A Study on Reliability-Based Selective Repeat Automatic Repeat reQuest for Reduction of Discrimination Time of P300 speller

Yusuke Kaneda
Nagoya Universuty
kaneda@cmplx.cse.nagoya-u.ac.jp

Hiromu Takahashi
Nagoya Universuty
takahashi@cmplx.cse.nagoya-u.ac.jp

Tomohiro Yoshikawa
Nagoya Universuty
yoshikawa@cse.nagoya-u.ac.jp

Takeshi Furuhashi
Nagoya Universuty
furuhashi@cse.nagoya-u.ac.jp

***Abstract*—Brain-Computer Interfaces (BCIs) control a computer or a machine based on the information of the signal of the human brain, and the P300 speller is one of the BCI communication tools. The P300 speller discriminates a character after averaging EEG data to improve the accuracy. Whereas Reliability-Based Automatic Repeat reQuest (RB-ARQ) is an error control method designed for BCIs, which makes a user keep thinking until a given reliability is satisfied and can improve the accuracy of BCI with a small loss of the discrimination speed. This paper proposes Reliability-Based Selective-Repeat ARQ (RB-SR-ARQ), which selectively requests a user to re-send the data based on the reliability of each data. The results show that the time required for thought discrimination can be reduced while the accuracy remains at a high level.**

I. Introduction

Recently, the study of Brain-Computer Interfaces (BCIs) has been researched actively. BCIs are interfaces which link the human brain and a computer. With the help of BCIs, it is expected that a patient like Amyotrophic Lateral Sclerosis (ALS) is able to control a computer and to communicate with other people. As for the measurement of brain activity, the electroencephalogram (EEG) has been used well for BCIs because it is noninvasive and inexpensive. The P300 speller [1], which was developed by Farwell and others, is one of the communication tools of BCI based on the P300 response, which is one of the event related potentials (ERPs) obtained from EEG. The P300 speller typically has 36 characters containing alphabets and numbers on 6×6 matrix (Fig.1), and one of the rows or the columns is highlighted at random. It discriminates a character using P300 evoked by the highlight of a row or a column which includes the character that a user attended. In this paper, the 12 highlights, 6 rows and 6 columns, are called 1 sequence. Usually a character is discriminated after averaging the EEG data corresponding to some sequences to improve the signal-to-noise ratio [2]. In fact, it is reported that the discrimination accuracy is increased by lengthening the time per discrimination, in other words, by increasing the number of data to be averaged [3], [4]. However, it is demanded that the time should be shorter while the accuracy is higher for the practicality of P300 speller.

On the other side, there are various error control methods at the field of information and communications. For example, Automatic Repeat reQuest (ARQ), which asks the transmitter to repeat code words, is one of the popular methods; and Reliability-Based Hybrid ARQ (RB-HARQ) [5], which requests retransmission based upon the reliability of each bit in code words, has been recently proposed. Also, Reliability-Based ARQ (RB-ARQ) [6], which customized RB-HARQ to be suitable for BCIs, has been proposed; and it is reported that RB-ARQ is an efficient error control method, i.e., RB-ARQ can reduce the discrimination time while the accuracy remains at a high level. In this paper, we examine the effectiveness of RB-ARQ when it is applied to the P300 speller, and show that RB-ARQ is more efficient than the averaging. Furthermore, we propose Reliability-Based Selective-Repeat ARQ (RB-SR-ARQ), which requests the data selectively, and compare with RB-ARQ.

A	B	C	D	E	F
G	H	I	J	K	L
M	N	O	P	Q	R
S	T	U	V	W	X
Y	Z	1	2	3	4
5	6	7	8	9	_

Fig. 1. P300 speller interface (6×6 matrix display)

II. Method

A. Reliability-Based Automatic Repeat reQuest

RB-ARQ is an error control method to improve the accuracy in BCIs, in which a user keeps thinking the same thought until a certain criterion is satisfied.

Let K be a set of thoughts, and $\boldsymbol{x}$ be the EEG data, belonging to $u \in K$. The posterior probability of $\boldsymbol{x}$ being in $k \in K$ is written as $P(k|\boldsymbol{x})$, and the label of $\boldsymbol{x}$ is predicted as follows:

$$\hat{u} = \arg\max_{k} P(k|\boldsymbol{x}), \tag{1}$$

where $\hat{u}$ denotes a predicted label. The maximum of the posterior probability is equivalent to the probability of correct discrimination; thus, it can be seen as a reliability of data. Let $X_T = \{\boldsymbol{x}_t | t = 1, 2, ..., T\}$ be a set of data at time T, the

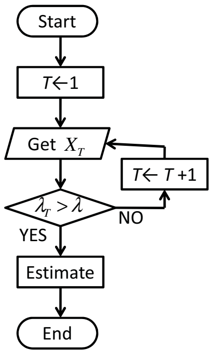


Fig. 2. Flow chart of RB-ARQ

reliability at time T λ_T can be calculated as follows:

$$\lambda_T = \max_k P(k|X_T) = \max_k \frac{\prod_t P(k|\boldsymbol{x}_t)}{\sum_{l \in K} \prod_t P(l|\boldsymbol{x}_t)}. \quad (2)$$

A user keeps thinking the same thought until $\lambda_T > \lambda$ is satisfied, where λ is a given threshold (Fig.2).

B. Application to P300 speller

The P300 speller discriminates a character by specifying the row and the column containing the P300 response.
Let $\boldsymbol{x}_R^r$ be a set of EEG data obtaind when the $r(\in \{1, ..., 6\})$th row is intensified. Using the posterior probabilities of a P300 appearing and not appearing at the rth row given $\boldsymbol{x}_R^r$, each of which is denoted by $P(1|\boldsymbol{x}_R^r)$ and $P(0|\boldsymbol{x}_R^r)$, the posterior probability of the P300 appearing at only the rth row given X_R $P(r|X_R)$ is calculated as follows:

$$P(r|X_R) = \frac{P(1|\boldsymbol{x}_R^r) \prod_{r \neq r'} P(0|\boldsymbol{x}_R^{r'})}{\sum_s (P(1|\boldsymbol{x}_R^s) \prod_{s \neq s'} P(0|\boldsymbol{x}_R^{s'}))}, \quad (3)$$

where $X_R = \{\boldsymbol{x}_R^1, \boldsymbol{x}_R^2, ..., \boldsymbol{x}_R^6\}$. The posterior probability of the P300 appearing at only the cth column $P(c|X_C)$ ($X_C = \{\boldsymbol{x}_C^1, \boldsymbol{x}_C^2, ..., \boldsymbol{x}_C^6\}$) is similarly obtained. The attended row and the column can be estimated by using these posterior probabilities as follows:

$$\begin{aligned} (\hat{u}_r, \hat{u}_c) &= \arg\max_{r,c} P(r, c|X_R, X_C), && (4) \\ &= (\arg\max_r P(r|X_R), \arg\max_c P(c|X_C)), && (5) \end{aligned}$$

where $\hat{u}_r$ and $\hat{u}_c$ are the estimated row and column, respectively. Let X_R^t be a set of EEG data corresponding to row flashes at tth sequence, and $X_{RT} = \{X_R^t | t = 1, 2, ..., T\}$ be a set of X_R^t at Tth sequence. Similarly, X_C^t and X_{CT} are also defined. Then, the reliability of (X_{RT}, X_{CT}) λ_T can be calculated as follows,

$$\lambda_T = \max_{r,c} P(r, c|X_{RT}, X_{CT}). \quad (6)$$

This is equal to the product of the reliabilities, i.e., $\lambda_{RT} = \max_r P(r|X_{RT})$ and $\lambda_{CT} = \max_c P(c|X_{CT})$.

C. Reliability-Based Selective-Repeat ARQ

There is an error control method Selective-Repeat ARQ (SR-ARQ) [7] that requests the sender to re-transmit only the data causing errors on the receiver side. We propose Reliability-Based Selective-Repeat ARQ (RB-SR-ARQ), which introduces the idea of requesting the data selectively into RB-ARQ. The discrimination time will be shortened by requesting rows or columns selectively, i.e., by decreasing the number of flashes in 1 sequence. The discrimination time can be decreased most greatly by requesting only the row and the column whose posterior probability is the highest among rows and columns, respectively. However, the signal-to-noise ratio of P300 would be small if only one row or one column is highlighted, because the amplitude of the P300 is known to be inversely proportional to the target stimulus frequency [8]. Therefore, it is needed to highlight rows and columns even though the posterior probability is not the highest. Thus, in this paper, the rows or the columns whose posterior probabilities $P(r|X_{RT})(r \in \{1, ..., 6\})$ or $P(c|X_{CT})(c \in \{1, ..., 6\})$ are lower than a certain threshold λ'_R or λ'_C ($\{r|\lambda'_R \geq P(r|X_{RT})\}$, $\{c|\lambda'_C \geq P(c|X_{CT})\}$) are highlighted at the same time. It could prevent the frequency of flashing a specific row or column becoming extremely high.

III. EXPERIMENT

A. Offline Experiment

1) Experimental Data: This offline experiment used the BCI Competition III data set II [9], which contained EEG data measured while two subjects (Sub A, Sub B) performed the P300 speller. EEG data was recorded with the sampling frequency of 240Hz, from 64 electrodes and a band-pass filter of 0.1 to 60Hz. The dataset contained 85 learning data and 100 test data. Note that 15 sequences were repeated per character.

2) Preprocessing and Classifier: The features for discrimination were extracted in accordance with the winner method in the competition [10]. The raw EEG data were filtered with a band-pass of 0.1 to 10Hz and down-sampled to 20Hz, then 14 data points corresponding to 0s to 0.65s after the flash were extracted. The extracted data were classified by using Linear Discriminant Analysis (LDA) in this experiment, though the winner method used Support Vector Machine (SVM) in [10], because the standard SVM do not provide the posterior probabilities.

3) Application of RB-ARQ and RB-SR-ARQ: RB-ARQ was applied to experimental data described above and the resultant of accuracy was examined while varying the threshold. In this paper, when the number of sequences T reached 15, the discrimination was done even if the maximum posterior probability did not exceed the threshold because of the provided data format. The accuracy when the standard avaraging was appled was also examined by varying the "fixed" number of sequences from 1 to 15. RB-SR-ARQ was applied to experimental data and the resultant of accuracy was examined while varying the threshold.

4) Performance Evaluation: This paper utilizes Utility [11] defined in Eq.(7), which is an alternative measure for the P300 speller to the information transfer rate (ITR).

$$U = \frac{(2P-1)\log_2(N-1)}{d} \tag{7}$$

where N is the number of classes (in this experiment, $N = 36$), P is the accuracy, and d is the discrimination time. Note that if $P < 0.5$, $U = 0$. In fact, Utility corresponds to the ITR when the spelling is done perfectly using BackSpace that can erase a mis-spelled character; thus, it is thought to be a practical performance measure to the P300 speller. In this paper, Utility was obtained by 10-fold cross varidation of learning data.

B. Online Experiment

This experiment was intended to evaluate how RB-ARQ works online.

1) Experimental Data: RB-ARQ was implemented in the P300 speller available in BCI2000 [12]. Four subjects (Sub 1, Sub 2, Sub 3, Sub 4) participated in this experiment, and their EEGs were recorded with the sampling frequency of 100Hz from five electrodes (Fz, Cz, Pz, O1, O2). The inter stimulus interval (ISI) was 80ms. Each session consisted of spelling 20 characters "THE QUICK BROWNY FOX" including spaces, and each subject performed eleven sessions. The first session was the learning phase: the number of sequences was fixed to be 15, and the optimum threshold λ for RB-ARQ and the optimum number of sequences for the averaging were determined so that the expected Utility was maximized by using 10-fold cross-validation. The rest sessions were the test phase: five sessions with RB-ARQ and five sessions with the averaging in an alternating order.

2) Preprocessing and Classifier: 65 data points corresponding to a time-window from 0.01s to 0.65s after the flash were extracted from the raw EEG data. After selecting the variable by Stepwise Linear Discriminant Analysis (SWLDA) [13], the test data were classified by LDA.

IV. Results and Discussions

Fig.3 shows the relation between the accuracy and the discrimination time in the cases of the averaging, RB-ARQ, and RB-SR-ARQ to competition data, and Fig.4 shows the relation between Utility and the discrimination time. The discrimination time was calculated using the mean value of such T that $\lambda_T > \lambda$ in the cases of RB-ARQ and RB-SR-ARQ, and the number of sequences in the case of the averaging. According to Fig.3, the longer the discrimination time is, the higher the accuracy is. Also, Fig.3 tells that the discrimination times of RB-ARQ are shorter than those of averaging at the same accuracy. Table I shows the comparison of the discrimination time at the accuracy of 90% in each method. This Table tells that the high accuracy of same degree compared with averaging is obtained while the discrimination time is fewer more than 7 seconds in the case of RB-ARQ. Also, according to Fig.4, it is found that the maximum of Utility in the case of RB-ARQ is bigger than that of averaging. Considering these results, the effectiveness of applying RB-ARQ to the P300 speller could be confirmed.

TABLE I
THE DISCRIMINATION TIME AT THE ACCURACY OF 90%

	Method	Time[s]
Sub A	Averaging	29.9
	RB-ARQ	18.7
	RB-SR-ARQ	13.1
Sub B	Averaging	21.6
	RB-ARQ	14.6
	RB-SR-ARQ	10.5

Fig.5 shows the probability density distribution of discriminant score of P300/non-P300 estimated by the Gauss kernel function [14], Fig.6 shows the Q-Q plot [15] of each distribution, and Table II shows the result of Shapiro-Wilk test [16] for each. According to Table II, it can be said that none of distribution is statistically the Gaussian distribution at the significant level of $\alpha = 0.05$. However, these distributions can be practically regarded as the Gaussian distribution because the test statistic W is nearly equal to 1, and each data point of Fig.6 is almost on the straight line. These distributions of Fig.5 also can be found the near Gaussian distribution. According to Fig.6, both ends of plotted data points are not on the line; and this would be the reason the P-value of Shapiro-Wilk test have become low.

Since none of artifact removal was applied, these data points could correspond to the time-windows containing artifacts such as eye movements. Therefore, these outliers could possibly be removed by applying an artifact removing method [17].

Table III shows the result of the accuracy and the discrimination time obtained from the online experiment, and Fig.7 shows the comparison of Utility between the averaging and RB-ARQ. The effectiveness of the application of RB-ARQ to P300 speller can be confirmed from the result that the Utility of RB-ARQ is bigger than that of averaging in each subject. Also, there is a significant difference (P-value = 0.029) in Utility between the averaging and RB-ARQ by repeated measures ANOVA at the significant level of $\alpha = 0.05$. It is needed that the experiment of P300 speller included BackSpace because the Utility premises erasures of mis-spelled characters using BackSpace which was not available in this experiment.

According to Fig.3, it is found that the discrimination times of RB-SR-ARQ are the shortest at the same accuracy in three methods. According to Fig.4, the effectiveness of the application of RB-SR-ARQ can be confirmed because the Utility of RB-SR-ARQ is bigger than that of RB-ARQ.

V. Conclusion

Reliability-Based Automatic Repeat reQuest (RB-ARQ) was applied to competition data of the P300 speller, it was shown that the discrimination time could be shortened while the accuracy was less decreased than the averaging; thus, the effectiveness of RB-ARQ was confirmed. Also, Reliability-Based Selective-Repeat ARQ (RB-SR-ARQ) was proposed, it

TABLE II
SHAPIRO-WILK TEST

		W-value	P-value
Sub A	P300	0.9978	3.2E-4
	non-P300	0.9988	1.1E-3
Sub B	P300	0.9951	1.7E-8
	non-P300	0.9906	$< 2.2E\text{-}16$

TABLE III
THE DISCRIMINATION TIME AND THE ACCURACY (ONLINE)

	Method	Time[s]	Accuracy[%]	Utility[bps]
Sub 1	Averaging	9.2	70.0	0.22
	RB-ARQ	13.1	95.0	0.35
Sub 2	Averaging	2.7	64.0	0.53
	RB-ARQ	3.0	73.0	0.79
Sub 3	Averaging	9.2	92.0	0.47
	RB-ARQ	6.4	92.0	0.67
Sub 4	Averaging	22.2	85.0	0.16
	RB-ARQ	20.3	89.0	0.20

was shown that the discrimination time of the competition data could be further shortened. In future works, an improvement of RB-SR-ARQ, and a verification of RB-SR-ARQ by online experiment are needed.

ACKNOWLEDGMENT

This work was supported by the Grant-in-Aid for Scientific Research (C) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, Grant number: 22500200.

REFERENCES

[1] G. Dornhege, *Toward brain-computer interfacing*. The MIT Press, 2007.
[2] U. Hoffmann, J. Vesin, T. Ebrahimi, and K. Diserens, "An efficient P300-based brain-computer interface for disabled subjects," *Journal of Neuroscience methods*, vol. 167, no. 1, pp. 115–125, 2008.
[3] R. Scherer, G. Müller, C. Neuper, B. Graimann, and G. Pfurtscheller, "An asynchronously controlled EEG-based virtual keyboard: improvement of the spelling rate," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, p. 979, 2004.
[4] D. McFarland, W. Sarnacki, and J. Wolpaw, "Brain-computer interface (BCI) operation: optimizing information transfer rates," *Biological psychology*, vol. 63, no. 3, pp. 237–251, 2003.
[5] J. Shea, "Reliability-based hybrid ARQ," *Electronics letters*, vol. 38, no. 13, pp. 644–645, 2002.
[6] H. Takahashi, T. Yoshikawa, and T. Furuhashi, "A study on application of reliability based automatic repeat request to brain computer interfaces," *Advances in Neuro-Information Processing*, pp. 1013–1020, 2009.
[7] M. Anagnostou and E. Protonotarios, "Performance analysis of the selective repeat ARQ protocol," *IEEE Transactions on Communications*, vol. 34, no. 2, pp. 127–135, 1986.
[8] E. Sellers, D. Krusienski, D. McFarland, T. Vaughan, and J. Wolpaw, "A P300 event-related potential brain-computer interface (BCI): the effects of matrix size and inter stimulus interval on performance," *Biological Psychology*, vol. 73, no. 3, pp. 242–252, 2006.
[9] http://www.bbci.de/competition/iii/.
[10] A. Rakotomamonjy and V. Guigue, "BCI competition III: Dataset II-ensemble of SVMs for BCI P300 speller," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 3, pp. 1147–1154, 2008.
[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," 2009.
[12] G. Schalk and J. Mellinger, "A Practical Guide to Brain-Computer Interfacing with BCI2000," 2010.
[13] D. Krusienski, E. Sellers, F. Cabestaing, S. Bayoudh, D. McFarland, T. Vaughan, and J. Wolpaw, "A comparison of classification techniques for the P300 Speller," *Journal of neural engineering*, vol. 3, p. 299, 2006.
[14] B. Silverman, *Density estimation for statistics and data analysis*. Chapman & Hall/CRC, 1998.
[15] D. Hoaglin, "Using quantiles to study shape," *Exploring data tables trends, and shapes*, 1985.
[16] S. Shapiro and M. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3-4, p. 591, 1965.
[17] R. Croft and R. Barry, "Removal of ocular artifact from the EEG: a review," *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 30, no. 1, pp. 5–19, 2000.
[18] E. Weisstein, *CRC concise encyclopedia of mathematics*. CRC Pr I Llc, 2003.

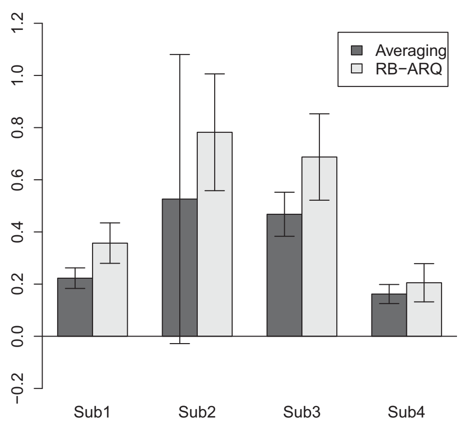
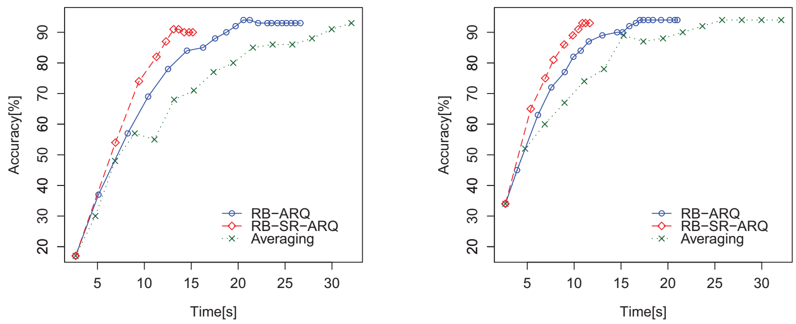
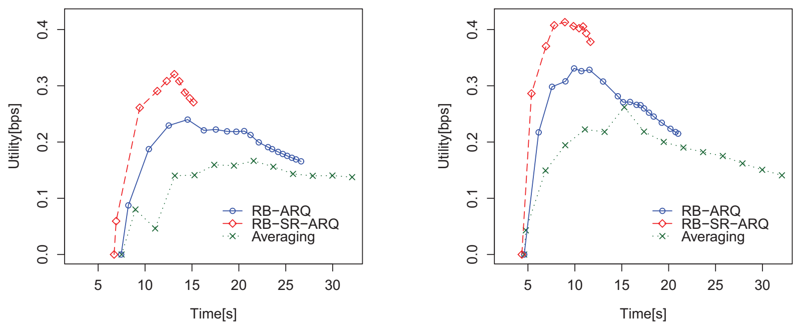


Fig. 7. Comparison of Utility



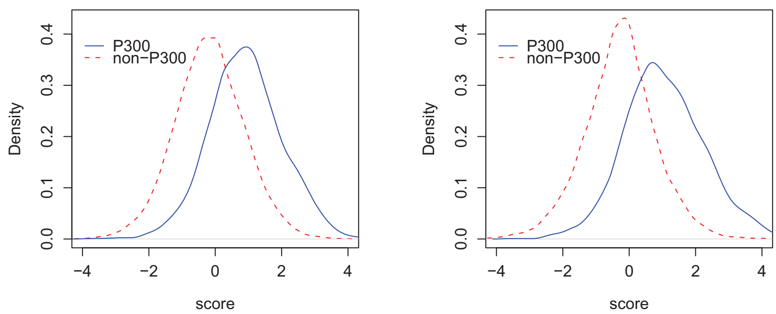
(a) Sub A (b) Sub B

Fig. 3. Accuracy and discrimination time



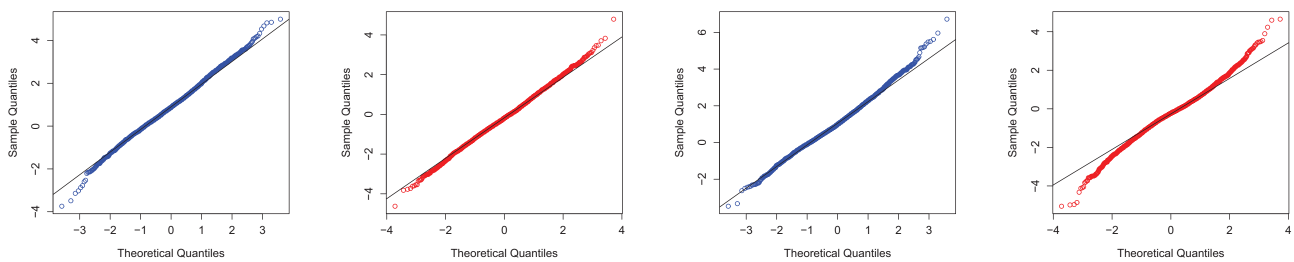
(a) Sub A (b) Sub B

Fig. 4. Utility



(a) Sub A (b) Sub B

Fig. 5. Probability density distribution of discriminant score



(a) P300(Sub A) (b) non-P300(Sub A) (c) P300(Sub B) (d) non-P300(Sub B)

Fig. 6. Q-Q plot