2004 IEEE International Workshop on Biomedical Circuits & Systems

HIGH PERFORMANCE P300 SPELLER FOR BRAIN-COMPUTER INTERFACE

Cuntai Guan, Manoj Thulasidas, Jiankang Wu

Institute for Infocomm Research, Singapore 119613
{ctguan,manoj,jiankang}@i2r.a-star.edu.sg

ABSTRACT

P300 speller is a communication tool with which one can input texts or commands to a computer by thought. The amplitude of the P300 evoked potential is inversely proportional to the probability of infrequent or task-related stimulus. In existing P300 spellers, rows and columns of a matrix are intensified successively and randomly, resulting in a stimulus frequency of 1/N (N is the number of rows or columns of the matrix). We propose a new paradigm to display each single character randomly and individually (therefore reducing the stimulus frequency to 1/(N*N)). On-line experiments showed that this new speller significantly improved the performance. Specifically, the new speller can reduce character classification error rate by up to 80% or double the information transfer rate compared to the existing P300 spellers.

1. INTRODUCTION

Brain-Computer Interface (BCI) allows people to interact directly with a computer using their brain signals. A comprehensive review on BCI can be found in [1]. The implication of BCI is it is probably the only resort for those people, who do not have control of their peripheral nerve and muscles, to access computer and communicate with outside world. Among various BCI systems, electroencephalography (EEG) is still the most common signal because of its non-invasive nature. Various signals can be extracted from EEG to develop BCI systems, for example, slow cortical potential [2], μ and β rhythms [3], synchronization and desynchronization evoked by motor imagery [4], P300 evoked potential [5], static-state visual evoked potentials [6], etc.

P300 potential is one of the well studied and most stable potentials. It is elicited by an oddball paradigm. That is, an anticipatory event creates a measurable potential difference at the central or parietal sites of EEG measurements. This positive potential typically occurs around 300 milliseconds after the event occurs.

Farwell and Donchin first proposed a P300 speller (FD-Speller) [5], in which a matrix of 6x6 cells is displayed to represent 26 letters and a few commands. The rows and columns of the matrix were successively and randomly intensified such that when the corresponding row or column with target cell was highlighted, a P300 signal was elicited. Donchin etc [7] further improved the information transfer rate by using a bootstrapping approach. Meinicke etc [8] first reported data-driven method for P300 speller. They achieved the highest information transfer rate up to the time of publication. Xu etc [10] reported good results on using ICA-based subspace projections for P300 speller. Brendan etc [11] studied the relationship between accuracy and various size of speller matrix. More recently, Mellinger etc [12] reported P300 speller can be used for amyotrophic lateral sclerosis (ALS) patients.

However, all these studies were based on FD-Speller. As we know, the amplitude of the P300 evoked potential is inversely proportional to the stimulus probability. In FD-Speller, the stimulus frequency for target cell is only 1/N (N is the number of rows or columns of the matrix). In this paper, we propose a new P300 speller paradigm. In this speller, each individual characters are displayed randomly, therefore reducing the stimulus frequency to 1/(N*N). This should elicit a higher P300. Our on-line experiments showed that a P300 speller based on this single display paradigm (called SD-Speller) significantly improved the character classification accuracy. Specifically, in our comparison experiments, six subjects participated in tests for both FD-Speller and SD-Speller. SD-Speller reduced character classification error rate by up to 80% compared to FD-Speller. Furthermore, SD-speller is more flexible in allowing us to design any type of cell layout, instead of square matrix in FD-Speller. Subjects also found SD-Speller caused less fatigue than FD-Speller during experiments.

In this paper, we evaluated various processing methods to optimize the speller performance. Machine learning method was adopted to construct classifier. On top of this, we evaluated two methods for on-line artifact removal as well as noise reduction based on principal component analysis (PCA). A new dynamic component was proposed to enhance the performance. The combinations of all these processing methods were evaluated and the best processing methods were selected. The character classification error rate was reduced by 23% after the algorithm optimization.

0-7803-8665-5/04/$20.00 ©2004 IEEE

2. PROBLEM DEFINITION

Given an EEG time sequence starting at time t, $x_t = [x_t(1)^T, \cdots, x_t(L)^T]^T$, where time t is synchronized with a visual stimulus, L is the number of channels, T denotes the matrix transpose, P300 detection can be formulated as the problem of statistical hypothesis testing. x_t can be classified into two classes:

H0: $x_t \in y_0$, if x_t contains P300 signal

H1: $x_t \in y_1$, if x_t does not contain P300 signal.

where null hypothesis y_0 is represented by EEG containing P300 signal, and the alternative hypothesis y_1 is represented by EEG not containing P300 signal.

An optimal hypothesis test is performed by the following likelihood ratio:

$$\gamma(x_t) = P(x_t / y_0)/P(x_t / y_1) \qquad (1)$$

where $P(x_t / y_0)$ is the likelihood of x_t containing P300 signal and $P(x_t / y_1)$ the likelihood not containing P300 signal. The hypothesis test is performed by comparing the likelihood ration $\gamma(x_t)$ to a predefined threshold.

In the context of P300 speller, the EEG sequence is synchronized with the intensification of target characters on which the user focuses. After each round of intensifications, one can determine the most probable character by which the highest P300 signals are evoked. In single trial case, the classification result is the character index i which satisfies:

$$\hat{i} = \arg\max_i \left\{ f(x_t^{(i)}) \right\}, i = 1, \cdots, N \qquad (2)$$

where N is the number of characters, $f(x_t)$ is a probability representation of x_t. Usually, several trials of data are used to obtain an ensemble average so as to enhance the classification accuracy. When K trials are available, we have

$$\hat{i} = \arg\max_i \left\{ f(x_{t_1}^{(i)}, x_{t_2}^{(i)}, \cdots, x_{t_K}^{(i)}) \right\}, i = 1, \cdots, N \qquad (3)$$

Assume all trials are independent, we obtain

$$f(x_{t_1}^{(i)}, x_{t_2}^{(i)}, \cdots, x_{t_K}^{(i)}) = \prod_{k=1}^{K} f(x_{t_k}^{(i)}) \qquad (4)$$

Here we use support vector machine (SVM) for classification, so we can obtain the following score

$$K(x_{t_k}^{(i)}, x) = \sum_{m=1}^{M} y_i \alpha_i k(x_{t_k}^{(i)}, x_m) + b \qquad (5)$$

where $k(\bullet)$ is the kernel function. We use Gaussian kernel in this study. We can then convert the margin score into a probability representation as follows,

$$f(x_{t_k}^{(i)}) = \frac{1}{1 + e^{-\mu[K(x_{t_k}^{(i)}, x) + \beta]}} \qquad (6)$$

In the case of Farwell-Donchin speller, rows and columns are intensified simultaneously. So the formulas in (2) and (3) are used to find the indices of rows and columns respectively. The character indices are then determined by the row and column indices.

3. P300 SPELLER

3.1 Farwell-Donchin Paradigm (FD-Speller)

In this speller, a subject was presented with a six by six matrix of characters (see Figure 1). The subject's task was to focus his attention on alphanumeric symbols in the matrix, one at a time. Rows and columns of this matrix were successively and randomly intensified for 100 milliseconds (ms), followed by 80 ms of non-intensification. Two out of twelve intensifications of rows and columns contained the desired character (i.e., one particular row and one particular column). EEG signal was sampled at 250Hz. 25 channels around central and parietal scalp were selected from all 64 channels we used to acquire EEG data.

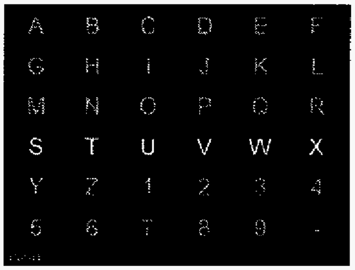

Figure 1. FD-Speller

3.2 Single Display Paradigm (SD-Speller)

Motivated by the fact that P300 is inversely proportional to the rare stimulus probability, we propose the following speller. Figure 2 shows the layout of such a speller, called SD-Speller in this paper. In this speller, all letters and digits are first displayed (left picture in Figure 2). When the speller starts, each single character is randomly flipped for 60ms (right picture in Figure 2).

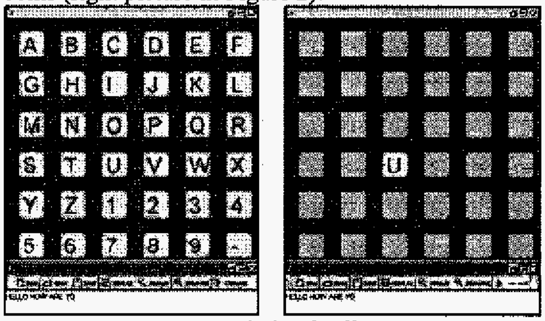

Figure 2. SD-Speller

As we want to compare this speller with FD-Speller, they both should work at the same speed. In other words, they should have the same time duration for one complete round of display of all characters. In FD-Speller, one round of display takes 180x12msec. Here in SD-Speller, it is 60x36msec. Under this timing configuration, we can compare their accuracy.

4. COMPARISON OF FD-SPELLER AND SD-SPELLER

We have conducted comparative experiments to show how SD-Speller outperformed FD-Speller.

In this section, we shall first briefly explain the platform and signal processing methods we were using in the study, and then present experiments settings and experimental results.

4.1 Signal Processing

Two on-line artefact removal methods were evaluated [9]. The first was a simple linear superposition model of the measured reference signals *u(n)* and the real EEG $\tilde{x}(n)$.

$$x(n) = \sum_{i=1}^{S} b_i u_i(n) + \tilde{x}(n) \qquad (7)$$

where S denotes the number of sites at which reference signals (EOG in our experiment) were measured. A further improved method is a difference model, which relaxes the constraint that EEG signal is uncorrelated with reference signals. The difference model is expressed as follows,

$$x(n) = y(n-1) + \sum_{i=1}^{N} b_i(u_i(n) - u(n-1)) + \tilde{x}(n) - \tilde{x}(n-1) \qquad (8)$$

PCA is a common method used to remove "noises". Here we applied PCA on the 25 selected channels and transformed them into 18 channels.

It is known that for different subjects, their P300 signals have different amplitude and latency. To enhance the differentiation between P300 signal and non-P300 signal, we propose to expand the data vector to include a "dynamic" component as follows:

$$x = [x(1)^T, \cdots, x(L)^T, \Delta x(1)^T, \cdots, \Delta x(L)^T]^T \qquad (9)$$

where $x_s = [x(1)^T, \cdots, x(L)^T]^T$ is the original EEG data vector, $\Delta x = [\Delta x(1)^T, \cdots, \Delta x(L)^T]^T$ is the dynamic component. This dynamic component is defined as:

$$\Delta x(n) = A \sum_{m=-W}^{W} m x(n+m) \qquad (10)$$

where A is a normalization constant, and the computation is performed over a window of $2W+1$. Δx is a least square estimate of the time derivative of EEG data. It supplies extra information to the classification. We will show that this new component is very useful to improve accuracy.

Classification is performed according to (3) to (6).

4.2 Experiment Settings

Six healthy subjects participated in experiments for both FD-Speller and SD-Speller. Each participant first underwent a training session which lasted around 20 min. During the training, the subject was requested to follow prompts on the screen to spell 42 characters. 10 rounds of intensifications were presented for each character. An SVM model was trained for each subject with these training data. Then, the subject started to do on-line test. The test task was also to spell 42 characters comprising all letters and digits. All data collected on-line were used in the accuracy evaluation without any manual data selection.

4.3 Optimization of Processing Methods

Before we compare the performance of the two paradigms, we first optimize our processing methods to get best classification accuracy. This study was carried out on FD-Speller.

Character classification error rates for various signal processing approaches and their combinations are listed in Table 1. These results are the average of all subjects with 10 ensemble average. The abbreviations used in Table 1 are ARm1: artefact removal,linear superposition model; ARm2: artefact removal, difference model; Dyn: dynamic data vector; PCA: principal component analysis.

Table 1. Character classification error rates for FD-Speller.

Methods	Error Rate %	Relative error rate change to Baseline %
Baseline	6.34	-
ARm1	6.83	7.73
ARm2	7.07	11.51
Dyn	5.37	-15.30
Dyn+ARm1	6.1	-3.79
Dyn+ARm2	**4.88**	**-23.03**
PCA	5.85	-7.73
PCA+ARm1	6.59	3.94
PCA+ARm2	7.32	15.46
PCA+Dyn	6.1	-3.79
PCA+Dyn+ARm1	6.83	7.73
PCA+Dyn+ARm2	5.61	-11.51

From the test results, we can see that the dynamic data vector helps reduce the error rate most significantly. This shows that the information of tracking the EEG waveform change is really helpful in the classification. It is interesting to notice that, artifact removal methods alone actually degrade the performance, but after using with dynamic data vector, we obtain the biggest accuracy gain (Dyn+ARm2). A relative error reduction of 23% was achieved. In the rest of the paper, this best set of algorithms will be used for evaluation.

4.4 Result Comparison

The average error rates of six subjects under various durations are listed in Table 2 for comparison. The durations range from 2.2 seconds to 21.6 seconds. This duration range corresponds to 1 to 10 trials for ensemble average. The error rate reduction from FD-Speller to SD-Speller is as high as up to 81% (for 22 sec case). In other words, when the classification accuracy is at 90% and 95% respectively, SD-Speller almost doubles the information transfer rate, as is shown in Table 3.

One of the reasons why SD-Speller outperforms FD-Speller can be explained by the comparison of dynamic range of P300 signals in the two spellers. The dynamic ranges shown in Table 4 were obtained by measuring the differences between P300 peaks and valleys preceding the peak (between 200msec and the peak points). All six subjects have clearly higher P300 signal in SD-Speller than in FD-Speller.

Table 2. Character classification error rates for FD-Speller and SD-Speller. The error reduction and t-test results are also listed. (EER – error rate reduction)

	Character Classification Error Rate %				
Time (s)	2.2	4.3	6.5	8.6	10.8
FD	67.97	49.37	36.38	26.13	19.44
SD	46.63	27.33	15.90	9.41	6.50
ERR %	31.40	44.65	56.28	64.00	66.55
t-test, ρ	<.001	<.001	<.001	<.001	<.001
Time (s)	13.0	15.1	17.3	19.4	21.6
FD	15.20	11.69	9.08	7.52	4.47
SD	4.63	2.95	2.71	1.83	0.81
ERR %	69.52	74.78	70.15	75.68	81.82
t-test, ρ	<.001	<.001	<.001	<0.002	<0.02

Table 3. Durations needed to achieve particular accuracy

Accuracy %	FD-Speller (s)	SD-Speller (s)
90	15	8
95	21	13

Table 4. Dynamic ranges for FD-Speller and SD-Speller.

Subject	FD-Speller(μv)	SD-Speller (μv)
1	2.736	3.459
2	3.976	6.042
3	4.872	6.605
4	3.589	5.321
5	5.744	7.012
6	3.310	3.924

5. CONCLUSIONS

We proposed a new speller paradigm, SD-Speller. It outperformed FD-Speller significantly in terms of classification accuracy under the same information transfer intervals. Furthermore, SD-Speller allows more flexible user interface design, for instance the letters and digits are not necessarily in a square matrix. This feature enables us to design better user-friendly applications. In our experiments, all subjects preferred SD-Speller to FD-Speller because the former caused less fatigue. We tested SD-Speller in norm lighting and even very bright conditions, and the performance did not change obviously. Our test also showed that SD-Speller is very robust to ambient noises. We tested in an exhibition hall with very loud sounds nearby (>70db), SD-Speller performed equally well as it did under quiet conditions.

6. REFERENCES

[1] Wolpaw J R, Birbaumer N, MacFarland D J, Pfurtscheller G and Vaughan T M, 'Brain-computer interface for communication and control'. *Clin. Neurophyiol.*, 113:767-791, 2002.

[2] Birbaumer N, Hinterberger T, Kübler A and Neumann N, 'The Thought-Translation Device (TTD): Neurobehavioral Mechanisms and Clinical Outcome', *IEEE Trans Neural Systems and Rehab Eng*, 11(2):120-123, June 2003.

[3] Wolpaw J R, McFarland D J, Vaughan T M and Schalk G, 'The Wadsworth Center brain–computer interface (BCI) research and development program', *IEEE Trans Neural Systems and Rehab Eng*, 11:204–7, 2003.

[4] Pfurtscheller G and Neuper C, 'Motor imagery and direct brain–computer communication', *Proc. IEEE* 89:1123–34, 2001.

[5] Farwell L A and Donchin E, 'Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potential', *Electroencephalogr and Clin Neurophy*, 70:510-523, 1988.

[6] Middendorf M, McMillan G, Calhoun G and Jones K S, 'Brain–computer interfaces based on steady-state visual evoked response', *IEEE Trans Rehabil Eng.* 8: 211–3, 2000

[7] Donchin E, Spencer K M and Wijesinghe R, 'The mental prosthesis: assessing the speed of a P300-based brain–computer interface', *IEEE Trans. Rehab. Eng.* 8:174–9, 2000.

[8] Meinicke P, Kaper M, Hoppe F, Heumann M and Rotter H, 'Improving Transfer Rates in Brain Computer Interfacing: A Case Study', *in Proc NIPS* 2002, Canada.

[9] Thulasidas M, Guan C, Ranganatha S, Wu J, Zhu X, Xu W, 'Effect of Ocular Artifact Removal in Brain Computer Interface Accuracy', *in Proc of 26th Annual Inter Conf IEEE EMBS*, 1-5, Sept, 2004, US.

[10] Xu N, Gao X, Hong B, Miao X, Gao S and Yang F, 'BCI Competition 2003—Data Set IIb: Enhancing P300 Wave Detection Using ICA-Based Subspace Projections for BCI Applications', *IEEE Transactions On Biomedical Engineering, 51(6): 1067-1072, June 2004.*

[11] Brendan Z. Allison and Jaime A. Pineda, 'ERPs Evoked by Different Matrix Sizes: Implications for a Brain Computer Interface (BCI) System', *IEEE Trans. Neural Syst. Rehabil. Eng.* 11:110-113, 2003.

[12] Mellinger J, Nijboer F, Pawelzik H, Schalk G, McFarland, D J, Vaughan T M, Wolpaw J R, Birbaumer N and Kübler A, 'P300 for Communication: Evidence from Patients with Amyotrophic Lateral Sclerosis (ALS)', *in Proc of the 2nd Inter BCI Workshop, 17-18 Sept, 2004, Graz, Austria.*