2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

A Study on Asynchronous System in P300 Speller based on User's Intention of Input

Kawai Kohei
Graduate School of Engineering
Nagoya University
Email: yokoi@cmplx.cse.nagoya-u.ac.jp

Tomohiro Yoshikawa
Graduated School of Engineering
Nagoya University
Email: yoshikawa@cse.nagoya-u.ac.jp

Takeshi Furuhashi
Graduated School of Engineering
Nagoya University
Email: furuhashi@cse.nagoya-u.ac.jp

***Abstract*—P300 speller is the communication tool based on Brain Computer Interfaces (BCIs) which allow users to input letters only by thoughts. It uses P300, one of the event-related potential (ERP), as the target feature. In P300 speller, another person starts and closes the system. Therefore, a user cannot switch P300 speller ON/OFF by himself/herself. To solve this problem, an asynchronous P300 speller which can control ON/OFF based on the user's intention of input is needed. In recent years, the intention classification method with additional pre-training has been proposed. In the additional pre-training, the classifier trains non-control state data which is recorded when the user does not input. However, the additional pre-training causes another burden and usage restrictions. In this paper, we propose and study an intention classification method using only training data in which a user inputs letters and an asynchronous system in P300 speller based on the user's intention of input.**

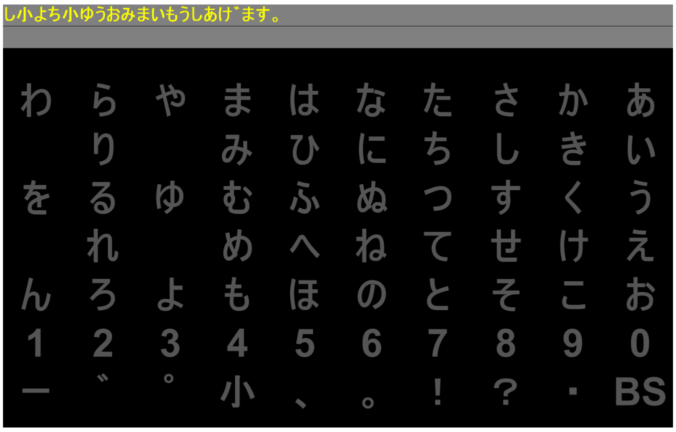

Fig. 1. Interface of P300 speller

I. Introduction

Brain-Computer-Interfaces (BCIs) allow users to control external devices without using muscles based on the brain signals such as electroencephalogram EEG [1]. BCIs are expected to be developed as a communication tool for severely paralyzed patients like those with amyotrophic lateral sclerosis (ALS) [2]. P300 speller [3] is one of the BCIs by which a user can input letters only by thoughts using P300, one of the event related potential (ERP), as the target feature. P300 speller generally employs the interface on which letters are allocated in the form of matrix (see Fig. 1). Each row and column is flashed one by one in random order, which is called stimulus presentation. A user concentrates on his/her desired letter by counting how many times it intensified. When the attended letter is intensified, the P300 is elicited. The system discriminates the user's desired letter that includes the P300 most likely as the target one. However, the patterns of P300 and its features are individually different. Therefore, just before an actual use, the classifier has to be trained with training data in which the user inputs a set of prepared letters (pre-training). Discriminant score for each recorded data is calculated based on the model generated in pre-training, and the system discriminates P300/non-P300 based on the discriminant score.

In P300 speller, another person starts and closes the system generally. Therefore, a user cannot switch P300 speller ON/OFF by himself/herself when the user wants to input letters or to stop the system. Asynchronous P300 speller which can control ON/OFF based on the user's intention of input is needed. Recently, several studies have addressed the issue of the asynchronous control of BCI. Zhang et. al. proposed a computational approach to implement an asynchronous P300-based BCI [4]. F Aloise et.al. proposed an asynchronous gaze-independent BCI [5]. Panicker et al. proposed an asynchronous P300 BCI with SSVEP-based control state detection [6]. In these approaches, in addition to general pre-training, the classifier trains non-intention training data which is recorded when the user does not input letters (non-control state), then the system classifies a user's intention of input and switches ON/OFF based on the classification results. However, in these methods, additional pre-training for non-control state is needed. Moreover, the non-control state in actual input basically has to match the state in the pre-training, or the classifier has to train several types of non-control state.

In this paper, we propose a classification method based on user's intention of input using only intention training data which is recorded when user input letters (control state), i.e. the proposed method does not need additional pre-training. We evaluate the performance of the proposed method in the experiment on classification of input intention and discuss the results.

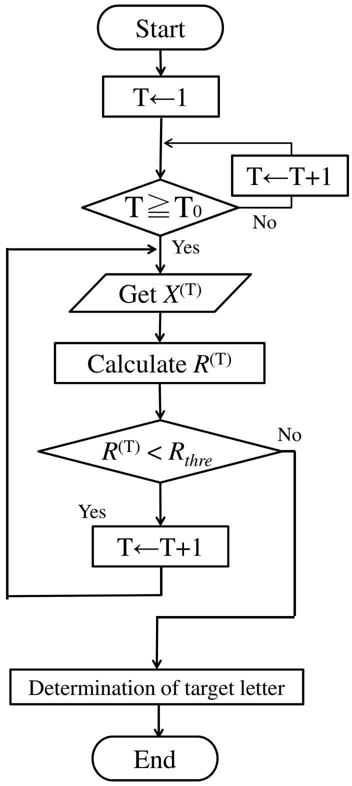

Fig. 2. Flow of Proposed method

II. PROPOSED METHOD

A. A process of Discrimination

Figure 2 shows the flow of the proposed method. In Fig.2, T is the number of sequences (in one sequence, every row and column flashes once in random order) on stimulus presentation. In the beginning or after inputting a letter, the minimum number of sequences on stimulus presentation to be classified is T_0. This is because of the low signal-to-noise ratio of P300. Figure 1 shows 7 10 matrix interface, and the number of candidates in Fig.1 is 17 (7 rows and 10 columns). When ith candidate flashes in tth sequence, the recorded EEG data is represented as $x_i^{(t)}$. In tth sequence, the EEG data denotes $\boldsymbol{x}^{(t)} = \{x_i^{(t)} | i = 1, 2, ..., 17\}$. When the classification of input intention is done, EEG data in several sequences are averaged and the averaged data are used to reduce the noises and improve the classification accuracy. After the stimulus presentation in T sequences, classified EEG data denotes $X^{(T)} = \{\boldsymbol{x}^{(t)} | t = T_s, T_s{+}1, ..., T\}$. T_s is the start of sequence for classification and defined by the following equation.

$$T_s = \begin{cases} 1 & (T < T_L) \\ T - T_L - 1 & (T \geq T_L) \end{cases}$$

T_L is the maximum length of sequences to be classified. When the system classifies user's intention into non-control state, stimuli of one more sequence are presented. Then the system classifies user's intention again using the latest $X^{(T)}$. This intention classification continues until the result of the classification becomes control state. When the system classifies user's intention into control state, the target letter is predicted based on the discriminant scores. In the determination of a target letter, row/column that has the highest averaged score among all rows/columns is identified, and the target letter is predicted by the intersection of the identified row and column.

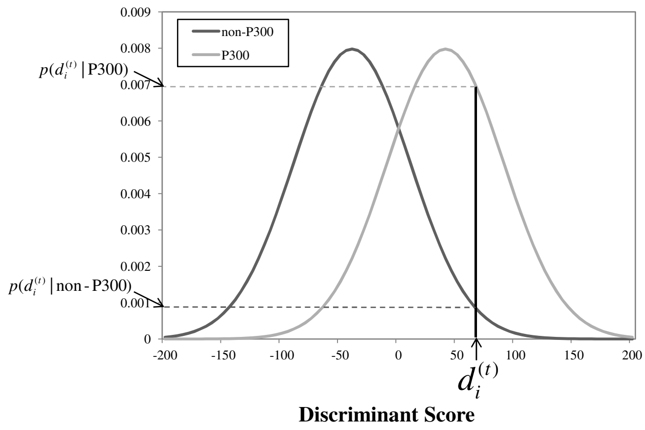

Fig. 3. Score distribution

B. model of P300 and non-P300 discriminant score

Discriminant score value is calculated as

$$d_i^{(t)} = w * x_i^{(t)} \tag{1}$$

where w is the weights assigned based on the training data and $x_i^{(t)}$ denotes the EEG data on ith candidate in tth sequence. In pre-training, a user inputs a set of prepared letters. Therefore, every training data is labeled as P300 data or non-P300 data. Averages of the P300 and the non-P300 discriminant scores (μ_{P300} $\mu_{non\text{-}P300}$) are calculated from the training data. Based on the previous the study [7], each score distribution is assumed as a normal distribution. Figure 3 shows an image of the probability density models of the discriminant score in a set of training data. We employ these score models to describe the likelihood of P300 and non-P300 by the following equations:

$$p(d_i^{(t)}|P300) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(d_i^{(t)} - \mu_{P300})^2}{2\sigma^2}\right\} \tag{2}$$

$$p(d_i^{(t)}|non\text{-}P300) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(d_i^{(t)} - \mu_{non\text{-}P300})^2}{2\sigma^2}\right\} \tag{3}$$

where the parameters for the normal distributions can be learned from the training data.

C. Classification of user's intention

When a user wants to input a letter, the P300 is elicited when the row/column including the target letter flashes. When a user has no intention to input a letter, P300 or a noise similar to P300 is rarely recorded intensively at a certain row/column. In the proposed method, input reliability $R^{(T)}$ is calculated by $X^{(T)}$.

The discriminant scores on the target row/column denotes $d_{target}^{(t)} = \{d_i^{(t)} | i = \hat{r}, \hat{c}\}$, where $\hat{r}$ and $\hat{c}$ are the estimated target row and column based on $X^{(T)}$. After the stimulus presentation in T sequences, the ensemble of discriminant scores on the target row/column denotes $D_{target}^{(T)} = \{d_{target}^{(t)} | t = T_s, T_s + 1, ..., T\}$. $R^{(T)}$ is defined by the following equation.

$$R^{(T)} = P(P300|D_{target}^{(T)}) = \frac{P(D_{target}^{(T)}|P300)P(P300)}{P(D_{target}^{(T)})}$$
$$= \frac{P(D_{target}^{(T)}|P300)P(P300)}{P(D_{target}^{(T)}|P300) + P(D_{target}^{(T)}|non\text{-}P300)} \quad (4)$$

We assume an even probability between $P(P300)$ and $P(non\text{-}P300) = 1/2$. The $P(D_{target}^{(T)}|P300)$ and $P(D_{target}^{(T)}|non\text{-}P300)$ represent the conditional probability density of observation when $D_{target}^{(T)}$ is given. We assume each EEG data is independent one another, so we have

$$P(D_{target}^{(T)}|P300) = \prod_{t,i} p(d_i^{(t)}|P300) \quad (5)$$

$$P(D_{target}^{(T)}|non\text{-}P300) = \prod_{t,i} p(d_i^{(t)}|non\text{-}P300) \quad (6)$$

where $p(d_i^{(t)}|P300)$ and $p(d_i^{(t)}|non\text{-}P300)$ is calculated by ep.(2) and eq.(3), respectively. In the classification of input intention, if $R^{(T)} > R_{thre}$ where R_{thre} is a preset threshold, the system classifies user's intention into control state. If not, the system classifies user's intention into non-control state.

III. Experiment

A. Data Description

In this paper, the interface containing Japanese characters shown in Fig.1 was employed for the P300 speller experiment. The offline experiment was done and it used a recorded dataset which contained EEG data measured by four subjects (Sub1 Sub2 Sub3, Sub4) The EEG data was recorded with sampling frequency of 1000Hz from nine electrodes based on the ten-twenty electrode system of the international federation [8] : Fz, Cz, Pz, O1, O2, P1, P2, C3 and C4, referenced to the linked ears, A1 and A2. The P300 speller implemented in BCI2000 [9], a general-purpose system for brain-computer interface research, was employed. The stimulus onset asynchrony (SOA) was 175ms. One letter consisted of ten sequences, and one sequence contained 17 (10 rows and 7 columns) stimuli. Two types of EEG data were recorded. One was the control state data in which the users were concentrating on the target letter (the users counted how many times the desired letter was intensified). The other was the non-control state data in which the users were paying no attention to a target but looking at the interface. In both cases, EEG for 40 letters (40 10 sequences) were recorded. These EEG signals were down-sampled to 100Hz, 12 data points after each stimulus corresponding to 0ms (50ms) to 600ms by averaging 5 data points in every 50ms were extracted. Stepwise Linear Discriminant Analysis (SWLDA) [10] was employed for the discrimination of P300/non-P300 in this experiment.

B. Experimental Settings

To evaluate the performance of the proposed method, the offline experiments using the data described in III-A were conducted. First, 10 letters (10 10 sequences) of the control state were utilized for the pre-training. The discrimination function and the score model of P300 and non-P300 were calculated by the training data. Then, two types of test data were generated and the proposed method was applied. One was consisted of the control state data of 20 letters (20 10 sequences). In the control state data, there could be three different classification results.

1) Correct Discrimination: the target letter was correctly detected.
2) Failure of Discrimination: The classification of the input intention was correct, but the target letter was not correctly detected.
3) Failure of classification: the system classified into a non-control state, i.e. it could not classify the data into the control state by the latest T_L(=10) sequences.

We calculated the accuracy of classification, the ratio of the number of sequences needed to classify the user's intention into control state over the number of sequences of stimulus and the information transfer rate (ITR). ITR indicates that how many bits of information is able to communicate effectively through the interface [11].

$$\text{ITR} = \frac{\log_2(N) + p\log_2(p) + (1-p)\log_2(\frac{1-p}{N-1})}{d} \quad (7)$$

where *p* denotes the accuracy of classification, *N* denotes the number of choices, i.e. N = 70 in this experiment, and *d* denotes the average time (minutes) to enter one letter in a session.

The other test data was consisted of the non-control state data of 10 minutes (202 sequences). We calculated the false positive rate (FPR) which indicates how many false events (the non-control state into the control state) the system detected on average within 1 minute.

These pre-training and intention classification were conducted for 50 times, and the results were averaged. In this experience, $T_0 = 3$, $T_L = 10$ and $R_{thre} = 0.99$.

IV. Result and Discussions

Table I shows the result in the control state data. In table I, Correct Discrimination achieved on average 89.3% (std.=5.9%), Failure of Discrimination achieved on average 10.5% (std.=5.8%) and Failure of Classification achieved on average 0.2% (std.=0.7%). It shows that incorrect classification was rarely done by the proposed method, while the number of sequences was 3.3 on average (std.=0.2). Figure 4 shows the result of ITR and Fig. 5 shows FPR in the offline experiment. A mean value of ITR was 31.06 bit/min (std.=3.66) and FPR achieved on average 0.11 event/min (std.=0.12).

TABLE I
RESULT OF CONTROL STATE DATA

	Correct Discrimination( )		Failure of Discrimination ( )		Failure of Classification( )		Number of sequences	
	Average	std.	Average	std.	Average	std.	Average	std.
Sub1	90.4	5.9	9.5	5.8	0.1	0.7	3.2	0.2
Sub2	84.9	5.4	14.7	5.2	0.4	1.4	3.5	0.3
Sub3	88.3	6.5	11.6	6.4	0.1	0.7	3.3	0.2
Sub4	93.7	5.7	6.3	5.7	0.0	0.0	3.1	0.1
Average	89.3	5.9	10.5	5.8	0.2	0.7	3.3	0.2

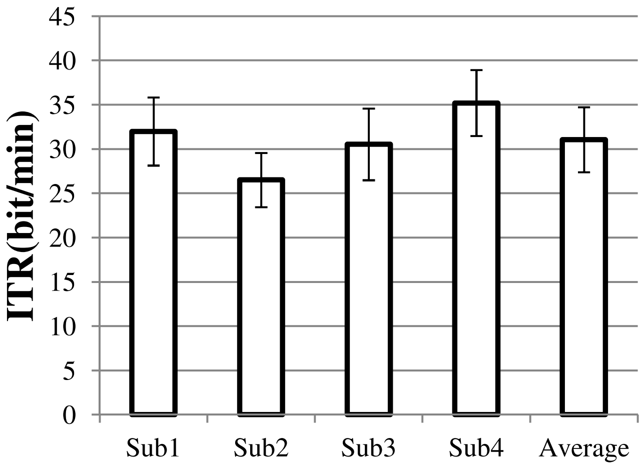

Fig. 4. Information transfer rate (bit/min)

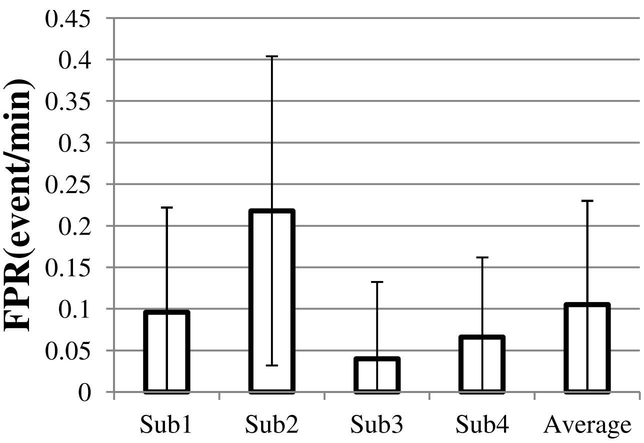

Fig. 5. False positive rate (event/min)

Zhang et al. reported on average 1.0 event/min in FPR, and a mean ITR of 20 bit/min in the offline experiment [4]. Though there were some differences in the experimental settings and the classification approach, these results in this experiment show that the proposed method using only control state data had good ITR with low FPR.

V. CONCLUSIONS

In this paper, we proposed and discussed the intention classification method using only training data in control state. In the offline experiment, the proposed method achieved on average ITR of 31.06 bit/min with a mean FPR of 0.11 event/min. This result showed that the performance of the proposed method was demonstrated good ITR with low false FPR in 4 subjects comparing with the conventional method. We will do online experiments by the proposed method and investigate the proposed method more.

ACKNOWLEDGMENT

This work was supported by the Grant-in-Aid for Scientific Research (C) no. 22500200 from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

REFERENCES

[1] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

[2] D. Kaub-Wittemer, N. Steinbuchel, M. Wasner, G. Laier-Groeneveld, and G. Borasio, "Quality of life and psychosocial issues in ventilated patients with amyotrophic lateral sclerosis and their caregivers," *Journal of pain and symptom management*, vol. 26, no. 4, pp. 890–896, 2003.

[3] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.

[4] H. Zhang, C. Guan, and C. Wang, "Asynchronous p300-based brain-computer interfaces: A computational approach with statistical models," *IEEE Transactions on Biomedical Engineering*, no. 6, pp. 1754–1763, 2008.

[5] F. Aloise, P. Arico, F. Schettini, S. Salinari, D. Mattia, and F. Cincotti, "Asynchronous gaze-independent event-related potential-based brain-computer interface," *Artificial Intelligence in Medicine*, vol. 59, no. 2, pp. 61–69, 2013.

[6] R. Panicker, S. Puthusserypady, and Y. Sun, "An asynchronous p300 bci with ssvep-based control state detection," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1781–1788, 2011.

[7] Y. Kaneda, H. Takahashi, T. Yoshikawa, and T. Furuhashi, "A study on reduction of discrimination time of p300 speller applied to reliability-based automatic repeat reQuest," in *6th Brain and Perception workshop*, 2009.

[8] H.H.Jasper, "The ten-twenty electrode system of the international federation," *Electroencephalography and clinical neurophysiology*, vol. 10, no. 2, pp. 371–375, 1958.

[9] G. Schalk, D. McFarland, T. Hinterberger, N. Birbaumer, and J. Wolpaw, "A general-purpose brain-computer interface (bci) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

[10] D. J. Krusienski, E. W. Sellers, S. Bayoudh, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "A comparison of classification techniques for the p300 speller," *Journal of Neural Engineering*, vol. 3, no. 4, pp. 299–305, 2006.

[11] B. Dal Seno, M. Matteucci, and L. Mainardi, "The utility metric: a novel method to assess the overall performance of discrete brain–computer interfaces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 1, pp. 20–28, 2010.