A Study on Classification Performance of Rapid Serial Visual Presentation with Small Choices

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

Hiromu Takahashi
Graduated School of Engineering
Nagoya University
Email: takahashi@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*—In the conventional P300-based Brain Computer Interface, characters or commands are matrix-arrayed and each row or column is randomly and successively intensified. In this interface, a user needs to move his or her eye gaze to the character or the command that the user desires, which could be a problem for those with severe motor disability. To solve this problem, the use of Rapid Serial Visual Presentation (RSVP) in the P300-based BCIs has been proposed. However, it is said that a longer classification time is needed in RSVP, because characters or commands are presented one by one in this method. However, the increase of the classification time depends on the number of choices. Thus, this paper compares the performance of RSVP and the conventional interface with fewer characters or commands. Moreover, reliability based selective stimulus presentation is applied to RSVP, and this paper studies on the improvement of classification time.**

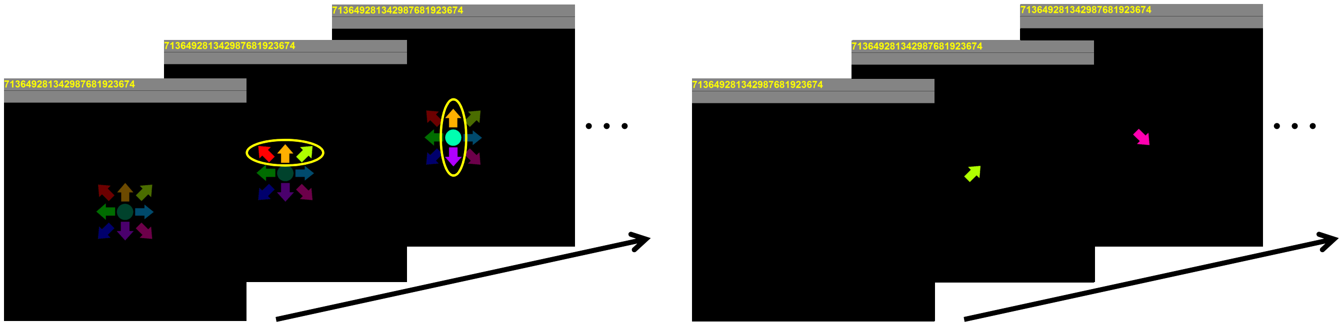
I. Introduction

Brain-computer interfaces (BCIs) allow a user to communicate and control external devices without using muscles [1]. BCIs are appealing to severely paralyzed patients like those with amyotrophic lateral sclerosis (ALS) [2]. Moreover they are also appealing to healthy people as amusement applications. The authors have been studying on the BCIs aimed at inputting characters [3] [4]. However, restoring ability of movement is thought to be significant as well as that of communication. Thus, the purpose of this study is the development of BCI for movement.

Generally, motor imagery based method is used in BCI aimed movement, because the operation is intuitive for users [5]. However, in this method, the user needs training for a few weeks to achieve a classification accuracy of 80% [1]. It is reported that ALS patients die within 3 to 5 years from onset [2]. Thus, the motor imagery based BCIs that require long training periods might be impractical for ALS patients whose progression of disease is fast. On the other hand, the P300, one of the event related potential (ERP), based BCIs do not require long training periods. In this regard, the P300-based BCIs could be superior to the motor imagery based BCIs. Because of this, the authors have been developing the P300-based BCls aimed at movement (P300 controller).

The P300 speller is one of the most popular P300-based BCIs. Generally, it employs a letter matrix interface, where each row and column is randomly flashed one by one prefixed number of times, while a user is just required to count how many times his or her desired letter is intensified. When the attended letter is intensified, the P300 is elicited. Thus, the letter whose flashes have most likely elicited the P300 is determined as the user's target letter [6]. However, the classification accuracy depends on eye gaze of user in the matrix interface [7]. Thus, it is thought to be unsuitable for the ALS patients who are difficult to move his or her eye gaze. To solve this problem, rapid serial visual presentation (RSVP) that presents letters one by one in sequence at the center of the interface was proposed [8]. This method is independent from eye gaze. However, classification time in the RSVP becomes longer than that in the matrix interface, because the number of choices that the RSVP presents at one time is smaller than that the matrix interface presents. On the other hand, it would not matter practically in the case that the number of total choices is small such as a movement control, because the number of presentations of stimuli is also small with the RSVP. Thus, the first purpose of this paper is the performance comparison between a RSVP interface and a matrix interface, both of which have 9 choices in P300 controller.

In conventional methods, the number of stimuli per letter is determined in advance. On the other hand, some methods in which the number is determined dynamically based on language model are reported and proven effective [9]. *Reliability-based automatic repeat request* (RB-ARQ) that have proposed by the authors is one of them [3]. Moreover, we proposed *reliability-based selective repeat ARQ* (RB-SR-ARQ) that is improved RB-ARQ [4]. In this method, stimuli are presented selectively and not randomly to increase the reliability effectively. In this paper, we apply RB-SR-ARQ to a RSVP. And, the second purpose of this paper is the study on how much RB-SR-ARQ can improve classification time with the RSVP.



(a) Matrix interface

(b) RSVP interface

Fig. 1. Interface

II. Method

A. Interface

In this paper, two interfaces as in Fig.1 are employed. In the case using the matrix interface (Fig.1(a)), the P300 is elicited when the row and the column including the target is intensified. On the other hand, in the case using the RSVP interface (Fig.1(b)), P300 is elicited when the target is presented in the center of the interface.

It is reported that when different choices are drawn in different colors, the recognition rate of the target increases in the RSVP interface [8]. Thus, in this paper, different choices were drawn in different colors. Specifically, three attributes of color [10]: hue, saturation and value were employed. When the hue, defined in an interval [0, 360), is been altering, chromatic color is been changing circularly. Thus, in this paper, the saturation and the value remained at the same values, and different hue were assigned to each choice so that they were arranged at regular intervals in the domain of definition.

B. Reliability-based ARQ

The posterior probability of each letter is calculated after presentation of a pre-fixed number of stimuli using the P300 speller implemented in BCI2000 [11]. The letter that has the maximum posterior probability λ_n is predicted as the target. However, in this case, even if the maximum posterior probability is almost 100% in the course of stimulus presentation, stimuli are kept presented for the pre-fixed number of times. Because of this, classification time will increase. To solve this problem, the authors have proposed *reliability-based automatic repeat request* (RB-ARQ). λ_i can be regarded as the reliability of the letter detection because it is probability that the target letter is correctly detected after the presentation of the *i*th stimulus. In RB-ARQ, stimulus presentation continues until λ_i become greater than an arbitrary threshold λ. Note that, the order of stimulus presentation is random (see the detail in [3]).

C. Reliability-based Selective Repeat ARQ

As mentioned in the section II-B, the order of stimulus presentation is random in RB-ARQ. On the other hand, stimuli are presented in such a way as to increase reliability, i. e., maximum posterior probability, effectively in RB-SR-ARQ.

Given *i*-1 stimuli are already presented, an expected reliability $E[\lambda_i]$, which is an expected value of the maximum posterior probability when the *i* th stimulus s_i can be presented, is calculated for all stimuli: all rows and columns in the matrix interface, and all choices in the RSVP interface. Let $s(j)'_i$ be the stimulus that has the *j* th largest expected reliability at the *i* th stimulus. Presenting $s(1)'_i$ is expected to increase the reliability most effectively [4]. However, stimuli are presented in the same way with RB-ARQ in the first two sequences, i.e., 9 different kinds of stimuli or three rows and columns per sequence, because stimuli can not be ordered at the beginning if each letter has an equal prior probability, and the influence of outliers is preferred to be reduced if contaminated in the first several epochs.

However, it is reported in [4] that the stimulus including a target letter (target stimulus) and the stimulus next to it were presented alternately by RB-SR-ARQ. Moreover, it can be simulated the target stimulus is presented continuously. On the other hand, it is reported that the second target can be hardly recognized if two different target stimuli or two identical target stimuli are presented with an interval of less than 500 ms [12]. If the stimulus onset asynchrony (SOA) is over 500 ms, this problem would become less significant, however, the classification time will becomes longer. Thus, this paper employed the following exceptions with SOA of 180 ms. If $s(1)'_i$ is not presented at the last two times, $s_i = s(1)'_i$. When $s(1)'_i$ is presented at either of the last two but $s(2)'_i$ is not at the last two times, $s_i = s(2)'_i$. If none of the above, s_i is randomly chosen from all stimuli except for $s(1)'_i$ and $s(2)'_i$. By these exceptions, it can be prevented that the target stimulus is presented continuously or iteratively.

III. Experiment

A. Experimental setting

In this experiment, the P300 speller implemented in the BCI2000 was employed. EEGs were recorded from five electrodes: Fz, Cz, Pz, O1, and O2 referenced to the linked-ears, with the sampling rate of 100 Hz using a Polymate AP216 (DIGITEX LAB. CO., LTD, Tokyo, Japan). The SOA was

TABLE I
EXPERIMENT SETTING FOR EACH SUBJECT

sub	electrode	epoch length[ms]
1	Fz, Cz, Pz	0-450
2	Fz, O_1, Cz, Pz	0-400
3	Fz, Cz, Pz	0-500
4	Fz, O_1, Cz, Pz	0-450

180 ms: each stimulus was presented for 100 ms with an inter-stimulus interval of 80 ms. A high-pass filter with a cut-off frequency of 1 Hz and then a moving average with a window length of 5 were applied to EEGs. After a step-wise variable selection, a subject-specific LDA classifier was trained using the learning dataset and it was employed in the succeeding test sessions.

Four healthy subjects (sub 1 - sub 4, all male) volunteered to participate in this experiment. In the learning and test sessions, they were required to try to enter the following 24 directions: three sets of all choices except for "stop" (the circle that is placed at the center of Fig.1(a)) which were ordered randomly (24 = (9-1)×3). First, the subjects performed the learning sessions using each interface in Fig.1. After subject-specific classifiers for each interface were trained, four test sessions were performed: two RB-ARQ sessions with either the RSVP interface or the matrix interface and two RB-SR-ARQ sessions with each interface. These four conditions were assigned in the order that brought a balance among subjects. In addition, "stop" was excluded from the task because it was much different from other choices visually and semantically.

The threshold was set to be 0.9 for all test sessions. Moreover, each subject performed four test sessions, each of which with 24 directions, with the maximum number of stimuli of 90, i.e., even if the reliability had not reached the threshold after the 90th stimulus, the stimulus presentation discontinued and the direction with the maximum posterior probability was selected as the target one. If the reliability reached the threshold before 2 sequences (12 stimulus presentation with the matrix interface, 18 stimulus presentation with the RSVP interface), the improvement of classification time with RB-SR-ARQ can not be confirmed. Thus, the epoch length and the electrodes used for the variable selection were adjusted so that the reliability was expected not to reach 90% before 5 sequences in this experiment. Table I shows the epoch length and the electrodes that were used for each subject.

B. Performance assessment

In this paper, *Information Transfer Rate* (ITR), often used in the performance assessment in the discipline of BCI, was employed as the performance measure [13].

$$\mathrm{ITR} = \frac{\log_2(N) + p\log_2(p) + (1-p)\log_2(\frac{1-p}{N-1})}{d} \qquad (1)$$

where p denotes the accuracy in a session, N denotes the number of choices, i.e., N = 9 in this experiment and d denotes the average time to enter one direction in a session.

IV. RESULT AND DISCUSSIONS

Table II shows the accuracy, the average number of stimuli and the ITR. Figure 2 shows the ITR of each subject, where RB means the RB-ARQ condition and SR means the RB-SR-ARQ condition.

The result of this experiment was significantly different among subjects. In the case of sub 1 and sub 2, the accuracy of classification with the matrix interface was approximately equal or greater than the threshold, i.e., 0.9, while the accuracy with the RSVP interface was smaller than that. On the other hand, in the case of sub 3 and sub 4, the accuracy of classification with the RSVP interface was much greater than the threshold, i.e., 0.9, while the accuracy with the matrix interface was smaller than that.

Let us compare the conditions in terms of the ITR, which is comprised of the accuracy and the number of stimuli. As shown in Fig. 2, in the case of sub 1 and sub 2, the ITR with the matrix interface was higher than that with the RSVP interface. On the other hand, in the case of sub 3 and sub 4, the ITR with the RSVP interface was higher than that with the matrix interface. As seen above, ITR showed the same tendency as the accuracy. A repeated measures two-way (RB vs. SR and matrix vs. RSVP) analysis of variance was conducted with the respect to the ITR and the result showed there was a statistically significant difference between the two interfaces (p(RB vs. SR) = 0.62, p(interface) = 0.032). Moreover, Fig. 3 shows the interaction plot[14], and it shows the better result of the RSVP interface than the matrix interface. Thus, the RSVP interface was more effective than the matrix interface in nine choices.

As previously described, the result of this experiment was highly dependent on each individual. For example, the results of sub 1 and sub 2 show that the matrix interface was more effective than the RSVP interface. To elucidate the cause of the bad accuracy of sub 1 and sub 2 in the RSVP interface conditions, the details of misclassification was investigated. Table III shows the result of misclassification in the RSVP interface for sub 1 and sub 2. In table III, a lot of misclassification in sub 2 have occurred after presentation of the maximum number of stimuli. Thus, it can be inferred that ninety stimuli were inadequate to reach the threshold for sub 2.

On the other hand, it makes no sense that this is the cause of bad accuracy of sub 1 in the RSVP interface and RB-SR-ARQ condition. Table III shows that the most misclassification of sub 1 in RB-SR-ARQ condition occurred next to the target in the hue circle, i. e., the choice whose hue was different from the target choice by 40 degrees. Moreover, we checked the order of stimulus presentation in this case, and found that the target stimulus, the stimulus similar to the target in color and a stimulus chosen randomly except for the former two stimuli were presented periodically and repeatedly. Thus, it can be inferred that the *Attentional blink* occurred; *Attentional blink* is the phenomenon that the second target can be hardly recognized if two different target stimuli are presented with an interval of less than 500 ms [12]. Moreover, it can be

TABLE II
ACCURACY, NUMBER OF STIMULI, ITR

		Accuracy[%]		Number of stimuli		ITR[bps]	
Interface		matrix	RSVP	matrix	RSVP	matrix	RSVP
sub 1	RB	83.3	79.1	59.6	56.6	0.114	0.0828
	SR	91.7	66.7	44.8	44.0	0.166	0.0838
sub 2	RB	87.5	75.0	36.3	78.6	0.166	0.0760
	SR	100	70.8	19.6	71.3	0.301	0.0718
sub 3	RB	87.5	100	64.9	27.1	0.121	0.267
	SR	70.8	100	70.9	27.0	0.0721	0.267
sub 4	RB	79.2	95.8	48.1	20.3	0.115	0.262
	SR	83.3	100	38.1	20.0	0.146	0.299

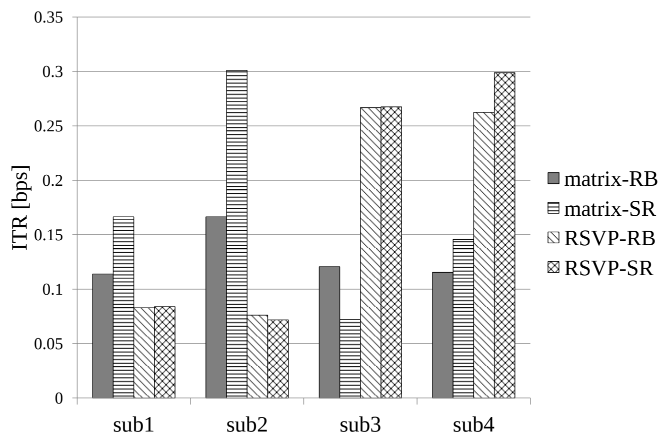


Fig. 2. Information transfer rate

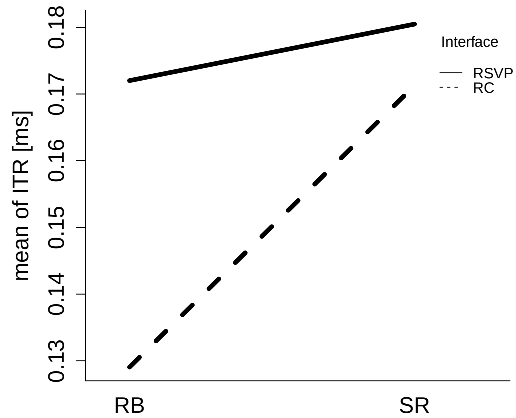


Fig. 3. Interaction plot

inferred that *Alternation effect* occurred; *Alternation effect* is the phenomenon that an alternation of stimuli produces small P300s [15]. These two phenomena were likely to occur, and it can be inferred that the accuracy of sub 1 in RB-SR-ARQ condition decreased because of these phenomena. On the other hand, it can be also inferred that the cause of bad accuracy of sub 1 in RB-ARQ condition was the same cause of sub 2 because the order of stimulus presentation was random in RB-ARQ condition.

In the future work, the choices have to be additionally differentiated to solve *Attentional blink* and then to avoid false recognition. Moreover, to solve *Alternation effect*, it might be effective to make the interval of the target presentation random in RB-SR-ARQ condition.

V. CONCLUSION

This paper scrutinized the effectiveness of the eye gaze independent interface (RSVP interface) which had 9 choices. It was concerned that the RSVP interface which presented the choices one by one increased classification time from that of the matrix interface. However, this paper showed that the RSVP interface was more effective statistically in this sense than the matrix interface. Thought the second purpose of this paper was to improve the classification time of the RSVP interface using *Reliability-based Selective Repeat ARQ*, the results of the experiment in this paper were not enough to show the effectiveness. Thus, the future work is the improvement of the RSVP interface and the application of RB-SR-ARQ to P300 controller.

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] Y. Kaneda, H. Takahashi, T. Yoshikawa, and T. Furuhashi, “A Study on Reliability-Based Selective Repeat Automatic Repeat reQuest for Reduction of Discrimination Time of P300 spellert,” *Soft Computing Intelligent Systems and International Symposium on Advanced Intelligent Systems*, pp. 982–986, 2010.

[4] H. Takahashi, T. Yoshikawa, and T. Furuhashi, “Selective Stimulus Presentation based on Target-to-target Interval for P300 Speller,” in *Proceedings of The 26th Symposium on Biological and Physiological Engineering*, pp. 137–142, 2011.

[5] J. del R. Millan, F. Galan, D. Vanhooydonck, J. P. E. Lew, and M. Nuttin, “Asynchronous Non-Invasive Brain-Actuated Controlof an Intelligent Wheelchair,” in *31st Annual International Conference of the IEEE EMBS*, pp. 3361–3364, 2009.

[6] 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.

[7] P. Brunner, S. Joshi, S. Briskin, J. Wolpaw, H. Bischof, and G. Schalk, “Does the’P300’speller depend on eye gaze?,” *Journal of neural engineering*. vol.7, 056013, 9 pages, 2010.

[8] L. Acqualagna and B. Blankertz, “A gaze independent spelling based on rapid serial visual presentation,” in *33rd Annual International Conference of the IEEE EMBS*, pp. 4560–4563, 2011.

[9] U. Orhan, D. Erdogmus, B. Roark, S. Purwar, K. E. H. II, B. Oken, H. Nezamfar, and M. Fried-Oken, “Fusion with Language Models Improves Spelling Accuracy for ERP-based Brain Computer Interface Spellers,” in *33rd Annual International Conference of the IEEE EMBS*, pp. 5774–5777, 2011.

[10] H. R. Kang, *Color technology for electronic imaging devices.* Society of Photo Optical, 1997.

TABLE III
THE RESULT IN RSVP INTERFACE CONDITION (SUB 1, SUB 2)

Sub	Presentation pattern	Accuracy	Number of classification	Number of misclassification	Maximum stimulus[1]	Similar stimulus[2]
1	RB	0.791	24	5	2	4
	SR	0.667	24	8	2	7
2	RB	0.750	24	6	4	4
	SR	0.708	24	7	5	2

[1] Number of misclassification after presentation of the maximum number of stimuli.
[2] Number of misclassification of the target direction by the choice that was next to it in the hue circle.

[11] 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.

[12] C. Cinel, R. Poli, and L. Citi, "Possible sources of perceptual errors in P300-based speller paradigm," *JBiomedizinische technik*, vol. 49, pp. 39–40, 2004.

[13] 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.

[14] D. G. Kleinbaum, L. L. Kupper, and K. E. Muller, *Applied regression analysis and other multivariable methods*. Duxbury, 1997.

[15] J. Ford, C. Duncan-Johnson, A. Pfefferbaum, and B. Kopell, "Expectancy for Events in Old Age:Stimulus Sequence Effects on P300 and reaction time," *Journal of Gerontology*, vol. 37, no. 6, pp. 696–704, 1982.