Improvement of spelling speed in P300 speller using transition probability of letters

Eri Samizo
Nagoya University
samizo@cmplx.cse.nagoya-u.ac.jp

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

Takeshi Furuhashi
Nagoya University
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 human's brain, and P300 speller is one of the BCI communication tools, which uses P300 as the feature quantity and allows users to select letters just by thinking. Because of the low signal-to-noise ratio of the P300, signal averaging is often performed to improve the spelling accuracy instead of the degradation of the spelling speed. In texts, there is some variability in the transition probabilities between letters. This paper proposes P300 speller considering the frequencies and the transition probabilities as the priori probabilities. It shows that the spelling speed is improved by the proposed method comparing with the conventional method.**

I. INTRODUCTION

Brain-Computer Interface is the system that controls a computer or a machine based on the information in signals from human's brain[1]. It is expected to be developed as a communication tool for seriously paralyzed patients like those with amyotrophic lateral sclerosis (ALS). Electroencephalogram (EEG) is most likely used for BCIs because it is noninvasive and inexpensive. P300 speller that is first introduced by Farwell et al. is one of the communication tools using P300 as a feature. P300 is one of the event-related potential (ERP) and it is elicited when a stimulus that a user attends to is provided. A user can choose and input letters just by his/her thoughts using P300 speller. It generally uses a letter matrix interface with visual stimulus. Each row and column flashes in random order one by one for a certain times. While thier flashing, the user concentrates the desired letter by counting how many times it flashes. Thereby, P300 is elicited when the row or column that contains the desired letter is flashed. Then the system discriminates the letter that includes P300 most likely as the target one[2].

However, signal-to-noise ratio of the P300 is small. Thus, averaging signals is needed[3][4], which improves the spelling accuracy instead of degrading the spelling speed. Practically, it is needed to input letters correctly in a short time to reduce user's burden. Conventionally, the number of flashing times, i.e., the number of stimuli, is fixed. To reduce the number of stimuli, Reliability-Based Automatic Repeat reQuest (RB-ARQ) has been proposed[5][6]. It is shown that this method can reduce spelling speed keeping spelling accuracy[7].

In RB-ARQ, the prior probability, the likelihood of each letter to be the target before the presentation of stimuli, is set equal for all letters. On the other hand, there is some variability among transition probabilities between letters in texts. In the area of understanding texts or voice recognition, the transition probabilities between letters are used for letter correction or the support of input and recognition[8][9]. In the field of designing keyboard array, they are also utilized to set up high-frequency letters at the place easy to tap[10].

In this paper, we propose P300 speller that considering the transition probabilities between letters as the prior probability in RB-ARQ. The experiments are done by two subjects with Japanese interface of P300 speller and the result shows the improvement of spelling speed by the proposed method comparing with the conventional one.

II. RELIABILITY-BASED AUTOMATIC REPEAT REQUEST

RB-ARQ is a method that presents stimuli randomly and sets the number of stimuli dynamically based on the maximum posterior probability[5][6][7].Suppose $\boldsymbol{x}_t$ denotes a feature vector from EEG data at time t, and let $X_T = \{\boldsymbol{x}_t | t = 1, 2, ..., T\}$ be a set of data at time T, the posterior probability at time T can be calculated as follows:

$$P(k|X_T) = \frac{P(k)\prod_t P(k|\boldsymbol{x}_t)}{\sum_{l\in K} P(l)\prod_t P(l|\boldsymbol{x}_t)} \quad (1)$$

In this equation, let K be a set of candidate letters and $k \in K$. And $P(k)$ is the prior probability that $\boldsymbol{x}$ belong to label k before the stimulus presentation, they are set equally. The posterior probability is obtained by multiplying the prior probability and likelihood. Maximum posterior probability at time T is defined as Eq.(2) using the posterior probability $P(k|X_T)$.

$$\lambda_T = \max_k P(k|X_T) \quad (2)$$

The maximum posterior probability is equivalent to the discriminant accuracy, which can be regarded as the reliability of data. λ is set as the threshold of reliability, and a user keeps thinking until λ_T becomes larger than λ.

III. PROPOSED METHOD

As mentioned above, the prior probability of RB-ARQ is set equal to all letters in the conventional method. This paper proposes the method to consider the transition probability of letters in text as the prior probability. Transition probability is the frequency of a letter in texts after the given preceding

978-1-4673-2743-5/12/$31.00

letter(s), it is given by the occurrence rate of N-gram character in an enormous quantity of text data. It is defined as below with N=1,2,...[11].

$$P(X_i|X_{i-N+1}^{i-1}) = \frac{P(X_{i-N+1}^{i})}{P(X_{i-N+1}^{i-1})} \tag{3}$$

Let X_i^j be a part of string from ith letter to jth letter in the character string $X_1X_2,...X_M$. $P(X_i|X_{i-N+1}^{i-1})$ is the conditional probability that ith letter becomes X_i when a string from $\{i-(N-1)\}$th letter to $(i-1)$th letter is given. $P(X_i^j)$ denotes the probability of the occurrence rate. 1-gram simply represents the probability of the occurrence rate of each letter. When N is 2, it is called Bigram, and it is called Trigram with N = 3. Using this probability, it is expected to improve the performance of inputting text in RB-ARQ. It is thought that the time until the posterior probability exceeds the threshold λ becomes shorter, because the letters with high occurrence rate have high prior probability.

IV. Experiment

A. Data description and preprocessing

This experiment used a recorded dataset which contained EEG data measured by two subjects (Sub A, Sub B) performed the P300 speller. EEG data was recorded with sampling frequency of 100Hz using Polymate AP216 (Digitex lab. co., ltd., Tokyo), from 5electrodes: Fz, Cz ,Pz, O1 and O2, referenced to the linked ears, A1 and A2 (Fig.1). In this experiment, the 7-by-10 letter matrix interface containing Japanese characters shown in Fig.2 was employed.

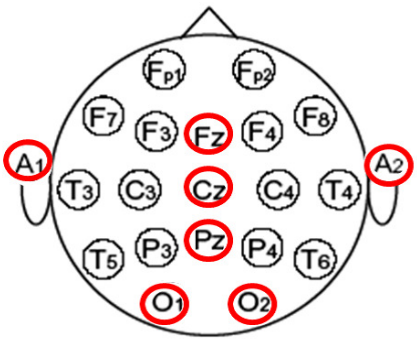


Fig. 1. Used electrodes

In this interface, < 小 > is used when the user wants to input small letters. For example, when the user wants to input <っ>, then he/she needs to select < 小 > before < つ >. And <BS> means backspace which deletes the preceding letter.

An input of one letter consisted of ten sequences (one sequence contained 17 (10 rows and 7 columns) stimuli), each stimulus was intensified for 100ms with 80ms interval between stimuli. Then, the EEG signals were down-sampled to 20Hz, 14 data points corresponding to 0s to 0.65s after each stimulus were extracted. The extracted data were classified

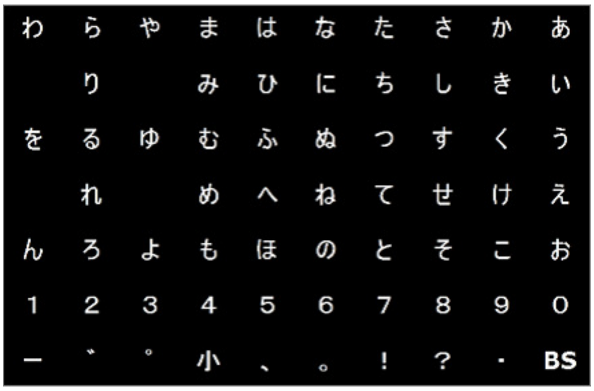


Fig. 2. User-interface

using Linear Discriminant Analysis (LDA) in this experiment. 20 letters were utilized for the learning session. The transition probability of 2-gram described in Ⅲ., which was made based on the web corpus of Japanese[12], was employed as the prior probability in RB-ARQ. In the calculation of Eq.(3), sonant marks and p-sounds were regarded as one character, and < 小 > was made using the number of the appearance of small letters.

B. Experimental settings

The experiment input the following five sentences (actually inputted by *Hiragana*) for 200 times in each test, and averaged the result. The number in () is that of letters in the sentence.

1. しょちゅうおみまいもうしあげます (19)
   Sho Chu O Mi Ma I Mo U Shi A Ge Ma Su
2. きょうもいちにちがんばりましょう (20)
   Kyo U Mo I Chi Ni Chi Ga N Ba Ri Ma Sho U
3. しんねんあけましておめでとうございます (22)
   Shi N Ne N A Ke Ma Shi Te O Me De To U Go Za I Ma Su
4. おたんじょうびおめでとうございます (23)
   O Ta N Zyo U Bi O Me De To U Go Za I Ma Su
5. みなさまのごけんこうをおいのりしております (24)
   Mi Na Sa Ma No Go Ke N Ko U Wo O I No Ri Shi Te O Ri Ma Su

When non-target letter was inputted, that is, discriminant result was wrong, <BS> would be selected at the next target to input full sentence correctly. The average number of stimulus presentation was calculated as changing the threshold λ in RB-ARQ, comparing the proposed method (it is called "Bigram") with the conventional method (the prior probability was set equal, called "Equal").

Though <BS> was counted as the number of the inputted letters, it was not included in "input result." When the threshold was low, however, the amount of wrong discriminant became large and it was difficult to get full correct input result. Thus, the maximum number of inputted letters was set quadruple of the number of letters in each test sentence.

Accordance Rate of input result and Reduction Rate of the number of stimuli, for the performance index in this experiment, are determined below.

$$Accordance\ Rate = \frac{\#\, of\ accorded\ letters\ of\ input\ result\ with\ test\ sentence}{\#\, of\ letters\ in\ test\ sentence} \tag{4}$$

$$Reduction\ Rate = \frac{\#\, of\ stimuli\ in\ Equal - \#\, of\ stimuli\ in\ Bigram}{\#\, of\ stimuli\ in\ Equal} \tag{5}$$

This paper also uses "Utility[13]" defined in Eq.(6) to evaluate accuracy and discrimination time at once.

$$U = \frac{(2P-1)\log_2(N-1)}{d} \tag{6}$$

where N is the number of classes (in this experiment, N=70, 10 rows $\times$ 7 columns), P is the accuracy, and d is the discrimination time. Note that if P<0.5, U=0. Utitlity represents the ITR when the spelling is done perfectly by using <BS> that can erase a non-target letter.

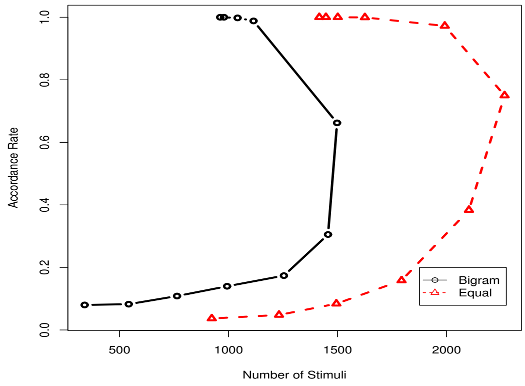
V. Results and Discussions

Fig.3 shows the results when each subject inputted the second test sentence in IV. B. In these figures, horizontal axis means the average of the number of stimuli and vertical axis means the average of Accordance Rate. Fig.4 shows the number of inputted letters of each subject with the second test sentence.

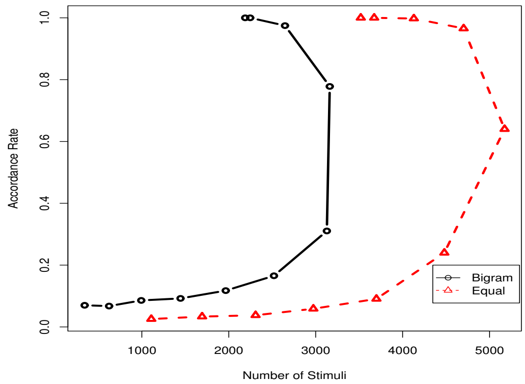
The small number of stimuli represents high spelling speed because the number of stimuli is in proportion to the discriminant (input) time. Higher Accordance Rate is better, essentially, it was expected to be 1.0, because this experiment especially aimed at inputting full correct sentence by using <BS>.

In Fig.3 and Fig.4, the discriminant time for each letter was short but there were a lot of miss discriminant while the threshold was low. Thus, Accordance Rate was still low even if the number of inputted letters reached the maximum number. On the other hand, Accordance Rate increased with the higher threshold and the number of inputted letters became less than the maximum number. However, the total discriminant time became longer due to the increased number of stimuli for each letter. After that, Accordance Rate became 1.0 because the number of miss discriminant became small and it was able to input the sentence accurately by using <BS>. At the same time, the frequency using <BS> was lower and the number of inputted letters became small. Thus, the total input time shortened even if discriminant time for each letter was long.

The discriminant time, the total number of stimuli, with the proposed method was shorter than that with the conventional method at every threshold. The results of other test sentences showed same features. Table I shows the average of the number of stimuli in each subject with the threshold λ=0.9



(a) SubA



(b) SubB

Fig. 3. Results of the Second Test Sentence

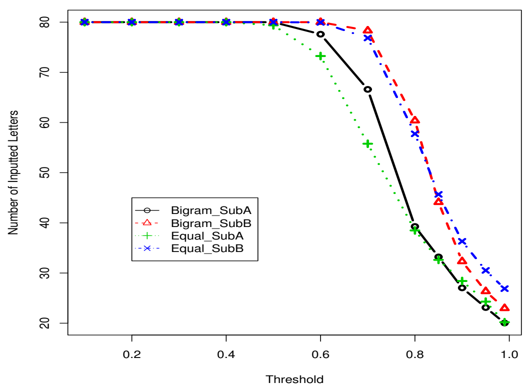


Fig. 4. Number of inputted letters

in RB-ARQ. Fig.5 shows the comparison of Utility between Bigram and Equal with λ=0.9. Table II shows the average of prior probability in Bigram. When the threshold λ was set at 0.9, Accordance Rate was 1.0 in all sentences, and the average prior probability in Equal was 0.0143.

TABLE I
NUMBER OF STIMULI (λ=0.9)

Test Sentence	SubA		SubB	
	Bigram	Equal	Bigram	Equal
1	1049	1385	2702	3427
2	961	1447	2246	3673
3	1162	1606	2902	3999
4	1116	1655	2874	4210
5	1366	1738	3295	4517
Average	1131	1566	2804	3965

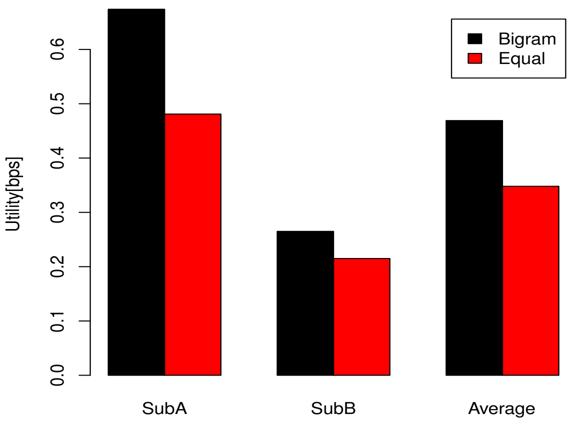


Fig. 5. Comparison of Utility

TABLE II
AVERAGE OF PRIOR PROBABILITY

Test Sectence	1	2	3	4	5	Average
Prior probability	0.125	0.132	0.171	0.167	0.074	0.126

Table I and Fig.5 show that discriminant time (the number of stimuli) of the proposed method was shorter in every test sentence than that of the conventional method and the performance in terms of Utility was higher. There was the significant difference in Utility between Bigram and Equal by ANOVA at the significant level of α = 0.05 (P-value = 0.000181). Reduction Rate of the number of stimuli shown in Eq.(5), was 27.8% for SubA and 29.3% for SubB, respectively. In Table II, it is confirmed that the average of prior probability of inputted letters in Bigram was ten times as high as that in Equal. These results show that using the transition probability of letters as the prior probability could make the total input time short keeping high accordance rate of input result by reducing discriminant time for each letter.

VI. Conclusion

This paper proposed P300 speller that considering the transition probabilities between letters as the prior probability in RB-ARQ. The experiments were done by two subjects with Japanese interface of P300 speller and the result showed the improvement of spelling speed keeping high accordance rate by the proposed method comparing with the conventional one. We will do online experiment by the proposed method and study the consideration of the transition probability of letters based on Trigram.

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, G. P. D.J. McFarland, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
[2] 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.
[3] R. Scherer, G. Muller, C. Neuper, B. Graimann, and G. Pfurtscheller, "An asynchronously controlled eeg-based virtual keyboard: improvement of the spelling rate," *IEEE Trans. on Biomedical Engineering*, vol. 51, no. 6, pp. 979–984, 2004.
[4] U. Hoffmann, J. Vesin, T. Ebrahimi, and K. Diserens, "Methods an efficient p300-based brain-computer interface for disabled subjects," *Journal of Neuroscience*, vol. 167, no. 1, pp. 115–125, 2008.
[5] H. Takahashi, T. Yoshikawa, and T. Furuhashi, "A study on application of reliability based automatic repeat request to brain computer interfaces," in *Proc. 15th Int. Conf. Neural Information Processing*, 2009, pp. 1013–1020.
[6] H.Takahashi, T. Yoshikawa, and T. Furuhashi, "Application of reliability-based automatic repeat request to multi-class classification for brain-computer interfaces," in *Proc. 2009 IEEE Int. Conf. on Fuzzy Systems*, 2009, pp. 1027–1032.
[7] Y. Kaneda, H. Takahashi, T. Yoshikawa, and T. Furuhashi, "Performance evaluation of reliability-based selective repeat automatic repeat request applied to p300 speller by online experiment," *TECHNICAL REPORT OF IEICE*, vol. 110, no. 461, pp. 35–40, 2011.
[8] K. Kukich, "Techniques for automatically correcting words in text," *ACM Comput. Surv*, vol. 24, no. 4, pp. 377–439, 1992.
[9] J. R. Ullmann, "A binary n-gram technique for automatic correction of substitution, deletion, insertion and reversal errors in words," *The Computer Journal*, vol. 20, no. 2, pp. 141–147, 1977.
[10] D. Kawashima, "A keyboard arrangement in romaji input," *The Bulletin of Tokai Women ' s University*, vol. 110, no. 461, pp. 35–40, 1991.
[11] W. B. Cavnar and J. M. Trenkle, "N-gram-based text categorization," in *Proc. of SDAIR-94, 3rd Annual Symposium on Document Analysis and Information Retrieval*, 1994, pp. 161–175.
[12] http://s-yata.jp/corpus/nwc2010/ngrams.
[13] B. D. 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, 2009.