A Speedy Hybrid BCI Spelling Approach Combining P300 and SSVEP

Erwei Yin, Zongtan Zhou, Jun Jiang, Fanglin Chen, *Member, IEEE*, Yadong Liu, and Dewen Hu*, *Senior Member, IEEE*

***Abstract*—This study proposes a novel hybrid brain–computer interface (BCI) approach for increasing the spelling speed. In this approach, the P300 and steady-state visually evoked potential (SSVEP) detection mechanisms are devised and integrated so that the two brain signals can be used for spelling simultaneously. Specifically, the target item is identified by 2-D coordinates that are realized by the two brain patterns. The subarea/location and row/column speedy spelling paradigms were designed based on this approach. The results obtained for 14 healthy subjects demonstrate that the average online practical information transfer rate, including the time of break between selections and error correcting, achieved using our approach was 53.06 bits/min. The pilot studies suggest that our BCI approach could achieve higher spelling speed compared with the conventional P300 and SSVEP spellers.**

***Index Terms*—Brain–computer interface (BCI), electroencephalogram (EEG), hybrid BCI, P300 potential, steady-state visually evoked potential (SSVEP).**

I. Introduction

A BRAIN–COMPUTER interface (BCI) is a type of system that can directly acquire signals from the human brain and translate them into digital commands that can be recognized and processed on a computer. A BCI provides a pathway for reestablishing communication and an environmental control capability to severely disabled persons [1], [2]. In the previous BCI research, most researchers have focused on BCI spellers with alphabetic writing systems [3]. Several different electroencephalogram (EEG) signals can be used for spelling in BCIs, such as P300 event-related potential (ERP) [4], steady-state visually evoked potential (SSVEP) [5], and sensorimotor rhythm (SMR) [6]. However, the SMR-based approach is slow and inaccurate and requires extensive training for multiple-item spelling. P300 and SSVEP potentials have been widely used in the BCI spellers due to their superior presence. The P300 potential is elicited by rare, task-relevant events and is often recorded at a latency of approximately 300 ms after the presented stimuli, an "oddball" paradigm [7]. The SSVEP is a periodic neural response evoked by certain visual repetitive stimuli [8]. Different people have different abilities to control the visual BCIs, especially for the difference between healthy persons and severely patients who lose their gaze control. To fulfill the needs of different users, P300 and SSVEP are often used to design the gaze-dependent or gaze-independent BCIs. Here, a BCI speller whose performance depends on eye gaze is a gaze-dependent BCI system, otherwise it is a gaze-independent BCI system [9], [10].

To make BCIs practical devices for communication and control, one of the main goals of current BCI speller-related research is to enhance the spelling speed under a reliable control. To fulfill this goal, several P300-based optimal approaches have been developed, focused mainly on the stimulus presentation paradigm design by decreasing the number of flashes per trial [11]–[15]. These P300-based approaches do reach higher spelling speed than traditional P300 speller; however, the performance of P300 speller does not leave much potential for further improvement due to the tradeoff between the spelling speed and the classification accuracy [16]. Indeed, previous studies have demonstrated that an SSVEP speller may achieve faster and more reliable control than a P300 speller [17]–[19]. The number of commands is limited by the number of stimulus frequencies when using a PC monitor [18], which restricts the application of this type of BCI speller. Several laboratories have proposed adaptive strategies that use multiple commands per selection [20], [21]. Although these approaches yield superb information transfer rates (ITRs) at the command level, the speed decreases significantly at the speller level because of the multiple commands per selection. Most previous studies that have examined the BCI spellers have only relied on the single-mode brain potential, which approaches a bottleneck in the spelling speed.

To further enhance the performance of BCI systems, recent work has validated a new framework for BCI research called the hybrid BCI. A hybrid BCI is a system that combines more than one BCI approach [22]. Because of the attention recently focused on the hybrid BCI systems, most recent papers presented in this research field have pertained to combinations of the SMR, P300, and SSVEP. Allison *et al.* developed a hybrid BCI based on the SMR and SSVEP [23]. Using this approach, the number

Manuscript received June 14, 2013; revised August 13, 2013; accepted September 8, 2013. Date of publication September 16, 2013; date of current version January 16, 2014. This work was supported in part by the National High-Tech R&D Program of China (863 Program) under Grant 2012AA011601, and in part by the National Natural Science Foundation of China under Grant 61375117. *Asterisk indicates corresponding author.*

E. Yin, Z. Zhou, J. Jiang, F. Chen, and Y. Liu are with the College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha 410073, China (e-mail: yinerwei1985@gmail.com; narcz@163.com; jiangjun@nudt.edu.cn; chenfanglin83@163.com; liuyadong1977@163.com).

*D. Hu is with the College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha 410073, China (e-mail: dwhu@nudt.edu.cn).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org. This includes a multimedia AVI format movie clip, which shows a natural scene. The size of the video is 11.7 MB.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2013.2281976

of illiterate subjects was reduced due to the improvement in the accuracy. The combination of SMR and P300 signals for target selection in 2-D cursor control was presented in [24]. The target selection was more reliable than that achieved using either SMR or P300 feature alone. These positive results for combined approaches have encouraged us to employ a hybrid BCI approach to increase the spelling speed.

The combination of P300 and SSVEP was first introduced by Panicker *et al.* to achieve asynchronous control [25] and was subsequently used to enhance the accuracy in [26]. Only single-frequency stimulation was employed to induce the SSVEP response in these two studies. In our previous research [27], we presented a hybrid BCI based on the fusion of multifrequency-based SSVEP and the P300 features, which produced a significant improvement in the accuracy. To the best of our knowledge, no hybrid BCI developed to increase the spelling speed has been reported to date.

In this paper, we propose a speedy hybrid BCI spelling approach that employs P300 and SSVEP as the signals that are used simultaneously to spell symbols. In this approach, the target symbol is identified by 2-D coordinates that are realized by the two brain potentials. More specifically, the P300 and SSVEP detection mechanisms are devised and integrated as two subspellers for identifying each dimensional coordinate of the target simultaneously. Thus, we only need N_1 flash codes for P300 and N_2 frequencies for an SSVEP to achieve the spelling of $N_1 \times N_2$ items. Furthermore, we design two types of BCI paradigms, termed the subarea/location (SL) and row/column (RC) modes, based on this approach. A comparison of the performance of these two modes is used to select the better paradigm for our approach and to study the influence of neighboring flashing. Finally, the practical information transfer rate (PITR) is used to determine the optimal number of stimulus trials for each subject for online spelling.

The remainder of this paper is organized as follows. The methodology, including stimulation paradigms, experimental setup, signal processing, and the selection of optimal trials, is described in Section II. The experimental results are presented and discussed in Sections III and IV, respectively. Section V concludes the paper.

II. Methodology

A. Stimulation Paradigms

To evoke the P300 and SSVEP potentials simultaneously, flash pattern mechanisms are composed of random flashings and periodic flickers. The random flashings are created by highlighting items using orange crosses in a pseudorandom sequence, and the periodic flickers are created using white rectangular objects which alternately appear on and disappear from a black background. Additional details of these hybrid stimulus mechanisms were provided in our previous study [23].

Here, we employ the classical standard BCI speller with a 6×6 matrix. The stimulus onset asynchrony (SOA) of the random flashes is 240 ms, which means that each flash is highlighted for 120 ms, with a 120-ms delay between flashes. One trial is defined as a complete cycle of flashes in which all of the stimulus code is performed once. Furthermore, we employ six frequencies, set at 8.18, 8.97, 9.98, 11.23, 12.85, and 14.99 Hz, in the design of a periodic stimulus mechanism. The selection of these frequencies was discussed thoroughly in the previous study [27]. Based on this combinational framework, we design two presentation modes using different stimulus sequences (see Fig. 1).

1) SL mode: [see Fig. 1(a)]. In the SL mode, the character matrix is divided into six subareas surrounded by dashed lines. All of the items in each subarea flash at the same frequency. The SSVEP feature is used to identify the subarea to which the target item belongs. At the same time, orange crosses highlight the same location in each subarea in a pseudorandom sequence. The location of the target item in the subareas is determined by the P300 feature. Thus, the target item is detected by the number of the subarea and its location.

2) RC mode: [see Fig. 1(b)]. In the RC mode, the flashing of items in the matrix is grouped into rows and columns. Each column surrounded by a dashed line flashes at same frequency to fix the column coordinate of the target character. Meanwhile, the rows of the matrix flash in a random order to fix the row coordinate of the target character. The intersection of the row and column determined by the SSVEP and P300 detection algorithms is identified as the target character.

B. Experimental Setup

1) Subjects: Fourteen healthy subjects (six females and eight males, aged 18–41 years, with a mean age of 28.7 years) participated in the study. None had uncorrected visual impairments or any cognitive deficit. Only two of the subjects had used the P300-SSVEP-based hybrid BCI prior to the study; the remaining 12 subjects were complete novices. All the subjects signed an informed consent form in accordance with the Declaration of Helsinki and were paid for their participation. The purpose of this study and the task required were explained to each subject in detail before preparation for the EEG recording.

2) Data Collection: The EEG signals were recorded from the surface of the scalp via ten-channel active electrodes placed at Cz, Pz, P3, P4, Oz, O1, O2, POz, PO7, and PO8, referenced to TP10 and grounded to Afz, based on the 64-channel extended international 10/20 system (see Fig. 2). The data were amplified and digitized by a BrainAmp DC Amplifier (Brain Products GmbH, Germany). The channels Cz, Pz, P3, and P4 were used for P300 detection, and the channels O1, O2, POz, PO7, and PO8 were used for the SSVEP detection. The channel Oz was used for both the P300 and SSVEP detection. The impedances were kept below 10 kΩ prior to recording using a conductive paste. The EEG signals were sampled at 250 Hz and filtered using a 50-Hz notch filter. The data collection, stimulus presentation, and experimental procedure associated with the BCI speller were controlled using the BCI2000 platform [28], which provides a Python interface for stimuli presentation and a MATLAB interface for signal processing.

3) Experimental Paradigm: The experiments were performed in a normal office room. The subjects were seated 70 cm in front of a 27$''$ LED monitor with a refresh rate of 60 Hz.

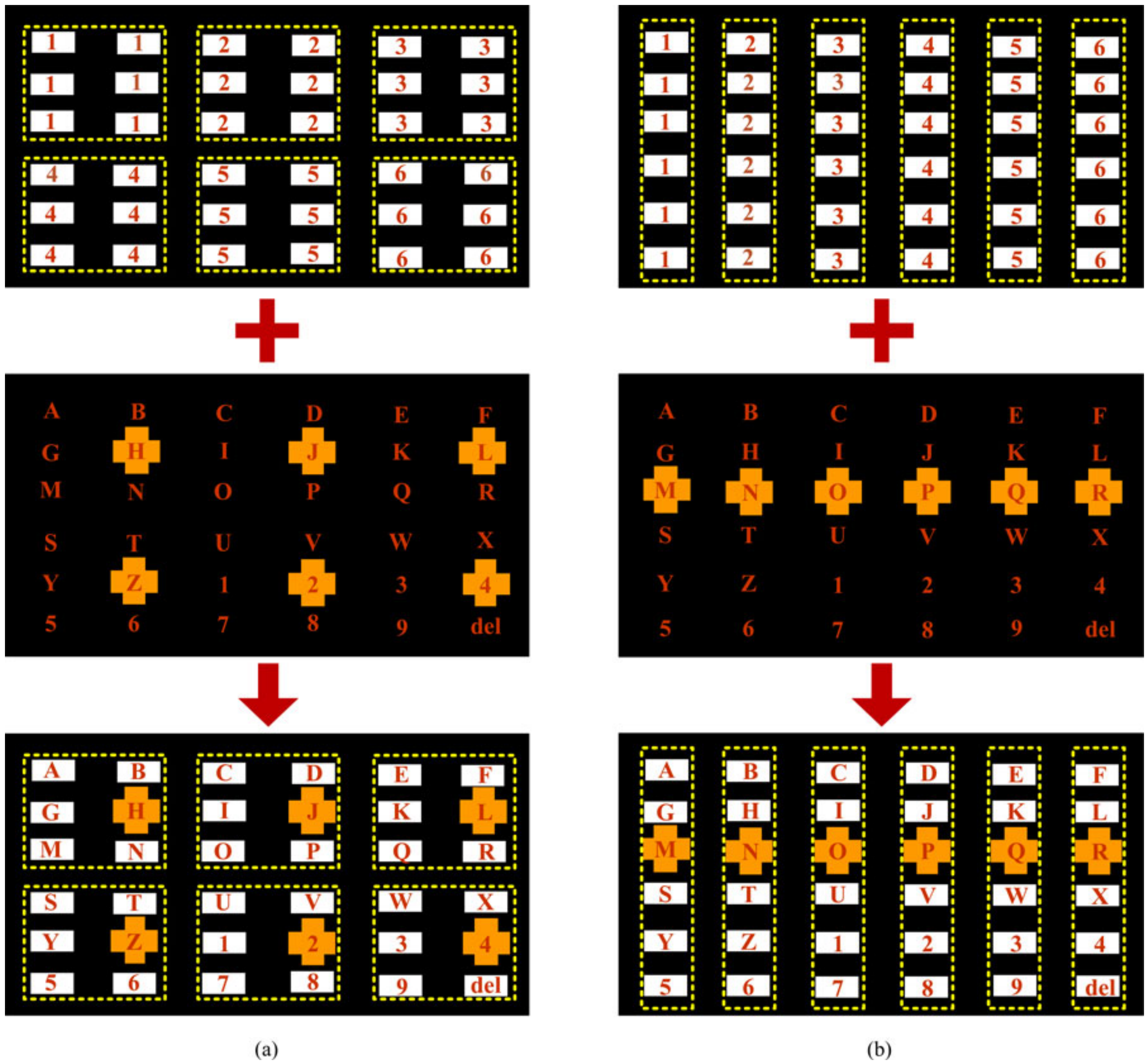

Fig. 1. Illustrations of the stimulus configuration for (a) SL mode and (b) RC mode. The flash pattern mechanisms are composed of random flashings and periodic flickers, represented by white rectangular objects and orange crosses, respectively. Moreover, all of the items within a subarea surrounded by a single-dashed line flash at the same frequency, and the orange crosses highlight synchronously at the same locations in all of the subareas surrounded by the dashed lines.

Each subject participated in four sessions on two different days within a one-week time period. To avoid fatigue bias, a random order of the sessions was used for each subject, and each subject was given a 5-min break for rest after each run. Furthermore, to avoid confusing the subjects, a 2-s break was given to allow the subjects to locate the next symbol. Additional details of the four sessions are described below.

1) In the P300 session, we employed the standard P300 speller with the RC paradigm. The P300 stimuli were presented by random flashing of the rows or columns using the orange crosses in a pseudorandom sequence. One trial consisted of 12 flashes (2880 ms in total). During the experiments, the subjects were instructed to maintain a mental count of the number of times the prompted symbol was highlighted. Each subject was required to perform six runs, each of which consisted of 14 symbols' spellings in random order to avoid adaptation (eight trials per symbol for a total of 84 symbols, approximately 80 min).
2) In the SSVEP session, the cells of the matrix flickered at six different frequencies, with the same frequency sequence as in the SL mode. The reason that we did not compose all of the cells of the matrix into six rectangular objects, as in [18], is that overly large white rectangular objects could be annoying to the subjects [29]. Note that the outputs of the SSVEP session were the number of frequencies because the target character could not be identified using only the SSVEP response. Only three runs (42 symbols in total) were performed for the offline analysis because SSVEP detection does not require the calibration data for the training classifier. During these runs, each subject was instructed to gaze at the prompted item for eight trials per symbol. Each trial consisted of six flashes (1440 ms in total). This session took approximately 30 min.
3) Both the SL and RC sessions were implemented with similar experimental processes consisting of a calibration

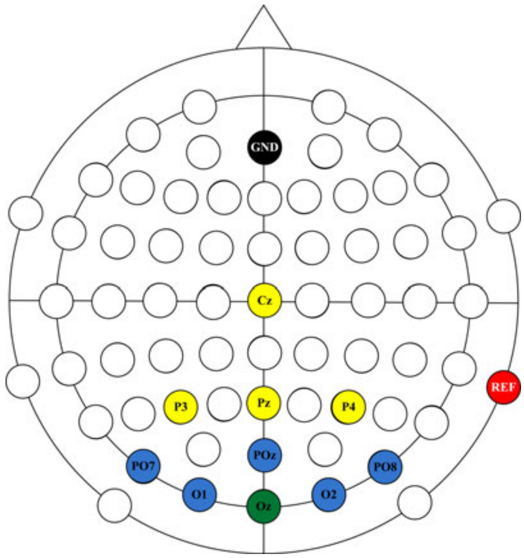

Fig. 2. Electrode configuration and locations used for the study. Ten channels are used for data collection according to the international 10/20 system. The reference and ground electrodes are TP10 (red holder) and AFz (black holder), respectively. The yellow and blue holders are employed for P300 and SSVEP detections, respectively, and the green holder is used for both the P300 and SSVEP detection.

phase and online test phase. During the SL and RC sessions, the subjects were instructed to simultaneously conduct the tasks performed in both the P300 and SSVEP sessions. In the SL and RC calibration phases, each subject had to implement six runs (eight trials per symbol for a total of 84 symbols). The calibration phases were conducted for the offline analysis and the selection of the optimal trial for the online test phases. Subjects were not shown a word to spell in the online phases. To provide a standardized spelling task, the subjects were required to spell their own names in Latin letters three times. The optimal number of trials for each subject was selected using the calibration data. During the online runs, when an incorrect symbol was detected, the subject had to correct the misspelling first using the "del" option located at the bottom right of the matrix and then respell the symbol. Both the SL and RC sessions took approximately 80 min including offline and online phases.

C. Signal Processing

We explored parallel P300 and SSVEP signal processing in the proposed hybrid BCI approach (see Fig. 3). First, the EEG signals from the P300 and SSVEP channels were separated. Second, these two types of signals were processed simultaneously by the respective P300 and SSVEP detection mechanisms. Finally, the target coordinates were determined based on the maximum score of the 2-D temporal frequency features, including the P300 and SSVEP. The target coordinates can be represented as follows:

$$(x, y) = \underset{i,j \in [1,\ldots,6]}{\arg} [\max(\text{Score}_i^{\text{P300}}), \max(\text{Score}_j^{\text{SSVEP}})] \quad (1)$$

where (x, y) denotes the target coordinates. The target coordinates are defined as the number of the subarea and its position in the SL mode, whereas the target coordinates represent the row and column coordinates in the RC mode. Furthermore, $\text{Score}_i^{\text{P300}}$ and $\text{Score}_j^{\text{SSVEP}}$ are calculated by the P300 and SSVEP detections, respectively. The complete signal processing approaches are summarized below.

1) P300 Detection: First, the EEG data from the P300 channels were filtered using a 0.1–45-Hz bandpass filter to eliminate signal excursion and high-frequency noise. Next, the 800-ms epochs of the data, starting from the onset of each stimulus, were extracted for the P300 feature analysis. After the feature extraction, the epochs of the calibration data were down-sampled from 250 to 25 Hz to reduce the feature size due to the low frequency of the information received for detection of the P300 response [30]. Then, a stepwise linear discriminant analysis (SWLDA) was performed to select suitable weighting parameters that best discriminated between the target and non-target flashes. The details of the algorithm are provided in [31]. SWLDA was used to train a classifier to calculate the scores of the P300 responses using the following equation:

$$Y_{ik}^{\text{p300}} = w^T X_{ik}^{\text{p300}} \quad (2)$$

where *i* and *k* denote the stimuli codes and the number of trials, respectively, and w represents the SWLDA classifier. Finally, the score of each stimulus code was calculated by averaging all of the scores according to the same code numbers

$$\text{Score}_i^{\text{P300}} = \frac{1}{K} \sum_{k=1}^{K} Y_{ik}^{\text{p300}} \quad (3)$$

where K is the total number of trials associated with the current decision.

2) SSVEP Detection: To enhance the signal-to-noise ratio, the data collected from the SSVEP channels were filtered using a 4–35-Hz bandpass filter. The canonical correlation analysis (CCA) was used to calculate the correlation between the stimulus frequency $(X_f(t))$ and the multichannel EEG data (Y^{ssvep}). CCA is an effective method for measuring the SSVEP response [32]. Here, the stimulus frequency is represented as a square wave periodic signal that can be decomposed into Fourier series of its harmonics

$$X_f(t) = \begin{bmatrix} \sin(2\pi f t) \\ \cos(2\pi f t) \\ \sin(2 \cdot 2\pi f t) \\ \cos(2 \cdot 2\pi f t) \\ \vdots \\ \sin(M \cdot 2\pi f t) \\ \cos(M \cdot 2\pi f t) \end{bmatrix}, \quad t = \frac{1}{S}, \frac{2}{S}, \ldots, \frac{N}{S} \quad (4)$$

where M is the number of harmonics, N is the number of sampling points per selection, and S is the sampling rate, set at 250 Hz. Finally, the score of each stimulus frequency was represented as follows:

$$\text{Score}_j^{\text{ssvep}} = r_j(X_{f_j}(t), Y^{\text{ssvep}}) \quad (5)$$

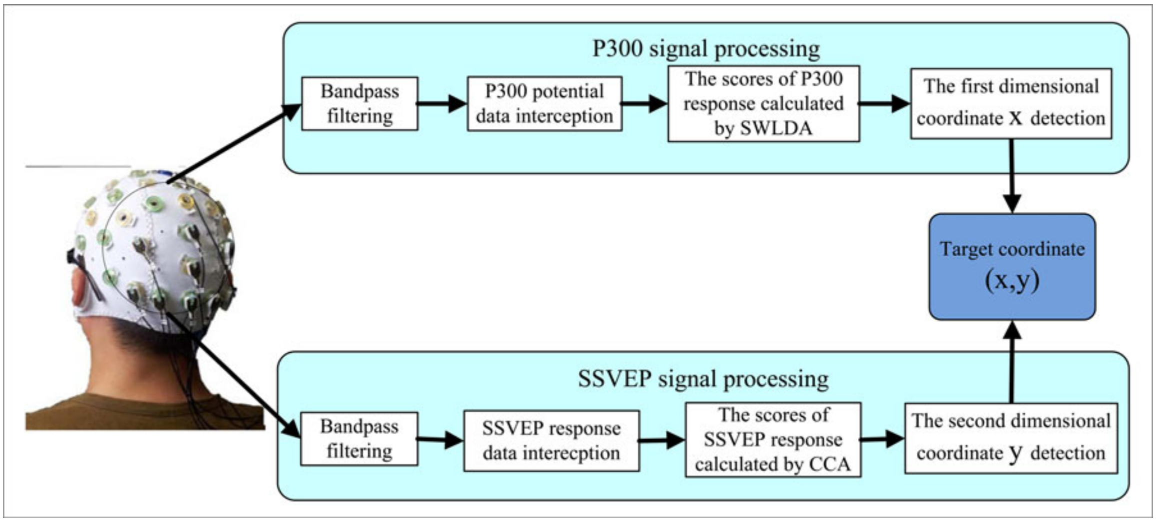

Fig. 3. Data analysis framework of our speedy hybrid BCI spelling approach.

where r is the correlation coefficient, f is the stimulus frequency, and j is the sequence number of the stimulus frequency. The correlation coefficient r is calculated using the function "canoncorr.m" in the MATLAB toolbox [33].

D. *Selection of Optimal Trials*

Due to the difference in the spelling capability among the users, the selection of trials for each subject leads to an optimization problem [34]. ITR is widely used in the BCI community as an important metric, which can be computed using the following formula [35]:

$$\mathrm{ITR} = \left\{ \log_2 N + P \log_2 P + (1-P) \log_2 \frac{1-P}{N-1} \right\} / T \tag{6}$$

where N is the number of options, P is the spelling accuracy, and T is the time interval per selection.

However, for practical applications, the speller must have an error-correcting function. In the online spelling process, two additional selections are required to correct a misspelling with the "del" option and select the correct symbol. When we finish the spelling task, after correcting all of the errors, we note that the accuracy P is now 100%. In this case, the ITR can be expressed as follows:

$$\mathrm{ITR} = \log_2 N/T. \tag{7}$$

Assuming that the classification accuracy remains stable at P, the spelling task consists of M symbol selections, and the user completes the spelling task by correcting all errors. Thus, the total number of selections M' can now be written as an iterative formula [36]

$$\begin{aligned} M' &= M + 2M(1-P) + 2(2M(1-P))(1-P) + \cdots \\ &= M\frac{1}{2P-1}, \quad (P > 0.5). \end{aligned} \tag{8}$$

The expected time needed per symbol then becomes the following:

$$T' = \frac{T}{2P-1}, \quad (P > 0.5). \tag{9}$$

Substituting (9) into (7), we obtain the following expression:

$$\mathrm{PITR} = \log_2 N/T' = \begin{cases} (2P-1)\log_2 N/T, & P > 0.5 \\ 0, & P \le 0.5. \end{cases} \tag{10}$$

T is computed from the following expression:

$$T = (S \cdot L \cdot R + I)/60 \tag{11}$$

where S is the SOA (240 ms), L is the number of flashes per trial, R is the number of trials per symbol, and I is the time break between selections (2 s).

This metric represents the information per minute a subject can correctly make in a reasonable real-world setting and is referred to as the PITR. In this paper, we selected the optimal number of trials by calculating the maximum PITR for each subject. The PITR was evaluated by a fivefold cross-validation procedure with the offline data. Fig. 4 presents the estimated optimal number of trials for each subject using the SL and RC modes.

III. Experimental Results

A. *Online Performance*

During the online experiments, each subject spelled his or her own name three times with the selected optimal number of trials. The number of task characters M_c and practical selections M_p were recorded directly. To facilitate PITR calculation, formula (9) was rewritten as follows:

$$T' = T \cdot \frac{M_p}{M_c}. \tag{12}$$

Considering that all of the subjects completed the spelling task successfully, substituting (12) into (7), the PITR of online

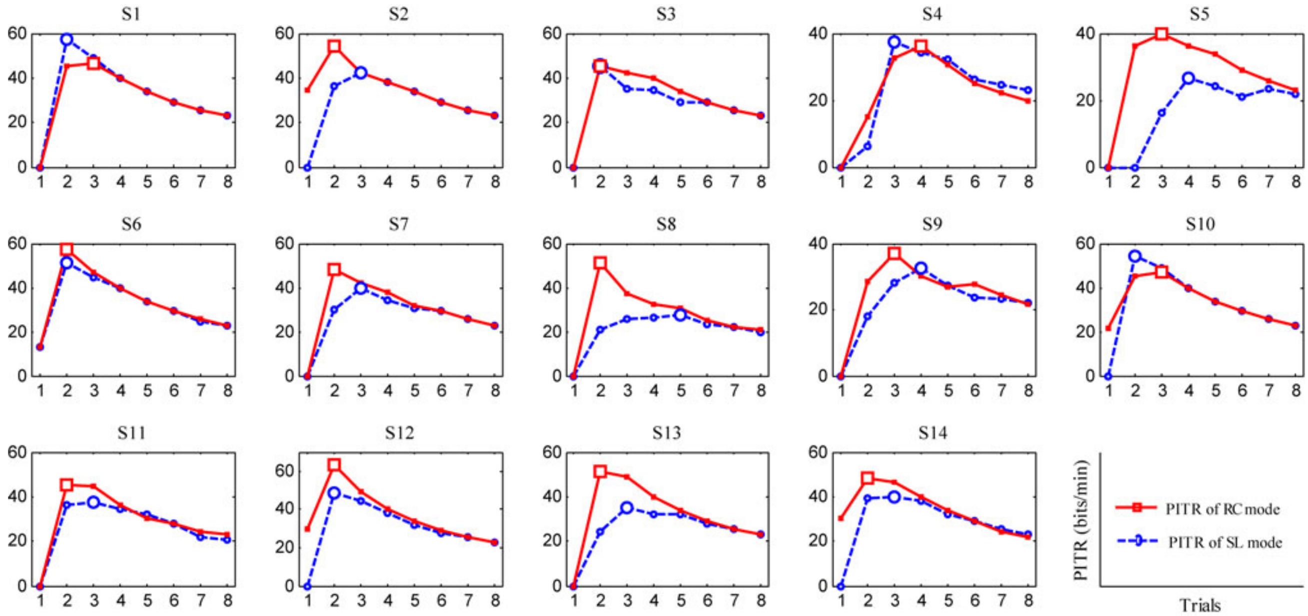

Fig. 4. Process of selecting optimal number of trials for each subject using the SL and RC modes. The curves were estimated with the offline data. In each subfigure, the markers represent the optimal number of trials that correspond to the maximum average PITR for each mode.

TABLE I
ONLINE PERFORMANCE COMPARISONS BETWEEN THE SL AND RC MODES

Subject	Task Characters	SL Mode			RC Mode		
		Optimal Trial	Practical Selections	PITR (bits/min)	Optimal Trial	Practical Selections	PITR (bits/min)
S1	30	2	30	63.56	3	30	49.08
S2	24	3	28	42.07	2	28	54.48
S3	30	2	40	47.67	2	34	56.09
S4	30	3	36	40.90	4	34	35.27
S5	30	4	38	31.56	3	36	40.90
S6	24	2	24	63.56	2	24	63.56
S7	30	3	36	40.90	2	34	56.09
S8	30	5	34	29.75	2	32	59.59
S9	24	4	42	22.84	3	32	36.81
S10	30	2	32	59.59	3	30	49.08
S11	39	3	45	42.54	2	39	63.56
S12	24	2	26	58.68	2	24	63.56
S13	18	3	18	49.08	2	18	63.56
S14	33	3	49	33.05	2	41	51.16
AVG	28.29	2.93	34.14	44.70	2.43	31.14	53.06
STD	5.09	0.92	8.53	13.03	0.65	6.18	9.88

The results were obtained from the online spelling tests using the selected optimal number of trials. The PITR (bits/min) included the time breaks between selections and error correcting.

spelling can be expressed as follows:

$$\text{PITR} = \frac{M_c \cdot \log_2 N}{T \cdot M_p}. \tag{13}$$

In addition, we performed a difference analysis of the performance of the two modes using paired t-tests.

The online performance of the two modes using our hybrid approach is summarized in Table I. The results demonstrate that the RC mode was significantly better than the SL mode. Specifically, for the RC mode, the average PITR increased significantly, from 44.70 to 53.06 bits/min ($p < 0.0128$), and the standard deviation decreased from 13.03 to 9.88 bits/min. In addition, the online results depend heavily on the selection of optimal trials. Although the average optimal number of trials for the RC mode was less than that for the SL mode (2.43 versus 2.93), the RC mode required fewer selections to complete the spelling tasks than the SL mode (31.14 versus 34.14, $p < 0.002$). That is, the spelling speed of the RC mode might be faster than that of the SL mode. More importantly, the highest peak PITR of 63.56 bits/min without a spelling error was achieved by four

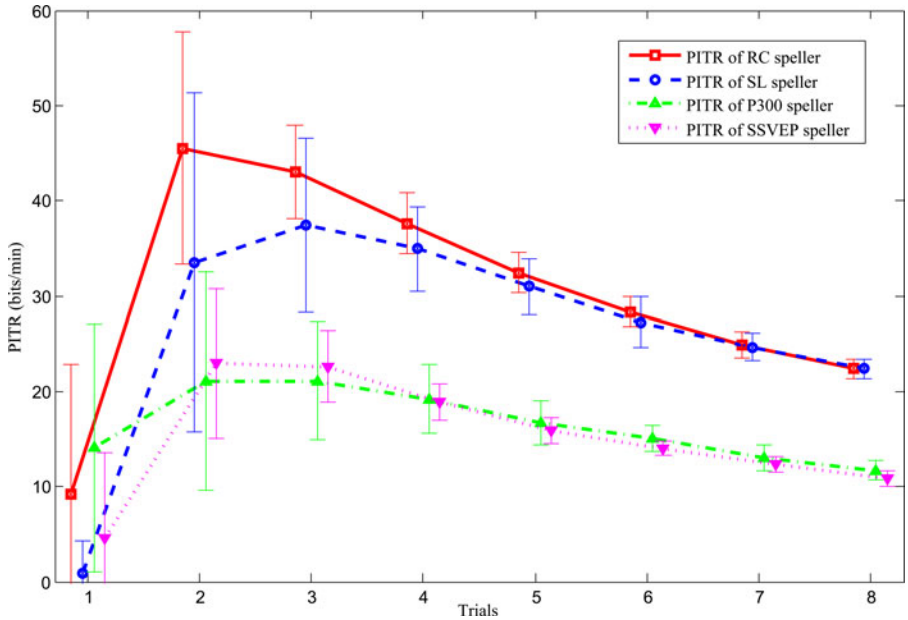

Fig. 5. System performance comparisons between the hybrid BCI approach (RC and SL spellers) and single-mode BCI approaches (P300 and SSVEP spellers) by the offline analysis. For clarity, the curves are shifted slightly left and right.

subjects using the RC mode, which indicates its potential in practical applications of the stimulus-driven BCI system.

B. *Offline Analysis*

Fig. 5 presents the results of offline PITR comparisons between our speedy hybrid BCI approach (SL mode and RC mode) and traditional BCI approaches (P300 and SSVEP spellers) conducted to confirm the superiority of our approach and to better understand the online results. The results indicate that the spellers using our speedy hybrid BCI approach achieved markedly higher average PITRs than both P300 and SSVEP spellers in all trials except the first trial, in which all of the spellers exhibited poor performance. In particular, using our approach (RC mode), the peak value of the average PITR was improved by 115.41% (45.50 versus 21.12 bits/min) and 98.55% (45.50 versus 22.92 bits/min) compared with the P300 and SSVEP spellers, respectively. Furthermore, in our approach, the PITR for the RC mode was noticeably higher than that for the SL mode over the first six trials. The error bars in Fig. 5, which represent the standard deviations of PITR achieved with different numbers of trials, indicate that smaller standard deviations are associated with the RC mode. This observation suggests that the RC mode achieves more stable system performance than the SL mode. These analysis results indicate that the RC mode using our hybrid approach yielded better system performance than the other three spellers.

IV. Discussion

A. *Potential Advantages of Speedy Hybrid BCI Approach*

With our speedy hybrid spelling approach, the target items are actually realized simultaneously by the P300 and SSVEP subspeller mechanisms. To the best of our knowledge, this hybrid strategy is the first multispeller system proposed in the BCI research field to date. As shown in Table II, if we employ an M-length flash code for the P300, together with N frequencies for the SSVEP, we can achieve the spelling of $M \times N$ items. For example, because the target character is selected with the P300 and SSVEP simultaneously, we only need to employ six flash codes for the P300 and six frequencies for the SSVEP to achieve the spelling of 36 items. Using a conventional P300 speller, it would require 12 flashes per trial to identify the target character. Therefore, the stimulus time decreases by half compared to that of the P300 speller. Furthermore, the number of target items of our hybrid strategy increase by six times compared with the SSVEP subspeller. According to the current reports, spelling of 36 items requires multiple commands per selection using the SSVEP speller, but this task can be achieved with a single command using our hybrid strategy.

TABLE II
Target Number That can be Coded by Different Lengths of P300 Code and Number of SSVEP Frequencies

P300 Code Length	SSVEP Frequency Number						
	2	3	4	5	6	···	N
2	4	6	8	10	12	···	$2N$
3	6	9	12	15	18	···	$3N$
4	8	12	16	20	24	···	$4N$
5	10	15	20	25	30	···	$5N$
6	12	18	24	30	36	···	$6N$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
M	$2M$	$3M$	$4M$	$5M$	$6M$	···	$M \times N$

The aforementioned theoretical analysis and the pilot study indicate that the hybrid BCI approach could be more effective than single-mode BCIs. Furthermore, our approach has the potential advantage of covering all of the commands of a standard keyboard with a speedy spelling speed, thus highly enhancing system performance in terms of PITR.

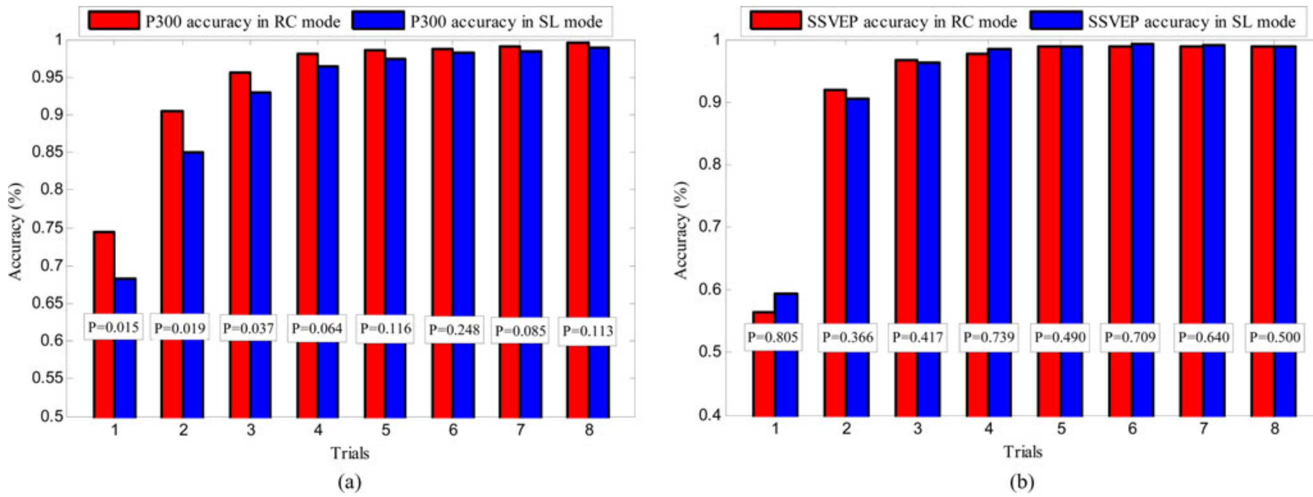

Fig. 6. Comparisons of the multiclass accuracy (one out of six) for the P300 and SSVEP between the RC and SL modes. (a) Comparison of P300 accuracy. (b) Comparison of SSVEP accuracy.

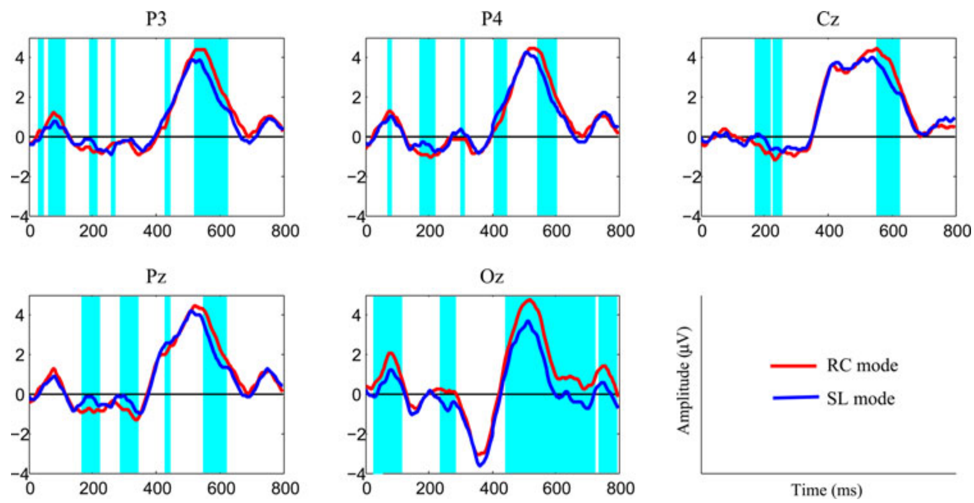

Fig. 7. ERP differences between the RC and SL modes. Grand average waveforms computed over all subjects at electrode locations P3, P4, Cz, Pz, and Oz.

B. Reasoning Behind the Performance Difference Between the RC and SL Modes

To determine why the RC mode performed better than the SL mode, we first conducted accuracy comparisons of P300 and SSVEP for these two modes. Note that the P300 and SSVEP accuracy in this paper do not represent the spelling accuracy but rather the multiclass accuracy (one out of six) of the two EEG features. As illustrated in Fig. 6, the P300 accuracies associated with the RC mode were higher than those associated with the SL mode for all trials [see Fig. 6(a)], whereas no noticeable differences were observed in the SSVEP comparison [see Fig. 6(b)]. More specifically, the p-values of the paired t-tests confirm that the P300 accuracy was significantly higher in the first three trials ($p < 0.05$), whereas there were no significant differences in the SSVEP comparisons for any trials.

According to [37], clear SSVEP suppression occurs when competing stimuli are visually closer than 4.5°. That is, the SSVEP will perform poorly when the competing stimuli are too close to the point of fixation. In our spellers (see the white rectangular objects employed to evoke SSVEP response in Fig. 1), all of the horizontal proximities of competing stimuli lie within 7.33°, indicating that there should be no influence from horizontal neighbor flicks. In contrast, in the vertical direction, neighboring stimuli are 4.09° apart, which is less than 4.5°. All of the white rectangular objects flash at the same frequency in the RC mode, indicating that there are no competing stimuli in this direction. In the SL mode, although some items have competing stimuli in the vertical direction, the offside of the competing stimuli has the stimuli flashing at the target frequency to offset the SSVEP suppression. This fact might be why there were no significant differences in the SSVEP accuracy between the SL and RC modes.

Fig. 7 demonstrates the ERP differences between the RC and SL modes. The waveforms were computed by averaging the P300 responses over all subjects for channels P3, P4, Cz, Pz, and Oz. The shaded areas indicate significant differences ($p < 0.05$) between the two modes. Most shaded areas appear at 400–650 ms, which indicates that the significant variations in ERP belong to the P300 component [38]. The RC mode produces significantly stronger P300 amplitudes than the SL mode.

In the P300 spelling approach, errors arise with the greatest frequency at locations adjacent to the target item. These errors are referred to as "adjacency-distraction errors" [13]. Compared to the RC mode, the SL mode has more flashes of nontarget items that are adjacent to the target (see Fig. 1). These nontarget flashes may not only attract the subject's attention and cause the adjacent items to be selected unintentionally but also reduce the surprise of target flashes, thereby decreasing the P300 amplitude. Based on these results, we conclude that the RC mode achieves higher overall accuracy than the SL mode due to the increased P300 accuracy.

C. Limitations and Future Directions

Our principal objective was to investigate a hybrid BCI approach as the foundation for a new type of speedy BCI speller. We employed the traditional P300 and SSVEP paradigms to create the hybrid strategy. The system performance remains unexplored. The limitations of the current approach and corresponding future directions are discussed below.

First, in addition to different individuals having different spelling capabilities when using the P300 and SSVEP spellers, the performance of any one individual varies to some extent from one recording session to the next [39], [40]. Furthermore, by comparing Table I and Fig. 5, as well as the results reported in [27], the online PITR was found to be higher than the offline results indicated, perhaps because the more rapid feedback encouraged the subjects to more actively focus their attention on the target. That is, the subjects' online performance may be underestimated during the selection of the optimal number of trials using the calibration data. The fixed optimal method does increase the performance with respect to traditional ways, while their performance gains still leave room for improvement, which has been observed by Scheuder *et al.* [34]. Consequently, to further improve the performance of our BCI speller, we will study the adaptive speller concept, in which the spelling speed may increase or decrease depending on the current individual's state.

Second, the individuals' abilities in eliciting the P300 and SSVEP potentials are different. Some subjects who succeed in producing P300 potentials may perform poorly in producing SSVEP potentials and vice versa. Because the target detection of our approach depends on the detections of the P300 and SSVEP subspellers simultaneously, one of these two potentials produced with poor accuracy will directly affect the spelling accuracy. To achieve more reasonable use of the multimodel information, future work should focus on the optimal design of the speller matrix size, including the lengths of the P300 flash codes and the number of SSVEP frequencies, based on individuals' performance with the P300 and SSVEP.

Third, the assessment of our speller was carried out under the assumption that the subjects always conducted the spelling task. In real-life settings, users should issue commands at their own pace. Asynchronous control is a type of approach that can detect whether the user intends to send information [38]. No BCI system that has achieved asynchronous control using multimodal brain signals has been reported to date. The fusion of P300 and SSVEP features may provide satisfactory performance for asynchronous control. To this end, future work should attempt to address the issue of asynchronous control as another important subject of study.

Finally, our BCI approach is a type of gaze-dependent BCI which is heavily depending on the users' ability of gaze control. Thereby, the performance in severely patients without the eye movements tends to be obviously lower than that in the healthy subjects. In order to provide more practical BCI spellers for severe patients, further studies are needed to employ the combination of P300 and SSVEP to design a more effective gaze-independent BCI approach.

V. Conclusion

In this paper, we proposed a new speedy hybrid BCI spelling approach that uses the P300 and SSVEP detection mechanisms simultaneously. In this approach, the target coordinates are identified by both the P300 and SSVEP subspellers. Additionally, we designed two BCI paradigms, involving SL and RC modes, based on this approach. To achieve reasonable online performance, we selected the optimal number of trials for each subject using the PITR calculated from the calibration data.

The experimental results with 14 subjects suggest that the spelling speed of our BCI approach was undoubtedly improved compared to that of traditional BCI approaches, e.g., P300 and SSVEP spellers, because of the information superposition of the bimodal brain potentials. Furthermore, the RC mode yielded higher PITR than the SL mode because of the increased P300 accuracy. The online average PITR, including the time for breaks between selections and for error correcting, was 53.06 bits/min. Four of the subjects achieved the highest peak PITR of 63.56 bits/min without spelling errors. Our hybrid BCI approach provides a new direction for the enhancement of the BCI speed.

References

[1] J. R. Wolpaw, D. J. M. N. Birbaumer, P. G, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.

[2] Y. Wang, Y. Wang, and T. Jung, "A cell-phone-based brain-computer interface for communication in daily life," *J. Neural Eng.*, vol. 8, p. 025018, 2011.

[3] H. Cecotti, "Spelling with non-invasive brain-computer interfaces-current and future trends," *J. Physiol. - Paris*, vol. 105, pp. 106–114, 2011.

[4] J. N. Mak, Y. Arbel, J. W. Minett, L. M. McCane, B. Yuksel, D. Ryan, D. Thompson, L. Bianchi, and D. Erdogmus, "Optimizing the p300-based brain-computer interface: Current status, limitations and future directions," *J. Neural Eng.*, vol. 8, p.025003, 2011.

[5] X. Gao, D. Xu, M. Cheng, and S. Gao, "A BCI-based environmental controller for the motion-disabled," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 137–140, Jun. 2003.

[6] K. R. Müller, M. Tangermann, G. Dornhege, M. Krauledat, G. Curio, and B. Blankertz, "Machine learning for real-time single-trial EEG-analysis:

From brain-computer interfacing to mental state monitoring," *J. Neurosci. Methods*, vol. 167, no. 1, pp. 82–90, 2008.
[7] L. A. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials electroencephalogram," *Clin. Neurophysiol.*, vol. 70, no. 6, pp. 510–523, 1988.
[8] F. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives," *Prog. Neurobiol.*, vol. 90, pp. 418–438, 2010.
[9] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk, "Does the 'P00' speller depend on eye gaze?" *J. Neural Eng.*, vol. 7, p. 056013, 2010.
[10] M. S. Treder, N. M. Schmidt, and B. Blankertz, "Gaze-independent brain-computer interfaces based on covert attention and feature attention," *J. Neural Eng.*, vol. 8, p. 066003, 2011.
[11] J. Jin, B. Z. Allison, E. W. Sellers, C. Brunner, P. Horki, X. Wang, and C. Neuper, "An adaptive P300-based control system," *J. Neural Eng.*, vol. 8, p. 036006, 2011.
[12] G. Townsend, J. Shanahan, D. Ryan, and E. W. Sellers, "A general P300 brain-computer interface presentation based on performance guided constraints," *Neurosci. Lett.*, vol. 531, no. 2, pp. 63–68, 2012.
[13] 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.
[14] J. Jin, P. Horki, C. Brunner, X. Wang, C. Neuper, and G. Pfurtscheller, "A new P300 stimulus presentation pattern for EEG-based spelling systems," *Biomed. Tech.*, vol. 54, no. 4, pp. 203–210, 2010.
[15] J. Jin, B. Z. Allison, E. W. Sellers, C. Brunner, P. Horki, X. Wang, and C. Neuper, "Optimized stimulus presentation patterns for an event-related potential EEG-based brain-computer interface," *Med. Biol. Eng. Comput.*, vol. 49, no. 2, pp. 181–191, 2011.
[16] C. L. Gonsalvez and J. Polich, "P300 amplitude is determined by target-to-target interval," *Psychophysiology*, vol. 39, pp. 388–396, 2002.
[17] O. Friman, I. Volosyak, and A. Graser, "Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 742–750, Apr. 2007.
[18] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method," *J. Neural Eng.*, vol. 6, p. 046002, 2009.
[19] H. Hwang, J. Lim, Y. Jung, H. Choi, S. Lee, and C. Im, "Development of an SSVEP-based BCI spelling system adopting a qwerty-style led keyboard," *J. Neurosci. Methods*, vol. 208, no. 1, pp. 59–65, 2012.
[20] H. Cecotti, "A self-paced and calibration-less SSVEP based brain-computer interface speller," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 18, no. 2, pp. 127–133, Apr. 2010.
[21] I. Volosyak, "SSVEP-based Bremen-BCI interface-boosting information transfer rates," *J. Neural Eng.*, vol. 8, p. 036020, 2011.
[22] G. Pfurtscheller, B. Z. Allison, C. Brunner, G. Bauernfeind, T. Solis-Escalante, R. Scherer, T. O. Zander, G. Müller-Putz, C. Neuper, and N. Birbaumer, "The hybrid BCI," *Front. Neurosci*, vol. 4, no. 30, p. 00003, 2010.
[23] B. Z. Allison, C. Brunner, V. Kaiser, G. R. Müller-Putz, C. Neuper, and G. Pfurtscheller, "Toward a hybrid brain-computer interface based on imagined movement and visual attention," *J. Neural Eng.*, vol. 7, p. 026007, 2010.
[24] J. Long, Y. Li, T. Yu, and Z. Gu, "Target selection with hybrid feature for BCI-based 2-D cursor control," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 1, pp. 132–140, Jan. 2012.
[25] R. C. Panicker, S. Puthusserypady, and Y. Sun, "An asynchronous P300 BCI with SSVEP-based control state detection," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 6, pp. 1781–1788, Jun. 2011.
[26] M. Xu, H. Qi, B. Wan, T. Yin, Z. Liu, and D. Ming, "A hybrid BCI speller paradigm combining p300 potential and the SSVEP blocking feature," *J. Neural Eng.*, vol. 10, p. 026001, 2013.
[27] E. Yin, Z. Zhou, J. Jiang, F. Chen, Y. Liu, and D. Hu, "A novel hybrid BCI speller based on the incorporation of SSVEP into the P300 paradigm," *J. Neural Eng.*, vol. 10, p. 026012, 2013.
[28] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: A general-purpose brain-computer interface (BCI) system," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, Jun. 2004.
[29] B. Z. Allison, T. Luth, D. Valbuena, A. Teymourian, I. Volosyak, and A. Graser, "BCI demographics: How many (and what kinds of) people can use an SSVEP BCI?" *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 2, pp. 107–116, Apr. 2010.
[30] H. Cecotti, B. Rivet, M. Congedo, C. Jutten, O. Bertrand, E. Maby, and J. Mattout, "A robust sensor-selection method for P300 brain-computer interfaces," *J. Neural Eng.*, vol. 8, p. 016001, 2011.
[31] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayoudh, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "A comparison of classification techniques for the P300 speller," *J. Neural Eng.*, vol. 3, no. 4, pp. 299–305, 2006.
[32] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 6, pp. 1172–1176, Jun. 2007.
[33] W. J. Krzanowski, *Principles of Multivariate Analysis: A User's Perspective*. New York, NY, USA: Oxford Univ. Press, 1988.
[34] M. Schreuder, J. Höhne, B. Blankertz, S. Haufe, T. Dickhaus, and M. Tangermann, "Optimizing event-related potential based brain-computer interfaces: A systematic evaluation of dynamic stopping methods," *J. Neural Eng.*, vol. 10, p. 036025, 2013.
[35] J. R. Wolpaw, N. Birbaumer, W. J. Heetderrks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: A review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 161–173, Jun. 2000.
[36] B. D. Seno, M. Matteucci, and L. T. Mainardi, "The utility metric: A novel method to assess the overall performance of discrete brain-computer interfaces," *IEEE Trans. Rehabil. Eng.*, vol. 18, no. 1, pp. 20–28, Feb. 2010.
[37] S. Fuchs, S. K. Andersen, T. Gruber, and M. Müller, "Attentional bias of competitive interactions in neuronal networks of early visual processing in the human brain," *NeuroImage*, vol. 41, pp. 1086–1101, 2008.
[38] H. Zhang, C. Guan, and C. Wang, "Asynchronous P300-based brain-computer interfaces: A computational approach with statistical models," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 6, pp. 1754–1763, Jun. 2008.
[39] A. Lenhardt, M. Kaper, and H. J. Ritter, "An adaptive P300-based online brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 2, pp. 121–130, Apr. 2008.
[40] M. Cheng, X. Gao, S. Gao, and D. Xu, "Design and implementation of a brain-computer interface with high transfer rates," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 10, pp. 1181–1186, Oct. 2002.

Erwei Yin was born in Heilongjiang, China, in 1985. He received the B.Sc. degree from Jilin University, Changchun, China, in 2008, and the M.Sc. degrees from the National University of Defense Technology, Changsha, China, in 2010, where he is currently working toward the Ph.D. degree.

His research interests include brain–computer interfaces and machine learning.

Zongtan Zhou was born in Henan, China, in 1969. He received the B.Sc., M.Sc., and Ph.D. degrees from the National University of Defense Technology, Changsha, China, in 1990, 1994, and 1998, respectively.

He became a Professor in 2007. His research interests include brain–computer interface, cognitive neuroscience, image/signal processing, computer/biologicalvision, and neural networks.

Jun Jiang was born in Zhejiang, China, in 1987. He received the B.Sc. and M.Sc. degrees from the National University of Defense Technology, Changsha, China, in 2009 and 2011, respectively, where he is currently working toward the Ph.D. degree.

His research interests include brain–computer interfaces and machine learning.

Fanglin Chen (M'12) received the B.S. degree of Engineering in automatic control science and technology, and B.S. degree (dual degree) of Science in applied mathematics, from Xi̓an Jiaotong University, Xi̓an, China, in 2006, and the M.S. and Ph.D. degrees from the Department of Automation, Tsinghua University, Beijing, China, in 2008 and 2011, respectively.

He has been on the faculty in the Department of Automatic Control, National University of Defense Technology, Changsha, China. His research interests include brain–computer interface, pattern recognition, classifier fusion, machine learning, computer vision, biometrics, and neuroimage.

Yadong Liu was born in Gansu, China, in 1977. He received the B.Sc. and Ph.D. degrees from the National University of Defense Technology, Changsha, China, in 2000 and 2006, respectively.

His research interests include image/signal processing, computer/biological vision, system identification and control, and cognitive science. He has more than 20 papers and two academic prizes in the areas of his interests.

Dewen Hu (SM'06) was born in Hunan, China, in 1963. He received the B.Sc. and M.Sc. degrees from Xi'an Jiaotong University, Xi'an, China, in 1983 and 1986, respectively, and the Ph.D. degree from the National University of Defense Technology, Changsha, China, in 1999.

From 1986, he was with the National University of Defense Technology. From October 1995 to October 1996, he was a Visiting Scholar with the University of Sheffield, U.K. He became a Professor in 1996. His research interests include cognitive science, brain–computer interface, image processing, and neural networks.

Dr. Hu is an Action Editor of *Neural Networks*.