of it triggering epileptic seizures in certain subjects (Regan *et al*. 1977).

Given these somewhat intractable disadvantages of P300- and SSVEP-based BCIs, attention in recent years has turned instead to systems based on ASSR. This is a brain-wave response to repetitive auditory stimuli. ASSR-based BCI involves two forms of auditory stimulation—single tone and machine voice—based on different modulation frequencies. This form of BCI has the advantage that neither initial training nor a visual display is required. However, it lacks versatility at present because it involves only two class of classification (Kim *et al*. 2011). In the present study, we investigate four-class classification accuracy in order to develop an improved ASSR-based BCI.

2. Method

2.1 Subjects

Five healthy male volunteers were recruited from students of the Graduate School of Science and Engineering for Education at the University of Toyama. Before the experiment, each subject was given a detailed written summary of the experimental procedures, and agreed to participate voluntarily by signing a subject consent form.

2.2 Auditory Stimuli

ASSR is an electrical brain response that is evoked upon hearing periodic amplitude-modulated sinusoidal tones, sound clicks, or a machine voice (TANAKA *et al*. 2013; Nakamura *et al*. 2013). It generally shows increased amplitude around the modulation frequency of the sound stream. The optimal modulation frequency has been reported as being 30–50 Hz, with the maximum amplitude response at ~40 Hz (Reyes et al. 2004). Therefore, we chose the four modulation frequencies of 32, 36, 40, and 44 Hz. The associated modulated auditory stimuli involved a Japanese machine voice saying ‘mae’ (32 Hz), ‘ushiro’ (44 Hz), ‘migi’ (40 Hz), and ‘hidari’ (36 Hz) so that the subjects could easily distinguish each sound. These stimuli were adjusted to be heard from different directions (front, back, right, and left) by combining them with a head-related transfer function. They were generated using SofTalk (Japanese software) and MATLAB at a sampling rate of 8 kHz, and were saved in the waveform (.wav) audio file format. The duration of each auditory stimulus and trial was 10 s.

2.3 Experimental task

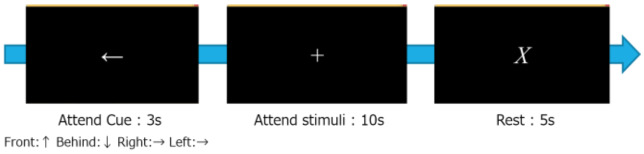

Figure. 1: Experimental task used in the study.

An individual subject sat on a comfortable armchair in a dimly lit room in front of a monitor while wearing earphones. The experimental task is shown schematically in Figure. 1. Firstly, the subject was presented with a 3-s visual cue to alert them to which directional voice they should listen for; this instruction was chosen randomly. After this initial cue, the four auditory stimuli were played simultaneously for 10 s through the earphones while the subject maintained visual attention on a cross (+) in the center of the monitor. This ended with a stationary ‘X’ being displayed in the center of the monitor for 5 s, during which time the subject could rest. The procedure was repeated for a total of 25 trials

2.4 EEG recording

Electrodes were attached to the subject’s scalp according to the International 10–20 system. EEG signals were acquired at four electrodes (Cz, Fz, T3, and T4) using a four-channel EEG acquisition system (ProComp infiniti, Thought Technology, USA). T3 and T4 are associated with the auditory cortical area. The choice of the other two electrodes was based on previous studies: Cz has been used in ASSR-based BCI research (Kim *et al.* 2011) and Fz has used in basic ASSR research (Griskova-Bulanova *et al.* 2013). The ground electrode was attached to the right earlobe, with the reference electrode on the left one. The EEG sampling rate was 256 Hz in each experiment.

2.5 Feature extraction

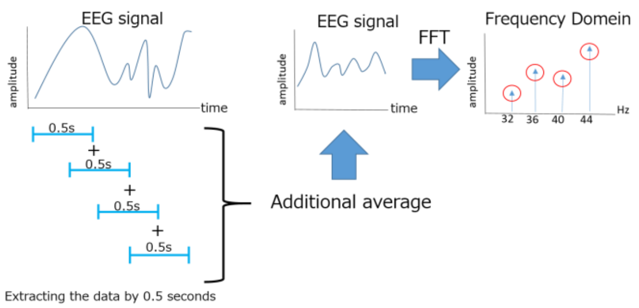

Figure. 2: the method of extracting features4

In this study, we used four auditory stimuli, each modulated by a different frequency. We measured the EEG signals, including ASSR, of each pattern. From the resulting EEG data, we extracted four features to verify the classification accuracy. We used a total of 4×25=100 trials to classify the EEG data sets for selective attention to stimuli from the front (25), back (25), right (25), and left (25). The frequency spectrums of each window length were calculated using the fast Fourier transform (FFT). The four extracted features are listed below.

Feature 1: amplitude of the corresponding frequency at each modulation frequency.

Feature 2: amplitude ratio of the corresponding frequency in each modulation frequency.

Feature 3: amplitude of the corresponding frequency and harmonic at each modulation frequency.

Feature 4: amplitude of the frequency of the EEG data obtained by synchronous addition at each modulation frequency.

Feature 4 involves additional averaged EEG data divided in the time domain using a 128-point analysis window and a 50% overlap. These EEG data are then added synchronously, as shown in Figure. 2.

The modulation frequencies used in this study are all integrally divisible by four so that a 64-point periodic analysis window arises in EEG data sampled at 256 Hz. This is not the case for, say, a 38-Hz sine wave, which therefore cannot be added synchronously. Consequently, we set the modulation frequencies as 32, 36, 40, and 44 Hz.

We measured auditory-stimulus EEG data at 256 Hz for 10 s for a single trial, thus creating data sets of 2,560 points each. Classification accuracy was determined using four different analysis windows. We compared the classification accuracy and analysis window length of the relationship using EEG measurements 1 s (256 points), 2 s (512 points), 4 s (1,024 points), and 8 s later (2,048 points).

In addition, we determined four types of classification accuracy by measuring the EEG signals from different electrode combinations: Cz only, Fz only, Cz and Fz, and all electrodes. We could not measure clear ASSR signals from T3 and T4, and hence these were excluded from use in isolation.

Using the above features, we determined classification accuracies using a 4-class nonlinear support vector machine (SVM) and leave-one-out cross validation. The SVM used the one-to-one approach and a Gaussian kernel.

3. Result

Table 1: Fz classification accuracy (%)

Time (s)	1	2	4	8
Feature1	35.3	40.4	35.3	42.6
Feature2	40.5	37.5	37.2	45.6
Feature3	41	37.5	39.5	44.2
Feature4	41.1	39.2	42	45

Table 2: Cz classification accuracy (%)

Time (s)	1	2	4	8
Feature1	39.5	39.6	42.2	43.8
Feature2	38.8	39	41.4	43.9
Feature3	39.2	37.5	37.4	41.8
Feature4	38.7	40.7	37.6	39.1

Table 3: Fz+Cz classification accuracy (%)

Time (s)	1	2	4	8
Feature1	42.1	39.1	38.9	44.3
Feature2	45.1	38.7	39	43.2
Feature3	44.2	36.9	38.5	42.1
Feature4	38.6	40.1	41.2	41.8

Table 4: All-channel classification accuracy (%)

Time (s)	1	2	4	8
Feature1	37.6	36.3	37.7	43.3
Feature2	38.2	39.6	39.5	43.4
Feature3	39.5	40.5	38.2	45.4
Feature4	37.8	37.6	41.6	46

The values of classification accuracy averaged over all subjects are listed in Tables 1–4. These show values that are consistently above that of the statistical-chance baseline (25%). The highest classification accuracy was obtained in all but three cases with a window length of 8 s. The average over all window lengths for each feature was joint highest for Features 3 and 4 (41.8%) when using the Fz electrode.

4. Discussion

We determined the classification accuracy of ASSR using a 4-class SVM. Given that the results shown in Tables 1–4 are consistently >25%, we consider that they show the possibility of 4-class classification in an ASSR-based BCI.

The highest classification accuracy was obtained in general when using an 8 s window length. However, this result is not particularly good in the context of an actual BCI because more time would be required to enter commands. Thus, it is considered necessary to make the tasks shorter in order to facilitate command input.

The classification accuracies reported here are lower in general than those of previous research. In this study, because we played four types of sound simultaneously during the input task, it may well have been difficult for subjects to concentrate effectively. This may have had a detrimental impact on the classification accuracy. Therefore, we consider it necessary to improve the classification accuracy by devising a better sound-presentation method for the input task. We intend to play each sound separately for a shorter time while requiring subjects to concentrate on individual sounds. In addition, we hope to improve the classification accuracy by moving each sound by several seconds.

5. Conclusion

Until now, the focus on ASSR-based BCIs has been on those with only two classes of classification. In this study, we addressed this by developing a more versatile system aimed at verifying 4-class classification accuracy. In our experiment, four machine voices with different modulation frequencies (32, 36, 40, and 44 Hz) were adjusted so as to be heard from different directions (front, back, right, and left), and then played simultaneously through earphones. Five subjects were instructed to concentrate on a single voice determined randomly by on-screen instructions shown during the interstimulus intervals. Features were extracted by analyzing the resulting EEG signals using an SVM to perform learning and classification. Our experimental results showed an overall average classification accuracy (~40%) above the statistical baseline (25%). We consider this result as showing the possibility of realizing a 4-class ASSR-based BCI. However, the classification accuracy of the present system would have to be improved before it could be used in a real environment. In future work, we intend to isolate the individual voices and play them separately for shorter times.

Reference

Dean J Krusienski, Eric W Sellers, Fran_cois Cabestaing, Sabri Bayoudh, Dennis J Mc-Farland, Theresa M Vaughan, and Jonathan R Wolpaw. A comparison of classi_cation techniques for the p300 speller. Journal of neural engineering, Vol. 3, No. 4, p. 299, 2006.

David Regan. Steady-state evoked potentials. JOSA, Vol. 67, No. 11, pp. 1475, 1489, 1977.

Do-Won Kim, Han-Jeong Hwang, Jeong-Hwan Lim, Yong-Ho Lee, Ki-Young Jung, and Chang-Hwan Im. Classification of selective attention to auditory stimuli: toward vision-free brain-computer interfacing. Journal of neuroscience methods, Vol. 197, No. 1, pp.180-185, 2011.

Eleanor Curran, Peter Sykacek, Maria Stokes, Stephen J Roberts, Will Penny, Ingrid Johnsrude, and Adrian M Owen. Cognitive tasks for driving a brain-computer inter-facing system: a pilot study. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, Vol. 12, No. 1, pp. 48, 54, 2004.

F Nijboer, EW Sellers, J Mellinger, MA Jordan, T Matuz, A Furdea, S Halder, U Mochty, DJ Krusienski, TM Vaughan, et al. A p300-based brain-computer interface for people with amyotrophic lateral sclerosis. Clinical neurophysiology, Vol. 119, No. 8, pp. 1909, 1916, 2008.

Han-Jeong Hwang, Jeong-Hwan Lim, Young-Jin Jung, Han Choi, Sang Woo Lee, and Chang-Hwan Im. Development of an ssvep-based bci spelling system adopting a qwerty-style led keyboard. Journal of neuroscience methods, Vol. 208, No. 1, pp. 59-65, 2012.

Inga Griskova-Bulanova, Kastytis Dapsys, Valentinas Maciulis, and Sidse M Arnfred. Closed eyes condition increases auditory brain responses in schizophrenia. Psychiatry Research: Neuroimaging, Vol. 211, No. 2, pp. 183-185, 2013.

Keita TANAKA, Shinya KURIKI, Iku NEMOTO, and Yoshinori UCHIKAWA. Auditory steady-state responses in magnetoencephalogram and electroencephalogram: Phenom-ena, mechanisms, and applications. Advanced Biomedical Engineering, Vol. 2, No. 0, pp.55-62, 2013.

Niels Birbaumer, Ander Ramos Murguialday, and Leonardo Cohen. Brain-computer interface in paralysis. Current opinion in neurology, Vol. 21, No. 6, pp. 634, 638, 2008.

Samuel A Reyes, Richard J Salvi, Robert F Burkard, Mary Lou Coad, David S Wack,Paul J Galantowicz, and Alan H Lockwood. Pet imaging of the 40 hz auditory steady state response. Hearing research, Vol. 194, No. 1, pp. 73-80, 2004.

T Nakamura, H Namba, and T Matsumoto. Classi_cation of auditory steady-stateresponses to speech data. In Neural Engineering (NER), 2013 6th InternationalIEEE/EMBS Conference on, pp. 1025-1028. IEEE, 2013.

Wolpaw, Jonathan R and Birbaumer, Niels and McFarland, Dennis J and Pfurtscheller, Gert and Vaughan, Theresa M,``Brain-computer interfaces for communication and control'', Clinical neurophysiology, Vol.113, No.6, pp.767-791 2002.