33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Estimation of Direction of Attention Using EEG and Out-of-head Sound Localization

Masashi Ebisawa, Masumi Kogure, Sho-hei Yano, Syu-ichi Matsuzaki, Yasuhiro Wada

Abstract—Brain-Machine Interfaces (BMIs) are being researched controlling external devices such as robots and computers by measuring the cranial nerve activity of the operator. The brain activities evoked by visual stimuli have been studied intensively. However, few studies have considered a BMI that uses the brain activities evoked by auditory stimuli. This study investigated whether a person's direction of attention can be estimated using an event-related potential (ERP) generated by selective attention to an auditory stimulus. An auditory stimulus and an out-of-head sound localization system that can create an audio image outside the head that is presented through an earphone were used instead of a loudspeaker system. This system was experimentally evaluated by presenting the subject auditory cues from one of six directions while the subject directed his attention in one direction. An EEG response similar to an ERP was observed. The direction of attention was estimated using support vector machine with an accuracy of 89.2[%] on average for the three subjects. This suggests that a BMI system based on the estimated direction of attention can be developed by using out-of-head sound localization.

I. INTRODUCTION

The use of a Brain-Machine Interface (BMI) is a promising approach to controlling a robot because it uses the operator's cerebral nerve activity [1],[2], not voluntary muscle activity. It is thus particularly attractive for people with a serious movement disorder such as people with amyotrophic lateral sclerosis or spinal cord damage. Its application to rehabilitation medicine as well as game software is expected.

There are invasive and non-invasive methods for measuring neural activity. A commonly used non-invasive method is electroencephalography (EEG). It uses the event-related potential (ERP), which reflects the electric and physiological reactions to internal and external stimulations. Potential P300 is especially useful because it is evoked by various types of stimuli (visual, auditory, etc.) with a latency of 250~500[ms]. P300 speller paradigm [3],[4] is a typical BMI using P300 latency. The P300 speller system presents a visual stimulus, such as a flashing character, and then predicts the character on which the user is focusing on the basis of the measured P300 latency.

The EEG potentials evoked by visual stimuli are often measured as the cerebral nerve activity for use in a BMI so that it is represented by the P300 speller. While there have been many studies on the use of a visual stimulus for a BMI, few studies have considered a BMI based on the brain activities associated with an auditory stimulus. In those studies, there was a problem obtaining the space required for the measurement. Moreover, there was a problem with synthesizing auditory stimuli outside the head. In the study reported here, we used an out-of-head sound localization system [5] that enables auditory stimuli to be produced outside the head. Stimuli are presented through earphones, which are similar to the way actual sounds are presented. We experimentally assessed the performance of a BMI system in which the direction of the sound source on which the user directed his or her attention was estimated (Fig. 1). Experimental evaluation of this system indicates that a user's direction of attention can be estimated from the EEG response.

Wada lab, Nagaoka University of Technology, 1603-1 Kamitomioka, Nagaoka, Niigata, 940-2188 Japan (e-mail: ywada@nagaokaut.ac.jp or s105013@stn.nagaokaut.ac.jp)

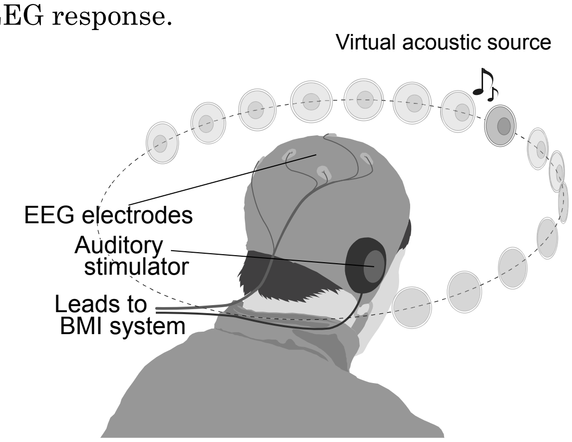

Fig. 1 BMI using out-of-head sound localization

II. METHOD

A. Subjects

Three healthy men, ages 22–24, participated in

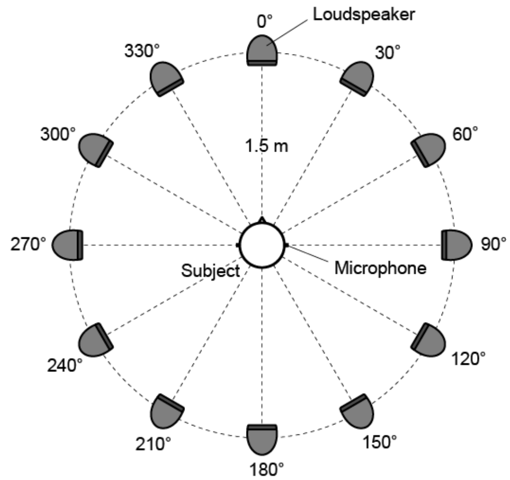

Fig. 2 Experimental setup for measuring transfer function for each subject

this study, which was approved by the Nagaoka University of Technology ethics committee.

B. Auditory stimuli

The out-of-head sound localization was created by presenting an auditory stimulus so that the subject eardrums were stimulated in a way similar to that by the sound from a loudspeaker. Since the shapes of the head and external auditory canal differed among the subjects, the transfer function for each subject was measured [6] using the setup shown in Fig. 2. Loudspeakers were arranged in a circle 1.5m from the subject at 30° intervals (12 speakers in total). A small microphone was attached next to each external ear canal entrance. The sound image positions corresponded to those of the loudspeakers.

A sound signal in which white noise was convolved with the measured transfer function was synthesized for each subject. The sound pressure level (in decibels) was regularized among the subjects by multiplying the coefficient (the acoustic pressure level for the left ear for direction 0° was adjusted to –25[dB]). The volume of the earphones was adjusted so that the acoustic pressure level for the left ear was 65[dB]. This was done using a head and torso simulator with a microphone attached at the eardrum position.

C. Equipment

We measured the EEG signals using a digital electroencephalograph (Biosemi ActiveTwo AD-box ADC-12) with 64 electrodes attached to the subject's scalp using a cap. The electrodes were placed in accordance with the international 10-20 system, and a reference electrode was attached to each earlobes.

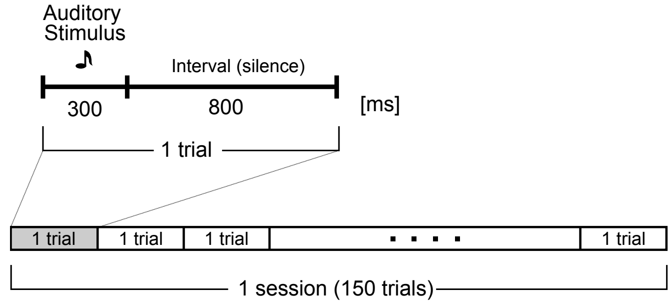

Fig. 3 Protocol for system assessment

D. Experiment task

The protocol experiment ally assesses the performance of our BMI system is shown in Fig. 3. Each trial consisted of a 300[ms] stimulus and an 800[ms] silence interval. The simulated sound source was in one of six directions: 30, –30, 90, –90, 150, –150° . A separate auditory cue was presented in a random sequence while the subject focused on one of the six directions. The subject counted the number of times this target cue was presented. It was presented for about 20% of the stimuli. To avoid the effects of visual sensation and eye blink, the subject was instructed to perform the counting task with eyes closed.

Ⅲ. RESULTS

To remove artifacts and noise, which are unrelated to brain activity, we pre-processed the raw data using third-order Butterworth band pass filters of 1~7[Hz]. The average for –100[ms] to 0[s] in each trial was set as the baseline.

Fig. 4 shows the average EEG wave forms for subject S1 measured at electrodes Fz, Cz, and Pz. The solid lines represent the target trials, and the dashed lines represent the non-target trials.

The EEG signals of S1 were the clearest and most stable among the three subjects. Prominent responses were observed in the target trials. We observed positive responses with about 300[ms] latency, indicating that the response differed between the target and non-target trials and that the direction of the auditory cue on which the subject is focusing can be estimated from EEG signals.

Ⅳ. ESTIMATION OF SOUND SOURCE DIRECTION

A. Support Vector Machine

Support vector machine (SVM) was used for

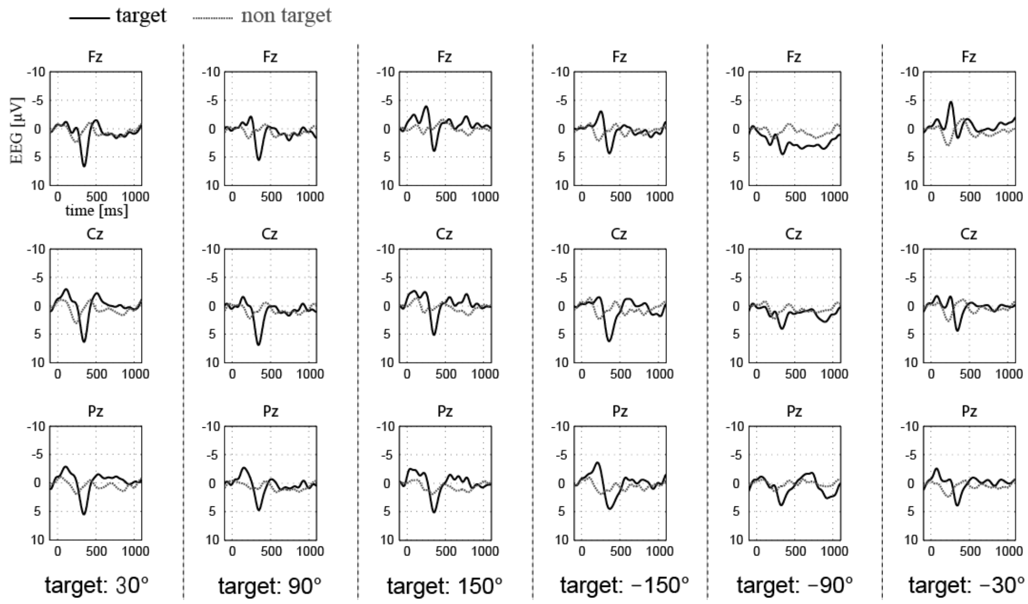

Fig. 4 Average EEG waveforms for subject S1 (six-direction discrimination task)

classification for estimating the target direction. SVM has been applied to binary distinction problems. The discrimination function of SVM is

$$f(x) = sgn\left(\sum_{i=1}^{l} y_i \alpha_i^* K(x, x_i) + b^*\right) \qquad (1)$$

where $x_i (i=1,\ldots,l)$ is a learning sample, $y_i (i=1,\ldots,l)$ is a teaching signal, and $\alpha^*_i (i=1,\ldots,l)$ is the optimum solution of the quadratic programming problem defined as

$$\max_{\alpha} \quad \sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$subject\ to \quad 0 \leqq \alpha_i \leqq C, i = 1, \ldots, l \qquad (2)$$

$$\sum_{i=1}^{l} \alpha_i y_i = 0$$

Linear kernel function K is defined as

$$K(x_i, x_j) = x_1^T x_2 \qquad (3)$$

Penalty parameter C of expression (2) is a positive integral constant. Threshold b^* is given by

$$b^* = \frac{1}{|I|}\sum_{i \in I}\left(y_i - \sum_{j=1}^{l} y_j \alpha_j^* K(x_i, x_j)\right) \qquad (4)$$

where I contains $0<\alpha^*_i<C$ in the support vectors.

B. Classification

SVM was used to classify the EEG waveforms into target and non-target trials. The feature vector used for SVM constituted the time series data from the 64 electrodes. We cut the data for 1100[ms] from the beginning of the auditory cue presentation for each trial. The 270 samples for each electrode were reduced to 27 by using the average for every ten samples. Therefore, the dimension of the feature vector was 1728 (27×64). Next, we normalized the data so that the maximum absolute value for each electrode was "1", which made the amplitude of each electrode constant. The results for which the amplitude exceeded 40[μV], which may have includes significant noise, were excluded from discrimination.

The SVM improved the S/N ratio of the EEG signals and thus facilitated the classification. We calculated the average EEG signal for 2–10 data samples and the EEG signal for a single trial. The signals were averaged using the same number of samples as for a single trial. Samples were repeatedly selected in a random sequence.

Discrimination of the target and non-target trials required estimation of the parameters α^* and b^* after setting an appropriate value of C. Thus, we separated the measurement samples into two categories: learning samples and evaluation samples. Each subject's data consisted of 368 samples for the target trials and 1432 for the non-target trials. Estimation regarding the learning used 276 samples of target trials and 808 samples of non-target trials. The remaining samples were used for the evaluation. To suppress the difference of the evaluation, we evaluated the samples used for learning ten times. The mean value was taken as the result.

C. Result

Fig. 5 shows classification accuracy against the number of samples used to calculate the average. The discrimination rate was the average of the discrimination rates for the target and non-target trials. The target discrimination rate was the number of evaluation samples divided by the number of target trials for all evaluation samples. It was calculated using a method similar to that for calculating the non-target discrimination rate. The average discrimination rate for the three subjects was 66.9[%] when data for a single trial were used and reached a maximum of 71.0[%]. When the best result for each subject was used to calculate the average, the average discrimination rate was 89.2[%]. It was over 22.3[%] when data for a single trial data were used. The best discrimination rate was 95.1[%] (average for ten samples) for subject s3.

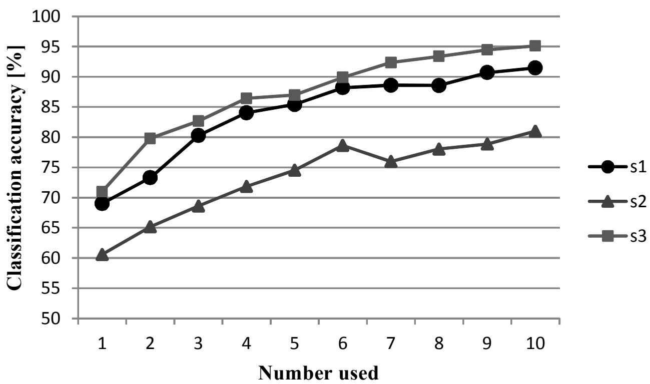

Fig. 5 Average classification accuracy against number of samples used to calculate average

V. DISCUSSION

While the best discrimination rate was 95.1[%] for S3, the best discrimination rate for S2 was 81.0[%]. This suggests that there are significant differences among subjects with out-of-head sound localization.

Since there are many features such as amplitude and latency that may affect the classification, further development is required for the pre-processing and the feature value to make the system more flexible. Since the difference in the discrimination rate between a single trial and the average was 22.3%, it is important to improve the discrimination rate for a single trial.

Therefore, we plan to search for an effective channel and/or time range for discrimination to achieve a flexible, high-performance system and to improve the estimation accuracy. We also plan to investigate the effect on the discrimination rate of using the training for audio image discrimination, and of using classification methods other than SVM.

VI. CONCLUSION

We investigated the ability of a BMI system using out-of-head sound localization to estimate a person’s direction of attention. The subject was given the task of counting the frequency of auditory stimuli from six possible directions. An EEG response with a positive peak and a latency of 300[ms] was observed during the target trials.

We classified EEG signals into two groups (target and non-target) by using support vector machine to classify the EEG waveforms into target and non-target trials. When data for a single trial were used, the average discrimination rate for three subjects was 66.9[%]. It was 89.2[%] when the best result for each subject was used and 95.1[%] for one subject in particular. These results indicate that a person’s direction of attention can be estimated from the EEG response.

REFERENCES

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain computer interfaces for communication and control," Clinical Neurophysiology vol.113, pp. 767-791, 2002.

[2] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: A General-Purpose Brain-Computer Interface (BCI) System," IEICE Transactions on Biomedical Engineering, vol.51, No. 6, pp. 1034-1043, 2004.

[3] E. W. Sellers and E. Donchin, ”A P300-based brain-computer interface:Initial tests by ALS patients,” Clinical Neurophysiology vol. 117, pp. 538-548, 2006.

[4] E. Donchin, K. M. Spencer, and R. Wijesinghe, ”The Mental Prosthesis:Assessing the Speed of a P300-Based Brain.Computer Interface,” IEICE Transactions on Rehabilitation Engineering, vol. 8, No. 2, pp. 174-179, 2000.

[5] S. Shimada and S. Hayashi, "Stereophonic sound image localization system using inner-earphones," Acustica, vol. 81, pp. 264-271, 1995.

[6] S. Yano, H. Hokari, and S. Shimada, ”A Study on Personal Difference in the Transfer Functions of Sound Localization Using Stereo Earphones,” IEICE Transactions on Fundamentals, vol. E83-A, No. 5, pp. 877-887, 2000.