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

A Novel Selective Stimulus Presentation for P300 Speller

Hiromu Takahashi, *Student Member, IEEE,* Tomohiro Yoshikawa, and Takeshi Furuhashi

Abstract— **The P300 speller is one of the brain-computer interfaces, allowing users to spell letters just by thoughts. Due to the low signal-to-noise ratio of the P300, however, stimuli are repeatedly presented so that EEG signals can be averaged, which improves the accuracy but degrades the speed. The authors have proposed to discontinue the stimulus presentation adaptively to the P300 response and have shown its superiority in the performance over the standard way that presents a pre-fixed number of stimuli. In addition to this adaptive stimulus termination, this paper proposes to select stimuli to be presented to avoid presenting redundant stimuli. Both off-line and on-line experiments show that the proposed method is more effective than our conventional method.**

I. Introduction

P300 speller is one of the brain-computer interfaces for communication, which typically employs a 6-by-6 letter matrix interface [1]. Each row and column is randomly flashed one by one, then the letter whose flashes have most likely elicited an event-related potential called P300 is determined as the user's desired letter. Due to the low signal-to-noise ratio of the P300, however, many flashes are necessary so that signals can be averaged, resulting in an improvement of the spelling accuracy in exchange for the spelling speed. Thus, our paper aims at improving the accuracy with a small loss of the speed.

One approach tries to improve the P300 classification by a powerful classifier or by feature selection [2]. Another is to modify the stimulus design from the row/column design [3], [4]. By contrast, our approach is to manipulate the sequence of the flashings; hence, it is easy to implement and could co-exist with other approaches. The authors have proposed a method that discontinues the flashings adaptively to the P300 responses, and have shown its superiority over the standard way where the number of stimuli is pre-fixed [5]. In addition, the present paper proposes a method that intensifies only selected rows and columns. Both off-line and on-line experiments show that the proposed method is more effective than our conventional method.

After some mathematical notations and our conventional method, Section II proposes an improved method: how stimuli are selected and why this is appropriate. Sections III and IV are experiments to evaluate the proposed method.

II. Methods

A. P300 speller

Let $\mathcal{L}$ be a set of selectable letters, and $\mathcal{R}$ and $\mathcal{C}$ be sets of rows and columns, respectively, then elements in $\mathcal{L}$ can also be represented by those in $\mathcal{R} \times \mathcal{C}$. In Fig. 1, for example, $\mathcal{L} = \{\mathrm{A}, \mathrm{B}, \ldots, \mathrm{BS}\}$, $\mathcal{R}$ and $\mathcal{C}$ are $\{1, 2, \ldots, 6\}$ and $\{7, 8, \ldots, 12\}$, respectively, and "T" can be represented by (4,8). Also suppose $\boldsymbol{x}_n$ is the nth EEG epoch corresponding to the nth stimulus denoted by $s_n \in \mathcal{R} \cup \mathcal{C}$ and it is associated with a P300 label $t_n \in \{0, 1\}$. Let $\theta_L = (\theta_R, \theta_C)$ be the user's target letter and suppose $\theta_L =$"T", then only the EEGs corresponding to flashes of the fourth row or the second column have a P300 label of 1. Let a sequence denote intensifications of six different rows and six different columns, several sequences are performed per letter.

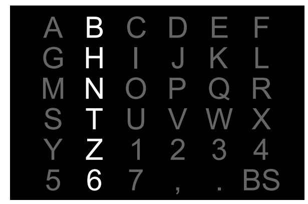

Fig. 1. User-interface while the second column is intensified.

B. Conventional method

Automatic repeat request (ARQ) is an error control scheme in the field of data transmission, in which the receiver asks the sender for re-transmission on error detection [6]. The authors have proposed reliability-based automatic repeat request (RB-ARQ) for BCIs, which employs the maximum posterior probability as the repeat criterion [5].

Suppose $\mathcal{X}_n$ denotes a set of EEG epochs observed up to the nth stimulus, i.e., $\mathcal{X}_n = \{\boldsymbol{x}_i\}_{i=1}^n$, then the maximum posterior probability given $\mathcal{X}_n$ can be represented as

$$\lambda_n = \max_{l \in \mathcal{L}} P(\theta_L = l | \mathcal{X}_n). \quad (1)$$

The details of how this is calculated can be found in [5]. Bayes classifier [7] selects such a letter $\hat{\theta}_L$ that

$$\hat{\theta}_L = \mathrm{argmax}_{l \in \mathcal{L}} P(\theta_L = l | \mathcal{X}_n), \quad (2)$$

where a hat means an estimated desired letter. The maximum posterior probability equals the expected classification accuracy; thus, it can be seen as the *reliability* of the classification result. RB-ARQ utilizes it as the repeat criterion, so as to control the expected accuracy rather than the number of sequences. Specifically, the stimulus presentation continues until λ_n becomes greater than an arbitrary threshold $\lambda \in [0\ 1]$, i.e., $\lambda_n > \lambda$, and considering practicality, the maximum number of stimuli is limited to $N_{\max}$. It is worth noting that the threshold λ is the lower bound of the expected accuracy.

Department of Computational Science and Engineering, Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 4648603, Japan takahashi@cmplx.cse.nagoya-u.ac.jp

C. Proposed method

The target accuracy can be set by the threshold λ; thus, the number of stimuli to reach the threshold needs to be reduced in order to improve RB-ARQ. Thus, we propose a method to choose rows and columns to be intensified so that the *reliability* is effectively increased. This paper terms this method reliability-based selective repeat automatic repeat request (RB-SR-ARQ) after selective repeat ARQ [6]. The question is how to choose stimuli to be presented.

Let $\lambda^R_{n(k)}$ be the kth largest posterior probability given $\mathcal{X}_n$ regarding rows and $\hat{r}_{n(k)}$ be the corresponding row, i.e., $\lambda^R_{n(k)} = P(\theta_R = \hat{r}_{n(k)}|\mathcal{X}_n)$ and $\lambda^R_{n(i)} \geq \lambda^R_{n(j)}$ $(i \leq j)$, and let analogous definitions hold for columns. Then, the *reliabilty* can be decomposed as follows:

$$\lambda_n = \lambda^R_{n(1)} \times \lambda^C_{n(1)} \tag{3}$$

$$= \max_{r\in\mathcal{R}} P(\theta_R = r|\mathcal{X}_n) \times \max_{c\in\mathcal{C}} P(\theta_C = c|\mathcal{X}_n). \tag{4}$$

Hereafter, events $\theta_L = l$ and $(\theta_R, \theta_L) = (r, c)$ are denoted by just l and (r, c) if not confusing. Suppose $n-1$ EEGs are already obtained, then the expected value of λ_n given $s_n \in \mathcal{R}$ can be written as

$$E[\lambda_n|s_n \in \mathcal{R}] = \lambda^C_{n-1(1)} E[\lambda^R_{n(1)}|s_n \in \mathcal{R}]. \tag{5}$$

By definition of the expected value, we get

$$E[\lambda^R_{n(1)}|s_n \in \mathcal{R}] = E[\max_{r\in\mathcal{R}} P(r|\mathcal{X}_{n-1}, \boldsymbol{x}_n)] \tag{6}$$

$$= \int \max_{r\in\mathcal{R}}\{P(r|\mathcal{X}_{n-1}, \boldsymbol{x}_n)p(\boldsymbol{x}_n|\mathcal{X}_{n-1})\}d\boldsymbol{x}_n, \tag{7}$$

where $p(\cdot)$ denotes a probability density. Note the probability density of a random variable $\boldsymbol{x}_n$ is conditioned on $\mathcal{X}_{n-1}$. Using the Bayes theorem and assuming each $\boldsymbol{x}_i$ is conditionally independent of the others given $\theta_R = r$, we get

$$P(r|\mathcal{X}_{n-1}, \boldsymbol{x}_n)p(\boldsymbol{x}_n|\mathcal{X}_{n-1}) = P(r|\mathcal{X}_{n-1})p(\boldsymbol{x}_n|r). \tag{8}$$

When we take the maximum of (8), we obtain the followings:

$$\begin{aligned} &\max_{r\in\mathcal{R}}\{P(\theta_R = r|\mathcal{X}_{n-1})p(\boldsymbol{x}_n|\theta_R = r)\} \\ &= \max[P(\theta_R = s_n|\mathcal{X}_{n-1})p(\boldsymbol{x}_n|\theta_R = s_n), \\ &\quad \max_{r'\in\mathcal{R}\setminus s_n}\{P(\theta_R = r'|\mathcal{X}_{n-1})p(\boldsymbol{x}_n|\theta_R = r')\}] \qquad (9) \\ &= \max[P(\theta_R = s_n|\mathcal{X}_{n-1})p(\boldsymbol{x}_n|t_n = 1), \\ &\quad \max_{r'\in\mathcal{R}\setminus s_n}\{P(\theta_R = r'|\mathcal{X}_{n-1})\}p(\boldsymbol{x}_n|t_n = 0)]. \qquad (10) \end{aligned}$$

If $\lambda^R_{n-1(k)} = P(\theta_R = s_n|\mathcal{X}_{n-1})$, i.e., the row to be intensified as the nth stimulus is the kth most likely the target row, (10) becomes as follows:

$$\begin{aligned} &\max_{r\in\mathcal{R}}\{P(r|\mathcal{X}_{n-1})p(\boldsymbol{x}_n|r)\} \qquad (11) \\ &= \begin{cases} \max[\lambda^R_{n-1(1)} f_1(\boldsymbol{x}_n), \lambda^R_{n-1(2)} f_0(\boldsymbol{x}_n)] & \text{if } k = 1, \\ \max[\lambda^R_{n-1(k)} f_1(\boldsymbol{x}_n), \lambda^R_{n-1(1)} f_0(\boldsymbol{x}_n)] & \text{otherwise,} \end{cases} \end{aligned}$$

where $f_{t_n}(\boldsymbol{x}_n) = p(\boldsymbol{x}_n|t_n)$. Substituting (8) and (11) into (7) and taking $\lambda^R_{n-1(1)}$ from the integral, we obtain

$$\begin{aligned} &E[\lambda_n|s_n \in \mathcal{R}] \qquad (12) \\ &= \begin{cases} \lambda_{n-1} \int \max[f_1(\boldsymbol{x}_n), \frac{\lambda^R_{n-1(2)}}{\lambda^R_{n-1(1)}} f_0(\boldsymbol{x}_n)]d\boldsymbol{x}_n & \text{if } k = 1, \\ \lambda_{n-1} \int \max[\frac{\lambda^R_{n-1(k)}}{\lambda^R_{n-1(1)}} f_1(\boldsymbol{x}_n), f_0(\boldsymbol{x}_n)]d\boldsymbol{x}_n & \text{otherwise.} \end{cases} \end{aligned}$$

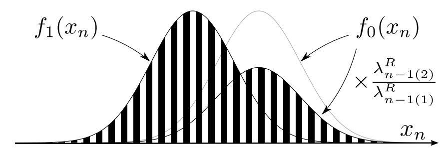

Fig. 2. The shaded area is the integral part in (12) if $k = 1$.

Since $\boldsymbol{x}_n$ is often a high-dimensional vector, it would be reasonable to approximate (12) by integrating only on the linear discriminant coordinate $x_n = \boldsymbol{w}'\boldsymbol{x}_n$ ($\boldsymbol{w}'$ is a transpose of the linear discriminant vector) [7]. Meanwhile, it has been reported that $f_0(x_n)$ and $f_1(x_n)$ are approximately Gaussian distributions with an equal variance (see Fig. 2) [5]; thus, they are exchangeable in (12), and then we get the following:

$$\begin{aligned} &E[\lambda_n|s_n = \hat{r}_{n-1(1)}] = E[\lambda_n|s_n = \hat{r}_{n-1(2)}] \geq \\ &E[\lambda_n|s_n = \hat{r}_{n-1(3)}] \geq \cdots \geq E[\lambda_n|s_n = \hat{r}_{n-1(|\mathcal{R}|)}], \end{aligned} \tag{13}$$

where $|\mathcal{R}|$ denotes the number of elements in $\mathcal{R}$. This suggests that flashing either the best or the second best most likely target row is expected to increase the *reliability* most effectively; accordingly, intensifications of the best two rows and two columns seem most effective. However, RB-SR-ARQ intensifies the best three rows and three columns, to ensure the target stimuli are rare events to elicit the P300. Due to this selective flashing, in this paper, "a sequence" is redefined to be a series of flashing rows and columns selected at the same timing.

Having said that, this selective stimulus presentation starts only after the two standard sequences, because neither rows nor columns can be ordered at the beginning and because the influence of outliers is preferred to be reduced if contaminated in the first several epochs. The whole procedure is described in Algorithm 1. It should be noted that there exists a duration between the stimulus onset of s_n and the completion of recording x_n; thus, a few epochs related to the last a few stimuli in the preceding sequence are not used for the stimulus selection. Also note that while RB-ARQ conducts the threshold decision per sequence, RB-SR-ARQ does it per stimulus, which makes the threshold decision more flexible and could further improve the performance.

III. Off-line experiment

The proposed method was compared with the conventional method, using bootstrapped samples from the dataset II in the BCI Competition III [8], which involved two subjects: Sub A and B. The bootstrap is useful for estimation of parameters and also of prediction error [9]; in fact, it has been used to estimate prediction error in the P300 speller [1].

The P300 amplitude is inversely proportional to the target stimulus probability [10]; hence, it should be smaller in the selective flashing sequence where the probability is one-third at most than that in the standard sequence where it is one-sixth. According to [10], the number of preceding non-targets is a decisive factor for the P300 amplitude. Thus, the bootstrap was performed considering not only P300 label t_n

Algorithm 1 The proposed method (RB-SR-ARQ)

Require: $0 \leq \lambda \leq 1$ and $N_{\max} \in \mathbb{N}$

$n \leftarrow 0$, $\mathcal{X}_0 \leftarrow \{\}$, $\mathcal{S} \leftarrow \mathcal{R} \cup \mathcal{C}$

repeat

$n \leftarrow n+1$

randomly take $s_n (\neq s_{n-1}$ if $n \geq 2)$ from $\mathcal{S}$

present s_n and sample $\boldsymbol{x}_n$

$\mathcal{X}_n \leftarrow \mathcal{X}_{n-1} \cup \{\boldsymbol{x}_n\}$

if $n = 12$ **then**

$\mathcal{S} \leftarrow \mathcal{R} \cup \mathcal{C}$

else if $\mathcal{S} = \{\}$ **then**

$\mathcal{S} \leftarrow \{\hat{r}_{n(1)}, \hat{r}_{n(2)}, \hat{r}_{n(3)}, \hat{c}_{n(1)}, \hat{c}_{n(2)}, \hat{c}_{n(3)}\}$

end if

until $n = N_{\max} \vee \lambda_n > \lambda$

return the estimated target letter $\hat{\theta}_L$ and n

TABLE I

THE NUMBER OF PRECEDING NON-TARGETS WHEN θ_L = "T"

n	1	2	3	4	5	6	7	8	9	10	11	12
s_n	6	2	8	5	3	11	10	4	7	9	1	12
t_n	0	0	1	0	0	0	0	1	0	0	0	0
z_n	0	1	2	0	1	2	3	4	0	1	2	3

but also the number of preceding non-targets z_n when $t_n = 1$ (see an example in Table I).

Each subject performed spelling 185 letters with 15 standard sequences; and the dataset was divided into two: first 85 letters, or 15300 epochs, for training a classifier, and the remaining 100 letters, or 18000 epochs, for test and they were bootstrapped. Feature vectors of dimension 896 (64 electrodes $\times$ 14 time points) were extracted, and then step-wise linear discriminant analysis (SWLDA) [1], [2] performed feature selection and calculated posterior probabilities. Algorithm 1 with $N_{\max} = 180$ was repeated for 1000 times, then the accuracy p and the average number of stimuli $\bar{N}$ were calculated as follows:

$$p = \frac{\sum_{j=1}^{1000} I(\theta_{Lj}, \hat{\theta}_{Lj})}{1000}, \quad \bar{N} = \frac{\sum_{j=1}^{1000} n_j}{1000}, \qquad (14)$$

where $I(a, b)$ gives 1 if $a = b$ and gives 0 otherwise, and θ_{Lj}, $\hat{\theta}_{Lj}$, and n_j are the jth randomly assigned target letter, the estimated letter, and the number of stimuli, respectively. This paper evaluates methods based on the *Utility* [11] defined as

$$U = \frac{2p-1}{d} \log_2(|\mathcal{L}| - 1), \qquad (15)$$

if $p > 0.5$, $U = 0$ otherwise. Note that d denotes the average duration per letter, which is proportional to the average number of stimuli. The *Utility* represents the information transfer rate when letters are spelled perfectly using the backspace ("BS" in Fig. 1) if needed.

Figure 3 compares the *Utility* obtained by both methods, varying the threshold $\lambda = \tanh(a)$ $(a = 0, 0.3, \ldots, 4.8)$. By doing this, the average duration d is approximately proportional to a (see the details in [5]). It shows the proposed method achieves about 10 % improvement if the threshold is appropriately determined to obtain the maximum *Utility*. Also it was found by investigating four possible combinations (2 ARQs $\times$ 2 threshold decisions) that this improvement is contributed mainly by the selective stimulus presentation not by the frequent threshold decision.

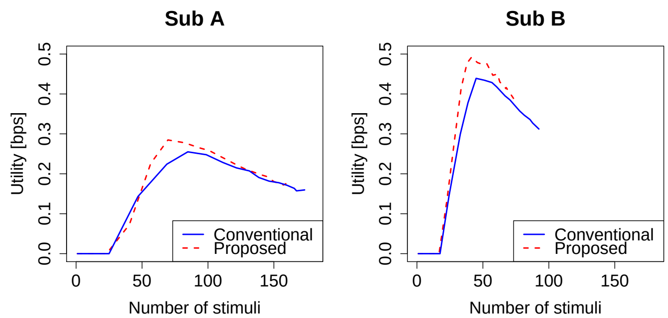

Fig. 3. Comparison of RB-ARQ and RB-SR-ARQ

IV. On-line Experiment

This experiment compared RB-ARQ and RB-SR-ARQ in actual use. In this study, English letters spelled using the P300 speller were transferred to another application, in which they were converted to Japanese letters. In addition to the comparison between RB-ARQ and RB-SR-ARQ, this study intended to compare a required number of English letters to type the same Japanese sentence with a predictive conversion and without it. However, the latter purpose is not the focus of the present paper; thus, the details are omitted.

A. Experimental settings

Six volunteer students (aged 22-26 years; one female and five males), participated in this experiment. They sat facing two LCD displays, one for the letter matrix as Fig. 1 and the other for Japanese input. Their 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). After a preprocessing similar to the off-line experiment, a feature vector of 65 dimensions (5 electrodes $\times$13 time points) was obtained.

Each subject performed a learning session, in which he or she was required to try to spell 25 English letters with ten standard sequences. Then a subject-specific SWLDA classifier was trained using this dataset. The threshold λ was determined so as to maximize the *Utility* by a ten-fold cross validation assuming to use RB-ARQ, i.e., it was optimized for RB-ARQ. Then each subject performed eight test sessions, in each of which he or she was required to correctly spell one of four Japanese sentences prepared for this experiment. Four conditions (2 ARQs $\times$ 2 conversion methods) were assigned to eight sessions in a randomized block design (a block is the four conditions), i.e., each sentence was spelled twice. The maximum number of stimuli was set to be 120, i.e., $N_{\max} = 120$ for both ARQs.

B. Results and Discussions

Table II shows the threshold common to both ARQs, the overall number of spelled English letters in four sessions

TABLE II

THRESHOLD, NUMBER OF SPELLED ENGLISH LETTERS, ACCURACY, AVERAGE NUMBER OF STIMULI, AND UTILITY.

Sub	Threshold λ	# English letters		Accuracy [%]		# stimuli (# sequences)		Utility [bps]	
		RB-ARQ	RB-SR-ARQ	RB-ARQ	RB-SR-ARQ	RB-ARQ	RB-SR-ARQ	RB-ARQ	RB-SR-ARQ
1	0.87	80	122	97.5	82.8	22.5 (1.9)	17.3 (1.4)	1.13	0.98
2	0.93	131	156	81.7	82.7	88.7 (7.4)	67.9 (9.3)	0.21	0.28
3	0.97	75	102	98.7	91.2	42.2 (3.5)	30.9 (3.1)	0.65	0.74
4	0.77	259	291	64.5	66.7	51.8 (4.3)	31.7 (3.3)	0.16	0.29
5	0.77	145	124	80.0	80.6	27.7 (2.3)	21.7 (1.8)	0.59	0.75
6	0.69	171	199	78.9	72.4	32.8 (2.7)	21.0 (1.7)	0.49	0.57

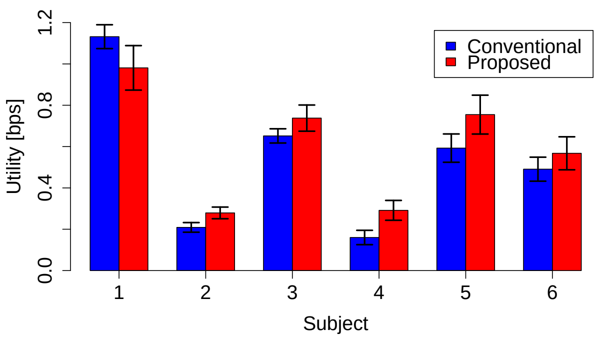

Fig. 4. Utility with an error bar of the standard deviation

(two with predictive conversion and the other two without it), the accuracy, the average number of stimuli (that of sequences), and the *Utility*. Additionally, Fig. 4 compares the *Utility* with an error bar of the standard deviation estimated by the bootstrap (1000 replicates) [9]. It is worth noting that the accuracy p and the average number of stimuli $\bar{N}$ were calculated by (14) with $j = 80$ for the Sub 1's RB-ARQ and they are related to the output of the P300 speller; thus, they should be independent from the conversion method. It is noticeable that the number of spelled English letters varied from subject to subject; in fact, it varied from session to session. It was because each Japanese sentence was correctly spelled using the backspace when necessary.

Table II tells the average number of sequences in RB-SR-ARQ sessions of Sub 1, 5, and 6 were less than two. This means that both ARQs were identical except for the frequency of the threshold decision, which does not make a significant difference according to the off-line experiment. Indeed, Fig. 4 does not clearly show the difference between the two methods in their *Utility*. On the other hand, we can find the Utility of RB-SR-ARQ was greater than that of RB-ARQ in all Sub 2, 3, and 4. Moreover, there was a statistically significant difference between the two ARQs at the significance level of 5 % ($p = 0.035$ by the paired t-test). Since the threshold λ was optimized for RB-ARQ, the performance of the proposed method could be improved if it is appropriately determined.

Since the P300 amplitude is determined by the number of preceding non-targets, presentation of only the best three rows/columns in a succeeding sequence might not be the best strategy, i.e., presentation of the best four rows/columns might be better. Therefore, the expected value of the *reliability* like the one in Section II-C needs to be more precisely analyzed, e.g., by taking the preceding non-targets into consideration, for further improvement.

V. CONCLUSION

This paper proposed reliability-based selective repeat ARQ (RB-SR-ARQ) for the P300 speller, which only flashes selected rows and columns. Both the off-line and on-line experiments show that the proposed method outperformed our conventional method. In future work, the proposed method will be applied for a larger letter matrix configuration, where many more of redundant flashes should be performed.

ACKNOWLEDGEMENTS

This work was supported by the Grant-in-Aid for JSPS Fellows No. 22-8417 and the Grant-in-Aid for Scientific Research (C) No. 22500200 from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

REFERENCES

[1] E. Donchin, K. M. Spencer, and R. Wijesinghe, “The mental prosthesis: Assessing the speed of a P300-Based Brain-Computer interface,” *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 8, no. 2, 2000.

[2] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, “Toward enhanced p300 speller performance,” *J. Neuroscience Methods*, vol. 167, no. 1, pp. 15–21, 2008.

[3] G. Townsend, B. K. LaPallo, C. B. Boulay, D. J. Krusienski, G. E. Frye, C. K. Hauser, N. E. Schwartz, T. M. Vaughan, J. R. Wolpaw, and E. W. Sellers, “A novel p300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns,” *Clin. Neurophysiol.*, vol. 121, no. 7, pp. 1109–1120, 2010.

[4] J. Hill, J. Farquhar, S. Martens, F. Bießmann, and B. Schölkopf, “Effects of stimulus type and of error-correcting code design on BCI speller performance,” in *Advances in neural information processing systems*, 2009, vol. 21, pp. 665–672.

[5] H. Takahashi, T. Yoshikawa, and T. Furuhashi, “Error control for performance improvement of Brain-Computer interface: Reliability-Based automatic repeat request,” *IEICE Trans. Inf. & Syst.*, vol. E94-D, no. 6, pp. 1243–1252, 2011.

[6] J. B. Anderson and S. Mohan, *Source and channel coding: an algorithmic approach*. Springer, 1991.

[7] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.

[8] B. Blankertz, K. R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pfurtscheller, J. R. Millán, M. Schröder, and N. Birbaumer, “The BCI competition III: validating alternative approaches to actual BCI problems,” *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 14, no. 2, pp. 153–159, 2006.

[9] B. Efron and R. Tibshirani, *An introduction to the bootstrap*. Chapman & Hall, 1993.

[10] C. J. Gonsalvez and J. Polich, “P300 amplitude is determined by Target-to-Target interval,” *Psychophysiology*, vol. 39, no. 03, pp. 388–396, 2002.

[11] 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 Trans. Neural Syst. Rehab. Eng.*, vol. 18, no. 1, pp. 20–28, 2009.