A High Performance Spelling System based on EEG-EOG Signals With Visual Feedback

Min-Ho Lee, John Williamson, Dong-Ok Won, Siamac Fazli, and Seong-Whan Lee, *Fellow, IEEE*

***Abstract*—In this paper, we propose a highly accurate and fast spelling system that employs multi-modal electroencephalography-electrooculography (EEG-EOG) signals and visual feedback technology. Over the last 20 years, various types of speller systems have been developed in brain-computer interface and EOG/eye-tracking research; however, these conventional systems have a tradeoff between the spelling accuracy (or decoding) and typing speed. Healthy users and physically challenged participants, in particular, may become exhausted quickly; thus, there is a need for a speller system with fast typing speed while retaining a high level of spelling accuracy. In this paper, we propose the first hybrid speller system that combines EEG and EOG signals with visual feedback technology so that the user and the speller system can act cooperatively for optimal decision-making. The proposed spelling system consists of a classic row-column event-related potential (ERP) speller, an EOG command detector, and visual feedback modules. First, the online ERP speller calculates classification probabilities for all candidate characters from the EEG epochs. Second, characters are sorted by their probability, and the characters with the highest probabilities are highlighted as visual feedback within the row-column spelling layout. Finally, the user can actively select the character as the target by generating an EOG command. The proposed system shows 97.6% spelling accuracy and an information transfer rate of 39.6 (±13.2) [bits/min] across 20 participants. In our extended experiment, we redesigned the visual feedback and minimized the number of channels (four channels) in order to enhance the speller performance and increase usability. Most importantly, a new weighted strategy resulted in 100% accuracy and a 57.8 (±23.6) [bits/min] information transfer rate across six participants. This paper demonstrates that the proposed system can provide a reliable communication channel for practical speller applications and may be used to supplement existing systems.**

***Index Terms*—Brain-computer interfaces (BCI), electroencephalography (EEG), electrooculogram (EOG), P300 speller, visual feedback.**

I. Introduction

SPELLER systems [1]–[6] provide a new method of communication for patients who suffer from neuromuscular diseases, such as locked-in-syndrome (LIS) and paralysis. These patients can only control part of their body or the movement of their eyes. For the past two decades, speller systems have seen considerable development within various research fields. Fundamentally, speller systems utilize two branches of technology: eye movement detection and BCI (Brain-Computer Interface).

There are several eye-based spelling systems using electrooculography (EOG) and eye-tracking. EOG measures the voltage fluctuations that result from eye movements. By exploiting these signal changes, EOG can be used to systematically track the eye-gaze direction of the user. Additionally, EOG also conveys highly recognizable information from eyelid movements such as blinking, gazing, and winking. At present, numerous studies have explored the possibility of designing EOG-based Human-Computer Interfaces (HCIs) and have had positive outcomes [7]–[10].

One of the problems with EOG interfaces is the lack of adequate support for a large number of classes due to a limited number of movement patterns. Deng *et al.* [8] proposed an EOG-based HCI system that detected variations of EOG signals for different directions in eye-movements (e.g., horizontal, vertical direction). Their paradigm contained six classes (go/stop and four directions) that could be decoded with an average accuracy of 90%. Usakli and Gurkan [11] proposed an EOG-based virtual keyboard system using four eye directions and double blinking. The spelling speed showed a very promising performance where five random letters were written in about 25s, but the system required great effort by the user because of the repeated EOG movements (two eye-blinks in the best case when choosing the first letter or, sixteen eye-movements per letter in the worst case). This means that the user must make anywhere from 10-80 EOG movements to spell a five letter word. To reduce the workload in EOG tasks, Huang *et al.* [12] and He and Li [13] proposed a one channel EOG system where the symbols were flashed sequentially in

Manuscript received January 16, 2017; revised July 1, 2017, January 27, 2018, and March 2, 2018; accepted March 13, 2018. Date of publication May 21, 2018; date of current version July 6, 2018. This work was supported in part by the Ministry of Science and ICT, South Korea, through the SW Starlab Support Program supervised by the Institute for Information and Communications Technology Promotion under Grant IITP-2015-1107 and in part by the Korean Government under Grant 2017-0-00451: Development of BCI based Brain and Cognitive Computing Technology for Recognizing User's Intentions using Deep Learning. *(Corresponding author: Seong-Whan Lee.)*

M.-H. Lee, D.-O. Won, and S.-W. Lee are with the Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, South Korea (e-mail: mh_lee@korea.ac.kr; wondongok@korea.ac.kr; sw.lee@korea.ac.kr).

J. Williamson is with the Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, South Korea, and also with Seokyeong University, Seoul 02713, South Korea (e-mail: jcw411@gmail.com).

S. Fazli is with Nazarbayev University, 010000 Astana, Kazakhstan (e-mail: siamac.fazli@nz.edu.kz).

Digital Object Identifier 10.1109/TNSRE.2018.2839116

TABLE I
SUMMARY OF THE PREVIOUS WORK ON EEG- AND EEG+EOG BCI SPELLER SYSTEMS

Ref.	Modality	Application	Experimental setup	Approach	result
[19]	EEG+EOG	Wheelchair	4 subjects, 8 EEG+1 EOG ch.	9 possible commands are generated by P300 signal and the BCI command is canceled by EOG signal.	Acc.: 83.4%, ITR: 3.79 bits/min
[20]	EEG+EOG	Wheelchair	4 subjects, 28 EEG+2 EOG ch.	6 directions of moving commands are produced by ERD signal and the EOG used for the stop command.	Acc.: 92%
[21]	EEG+EOG	Robot	13 subjects, 8 EEG+4 EOG ch.	8 commands by P300, 7 commands by EOG signals are combined for robot movement.	Acc.: 98.08%
[22]	EEG+EOG	Exoskeleton	6 subjects, 5 EEG+2 EOG ch.	Combining the ERD signals and eye left/right movement task for exoskeleton control.	TRI-HFT: 50.33(10.2)
[23]	EEG+EOG	Speller	10 subjects, 3 EEG+4 EOG ch.	The layout divided in half. The EEG and EOG used for classifying the target symbol and region, respectively.	Acc.: 90.6%, ITR: 31.2 bits/min
[24]	EEG+EOG	Navigation	4 subjects, 15 EEG+2 EOG ch.	The EOG used to classify one of the three directions and the target symbol was selected by the motor-imagery task.	Acc.: 89.3%, ITR: 24.7 bits/min
[25]	EEG+EOG	Navigation	1 subjects, 8 EEG+2 EOG ch.	The 9 directions divided into dual-monitor. The EOG/EEG used to select the target monitor and direction, respectively.	Acc.: 80%
[26]	EEG+EMG	Speller	10 subjects, 9 EEG+1 EMG ch.	The layout were divided into 4 regions. The SSVEP/EMG used to select the target symbol and the region, respectively.	Acc.: 81.0%, ITR: 83.6 bits/min
[13]	EOG	Speller	8 subjects, 1 EOG ch.	The blink task was decoded by combining of the SVM and waveform detection algorithm in speller layout.	Acc.: 94.4%, ITR: 68.6 bits/min
[12]	EOG	Wheelchair	4 subjects, 1 EOG ch.	The Thirteen buttons were flashed sequentially, and target were chosen by the eye-blinking task.	Acc.: 96.7%, RT: 3.53s
[27]	EEG	Speller	10 subjects, 12 EEG ch.	Visual stimulus was presented with faces, emotion, and movement	Acc.: 91%, PBR: 26.3 bits/min
[28]	EEG	Speller	9 subjects, 12 EEG ch.	The efficient target-to-target interval was applied on the P300 speller.	Acc.: 94.8%, PBR: 29.9 bits/min
[29]	EEG	Speller	15 subjects, 29 EEG ch.	The random set presentation and visual stimuli of face image are applied on the p300 speller.	Acc.: 97%, ITR: 35.2 bits/min
[30]	EEG	Speller	17 subjects, 10 EEG ch.	The visual stimuli of green familiar faces used on the p300 speller .	Acc.: 81.6%, ITR: 39.0 bits/min
[31]	EEG	Speller	15 subjects	Dynamic stopping classification using a hidden markov model with language information.	Acc.: 92.3%, ITR: 42.3 bits/min
Ours	EEG+EOG	Speller	20 subjects, 29 EEG+1 EOG ch.	Real-time visual feedback from the P300 features and select the target character by EOG signal.	Acc.: 97.6%, ITR: 39.6 bits/min
Ours	EEG+EOG	Speller	6 subjects, 4 EEG+1 EOG ch.	Combines our method with the previous approaches [19],[21].	Acc.: 100%, ITR: 57.8 bits/min

a predefined one-by-one manner. The user performed a blinking task to generate a selection when the target symbol was flashed. The decoding accuracies of both systems exhibited a strong performance of 94.4% [12] and 96.7% [13] for the 40 symbols in the speller and the 13 symbols of wheelchair system, respectively (see Table I).

Brain-based spelling systems are less physically demanding than the EOG systems because they use components of the event-related potential (ERP) [9], [10]. ERPs are brain responses to specific cognitive tasks. The P300 is an ERP component with a strong deflection and has been widely investigated over the past few decades. Specifically, the P300 is a positive deflection in the electroencephalography (EEG) signal over the central and parietal cortex. It occurs approximately 300 ms after a task-relevant stimulus is detected. In an EEG-based speller system, the subject mentally attends to a target letter, and the system infers the letter based on the subject's brain response because the target letters will show a stronger P300 response as compared to non-target stimuli. No additional input from the subject is needed. In conventional speller systems, for instance, a letter-matrix composed of alphabetical letters and digits is shown on a screen. The subject attends to the target letter they wish to spell while the matrix is displayed. During this period, rows and columns are flashed consecutively (classic row-column (RC) paradigm [6]). When the intended letter is flashed, an elevated P300 signal can be detected from the subject's brain without the need for any muscle movements.

While EEG-based speller systems do not require any muscle movements, the spelling speed and accuracy is much lower than an EOG-based speller due to the poor signal-to-noise ratio of the EEG signal. Due to the low signal-to-noise ratio in single-trial EEG signals, the detection of target symbols from a single trial is very difficult. Therefore, the target and non-target symbols must be flashed several times for each choice. The epochs corresponding to each target and non-target trials are then averaged over time in order to improve the spelling (decoding) accuracy. Consequently, the typing speed depends on how many trials are averaged for a desired spelling accuracy, thereby the development of fast and accurate spelling systems is the most important factor for ERP spelling studies.

Additional BCI research has investigated the utility of visual feedback which is a subtype of biofeedback (or neurofeedback) that uses real-time displays of brain activity so that subjects may perform self-regulation of their brain signals. It has been commonly employed in clinical therapies for Attention Deficit Hyperactivity Disorder (ADHD) [14], stroke [15], and user training in BCIs [16]. Previous studies have shown that visual or auditory feedback positively impacts typing speed, accuracy, and subjective experience [17].

Given the advantages and drawbacks of the two core technologies, EOG and EEG, combining the two devices into a 'hybrid' BCI system can maximize typing speed and decoding accuracy. Usakli *et al.* [18] compared the spelling performance of the EOG and EEG modalities independently. The EOG and EEG systems required 24.7 s and 105 s, respectively, for writing five letters. At the time, the authors suggested the concept of a hybrid EEG-EOG system in furtherance of efficient human-computer communication by using the characteristic signals of the EEG and EOG systems (e.g., signal waveforms, features, required practice).

Lee *et al.* [32] proposed a hybrid spelling system that combines EEG signals with a webcam-based eye-tracking system. A QWERTY keyboard layout was presented on a monitor divided into three sections, corresponding to the right, left, and middle fields of vision. The information obtained from eye-gaze direction was then used to prevent any errant spellings by the system. In short, this hybrid system focused on the reduction of errors in EEG-based spellers. The authors reported a reduction in the error rate by 40%. Similar approaches where the visual stimulus were divided into a number of sections with the EOG and EEG modalities used to classify the target section and the symbol, respectively, have been proposed based on the dual-monitor [25] and half checkerboard paradigms [23].

Lin *et al.* [26] proposed a hybrid BCI system based on the steady-state visually evoked potential (SSVEP) and electromyography (EMG). The speller layout was divided into four sections with the same frequency set. The target symbol was classified by the SSVEP signal in the selected section which was classified using the EMG signals from the forearm. The classification accuracy and ITR at 81.0% and 83.6 bits/min was a significantly enhanced performance compared to the single modality.

Other types of hybrid BCI systems have used an EOG signal as the main modality for decoding a target symbol [21] or independently utilized each modality to extend the available number of classes [24]. Ma *et al.* [21] proposed an EOG-EEG hybrid HCI interface for robotic controls. Seven eye movement patterns and five ERP targets were classified (12 classes in total) and used to control robot behaviors such as go/stop, turn, and dance among others. Results from the thirteen subjects indicated an EEG decoding accuracy rate of 70% after the first sequence and an EOG accuracy of 92%. Jiang *et al.* [24] proposed an EOG+event-related desynchronization (ERD)-based BCI system in a three target selection interface. The system first classified a target symbol by estimating the direction of the eyes based on the EOG signal, then the system executed a command when the user generated ERD components by performing a motor-imagery task. The accuracy and completion time for target selection showed 89.3% and 24 s, respectively.

Eye-based speller (EOG or eye tracker) systems are sourced from eye movements. These signals show significant patterns that can lead to high classification accuracy, but the system requires a high level of voluntary eye movement to control a target. ERP spellers, on the other hand, do not require any specific movements by the user but suffer from relatively low classification accuracy due to small signal amplitudes and a low signal-to-noise ratio (SNR) in the EEG. While it is possible to increase the accuracy of ERP spellers with visual feedback, it is often applied merely as a confirmation of the decision reached by the system.

While there have been numerous types of EEG-EOG approaches in hybrid BCI systems, the target applications have mostly been in robotics systems such as wheelchair [19], [20], robot arm [21], or exoskeleton [22] control. Details of these hybrid EEG-EOG BCI systems are included in Table I.

To the best of our knowledge, this study is the first to propose an EEG-EOG hybrid BCI speller system with the novel idea that the subject can actively select the target letter by using the visually informed classification result. The system presents a brain-based classification score for the candidate character to inform the subject about the current prediction of the system. If the ongoing classification is correct, the human actively selects the target letter through a simple movement task, specifically an eye wink in our case. We term this paradigm row-column visual feedback (RC-VF).

By adding active human feedback to the system, the required quantity of brain signal samples for reliable classification is minimized. As a result of the system design, the system performance, namely accuracy and spelling speed, has greatly increased. At the same time, only minute eye activity is required. The average classification accuracy and information transfer rate (ITR) across 20 participants were found to be 97.6% and 39.6 ($\pm$13.2) [bits/min] so 8.2 s were required for the spelling of each character on average.

Additionally, a second experiment was designed to enhance the spelling performance and to increase convenience as a real-world application by applying the proposed RC-VF method with other recent advances in ERP spelling systems. The purpose of the experiment was to maximize the performance of our novel ERP spelling system while demonstrating that our system could be combined easily with previous approaches that have led to additional performance increases. First, we modified the presentation matrix type (e.g. color, size, face image, rate, and motion) as found in [27], [33], and [34]. Secondly, we varied the original experimental setup in regards to electrode montages, inter-stimulus interval (ISI), stimulus onset asynchrony (SOA), and target-to-target interval (TTI) [28], [35]. By reducing the number of channels to only 4, we increased convenience and practicality, and by using an alternative EOG signal for classification, we increased comfort and ease of use. Most importantly we modified the classification strategy by the number of allowed sequences in order to guarantee a promising performance. The results show 100% spelling accuracy and 57.8 ($\pm$23.6) [bits/min] ITR corresponding to 6.2 s for spelling a character across six participants. These results compare very favorably to other current state-of-the-art ERP spelling devices [28]–[31], [36]–[38].

II. Materials and Methods

A. Participants and Data Acquisition

Twenty healthy subjects (aged 24-32, 12 males, 8 females) participated in the experiment. None of the participants had a previous history of psychiatric, neurological, or other pertinent diseases. Participants were seated 60 ($\pm$5) cm in front of a 19" LCD monitor (60 Hz refresh rate, 1280$\times$1024 resolution) on which the stimuli were presented. EEG was recorded with 24 channel Ag/AgCl actiCAP EEG and EOG electrodes (Brain Products, Germany) at a sampling rate of 500 Hz. The following were the measured channels according to the

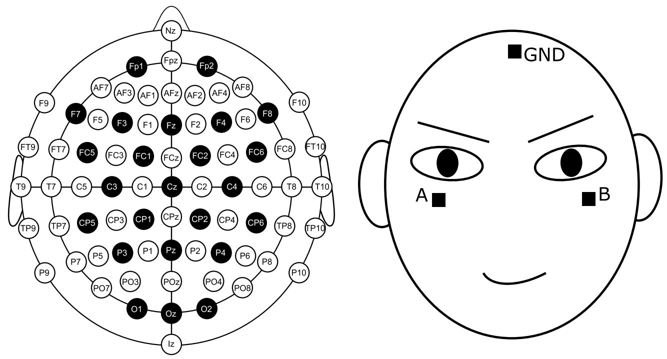

Fig. 1. EEG and EOG channel locations. International 10-20 system (24 EEG electrodes (black circles), reference: FCz, ground: Fpz, A: EOG1, B: EOG2).

International 10-20 system: Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, C3, Cz, C4, CP5, CP1, CP2, CP6, P3, Pz, P4, O1, Oz, O2, EOG1, and EOG2 (see Figure 1). EOG electrode configurations were as shown in Figure 1, with two bipolar electrodes Fp2-A (EOG1) and Fp1-B (EOG2) used for recording movements of each eye. EOG signals were then down-sampled to 32 Hz with a 5*th* order digital Chebyshev filter.

B. Experiment Stimuli and Paradigm

The interface layout of the speller followed the original design of the RC speller [6]. Thirty-six character symbols ('A'-'Z', '1'-'9', '_') were equally spaced on screen in a configuration of 6 rows and 6 columns (see Figure 2-A). Individual columns/rows of characters were flashed sequentially. A single iteration of flashing all rows and columns was considered a sequence. As such, each sequence consisted of 12 flashes: 6 flashes from rows and 6 flashes from columns. For each selection of a target character, a maximum of ten sequences (i.e. 120 flashes) was allotted without prolonged inter-sequence intervals. The paradigm code was developed with the Psychophysics Toolbox (http://psychtoolbox.com) and OpenBMI [39], [40] in Matlab (MathWorks; MA, USA).

Across all spelling conditions, each line of characters was flashed for 200 ms, followed by an inter-stimulus interval (ISI) of 50 ms corresponding to 3 s for one sequence (250 ms$\times$12). The paradigm was divided into two phases: the training phase and the test phase. The training phase was composed of two sessions; one was for calibrating the EEG-based classifier, and the other was for estimating the parameters of the EOG detector for the subject. The test phase was divided into four sessions: two RC-VF sessions, then one RC session, and then a final free spelling RC-VF session.

The two RC-VF sessions and one RC session were copy-spelling procedures that prompt the user to pay attention to a pre-defined letter sequence which was the same as the calibration session. The copy-spelling sessions were done to compare the performance differences between the conventional RC speller and proposed RC-VF speller. The additional free spelling session with the RC-VF speller was conducted to investigate whether the proposed RC-VF system successfully works in a free spelling environment where environmental noise and task difficulty could affect the result.

C. Offline Training Phase: RC Spelling

In the EEG calibration session of the training phase, subjects were instructed to relax their muscles and minimize movements. Subjects were tasked to spell the given phrase, 'BRAIN_COMPUTER_INTERFACE' (24 characters including spaces) by gazing at the target characters on the screen. Throughout the session, the task sentence was displayed in the top center area of the screen with the target letter emphasized, and the speller was displayed in the area below (see Figure 2-A). The training phase was done offline, and no feedback was given to the subject while the EEG signals were collected. Subjects were instructed to count the number of times each target character had flashed. For the offline data analysis of the training data, EEG and EOG data were band-pass filtered between 0.1 and 25 Hz with a 5*th* order Butterworth digital filter. The EEG data were segmented from -200 to 800 ms with respect to stimulus onset. Then, baseline-correction was applied by subtracting the mean amplitudes in the -200 ms to 0 ms pre-stimulus interval from all segmented data. Eight discriminant intervals were extracted with a heuristic based on signed r-values [41], [42]. Mean amplitudes were calculated and used as ERP features. These subject-dependent spatio-temporal feature vectors were therefore formed with 192 dimensions (i.e., 24 channels$\times$8 features). After that, the linear discriminant analysis (LDA) [43] classifier was trained using these feature vectors.

$$\mathbf{w} = \text{argmax}_{\mathbf{w}} \frac{\mathbf{w}^{\mathrm{T}} \mathbf{S_B} \mathbf{w}}{\mathbf{w}^{\mathrm{T}} \mathbf{S_w} \mathbf{w}} \tag{1}$$

where $\mathbf{S_B}$ and $\mathbf{S_w}$ denote the between-class and within-class covariance matrices, respectively and $\mathbf{w}$ the hyper-plane for separation of the classes, which can be obtained by maximizing the Rayleigh coefficient.

$$f(\mathbf{v}) = \mathbf{w}^{\mathrm{T}} \cdot \mathbf{v} + b \tag{2}$$

where b is a bias term and $\mathbf{v}$ is the ERP feature vector containing EEG data for a single trial $\mathbf{x}$. $f(\mathbf{v})$ is the classification output from the single trial's feature vector.

D. Offline Training Phase: EOG Calibration

EOG electrodes were placed as shown in Figure 1. EOG1 recorded the winking of the right eye while electrode EOG2 recorded the winking of the left eye. Subjects were tasked to choose the eye they were more comfortable winking with and then wink thirty times with the chosen eye. In this study, amplitude values acquired during the process of winking were then used to determine the subject-dependent threshold for detecting wink executions [8], [21]. All subjects were instructed to wink 30 times following a cue with the chosen eye. Thirty trials of EOG signal were then corrected and averaged to determine subject-specific threshold values for winking detection. The maximum and minimum amplitude value (i.e., th_{min} and th_{max}, respectively) were heuristically determined based on the peak of averaged EOG signals and its standard deviation (STD) (see Figure 3). The values of the th_{min} and th_{max} were selected at a level not exceeding

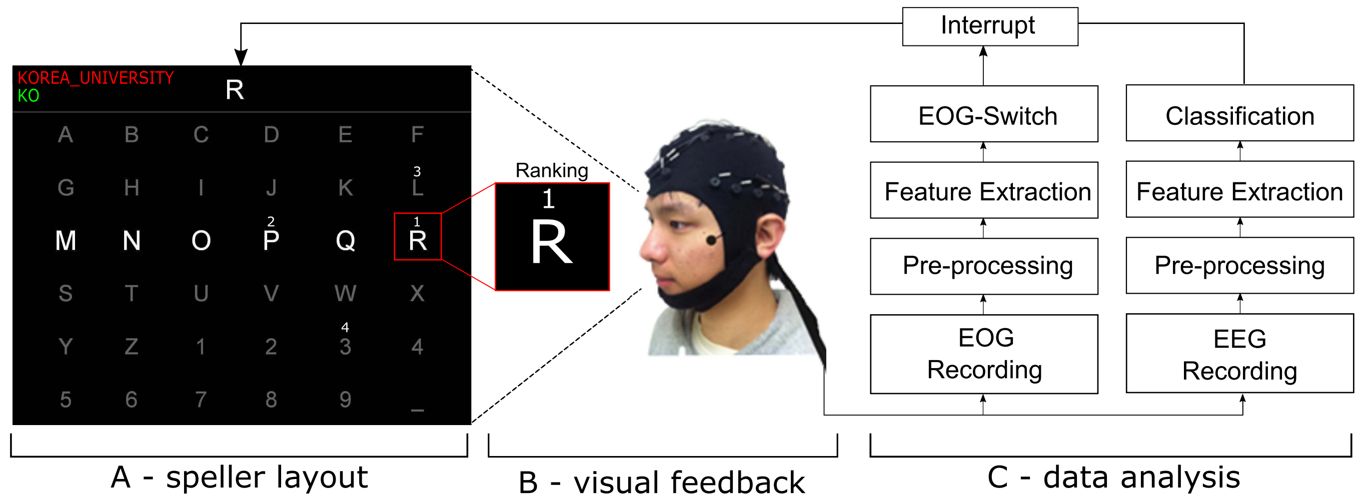

Fig. 2. RC-VF speller system architecture. A: design of the row-column speller B: expansion of the target character with ranking score as visual feedback C: data analysis of EEG and EOG signals.

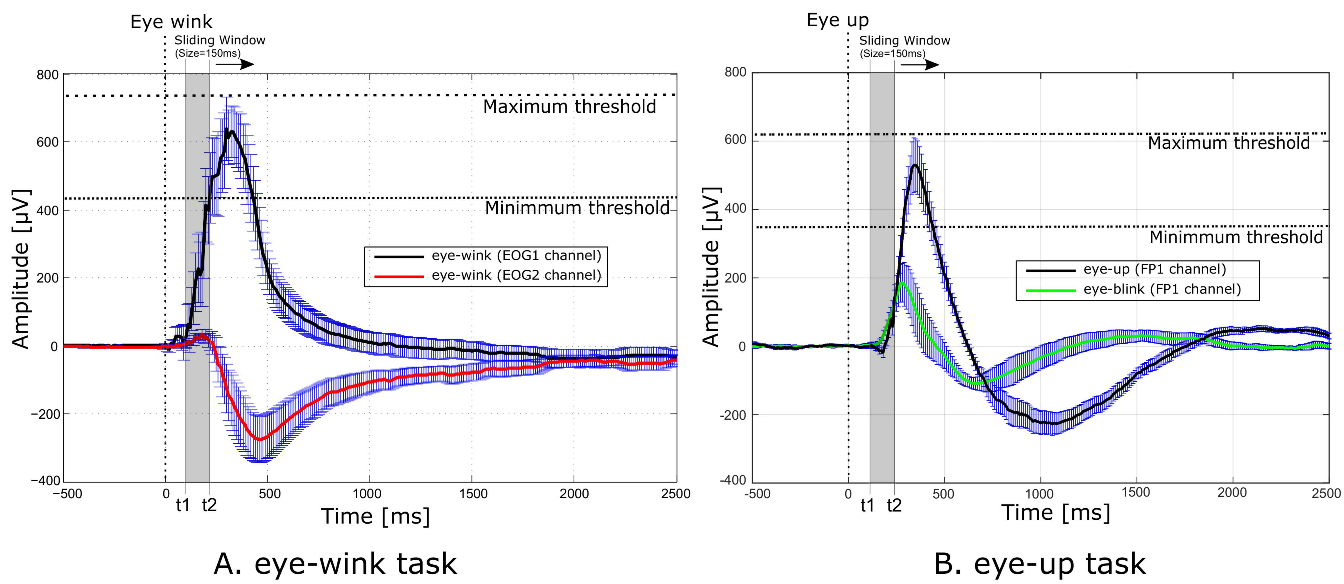

Fig. 3. Procedure for EOG detection. In the left plot, black and red lines correspond to the averaged EOG signals from the EOG1 and EOG2 channels during the winking task. In the right plot, the black line indicates the averaged EOG signals for the eye-up task from the FP1 channel, and the green line indicates the averaged EOG signals for the natural eye-blink task from the same channel of FP1. Standard deviation of the signal is given by blue bars. The gray-patch indicates a sliding window (size of 150ms) along with the time courses of EOG signal.

two standard deviations from the peak of averaged EOG signals. After determining the threshold parameter, subjects were asked to perform various strong head movements and blinking to check whether the EOG detector was working properly. For most subjects, the threshold values lied within the 200-400 μV amplitude range.

To detect the winking task in the real-time environment, a sliding window (size of 150 ms) was created (see Figure 3), and the EOG detector generated a command every 0.05 s (i.e., step size=50ms). Assume that $x_n(t)$ is original EOG$_n$ data at a time point t, and n ($n = 1$ or 2, corresponding to the EOG1 and EOG2) is the subject's selected channel. The mean amplitude value ($\mathbf{mean}\,[E]_{t1}^{t2}$) within the time intervals of $t1$ and $t2$ is calculated as an EOG feature. Finally, the EOG detector was generated when the value of $\mathbf{mean}\,[E]_{t1}^{t2}$ was ranged in the th_{min} and th_{max}.

$$th_{\mathbf{min}} \le \mathbf{mean}_{t\in\langle t_1,t_2\rangle} E(x_n) \le th_{\mathbf{max}} \tag{3}$$

Please note, that winking and blinking may have a similar signature in the EOG signal to some degree. However, when winking these changes occur predominantly in one EOG channel only. Furthermore, blinking is a relatively short implicit process as compared to conscious winking, which leads to stronger and more prolonged activations in EOG (see Figure 3). Afterward, subjects were given a short online practice session (<2min), so that they could familiarize themselves with using the on/off control of the EOG detector.

E. Online Test Phase: RC Spelling

In the RC session of the online test phase, subjects were instructed to spell 'KOREA_UNIVER-SITY' (16 characters including space) using the RC speller. The target sentence was colored with red and presented in the top left corner of the screen. During online data analysis, the real-time data was acquired from the EEG amplifier with a down sampling rate of 100 Hz. The continuous EEG data was then segmented from −200 to 800 ms with respect to the stimulus presentation and band-pass filtered in the same frequency band as the offline data analysis (0.1 to 25 Hz). Baseline correction and ERP feature extraction were applied, and the classification output was calculated using the LDA classifier, which was estimated previously from the training data set.

In performing RC spelling, the highest scored character was estimated as the target in individual sequences. The decoding accuracy was therefore validated in each sequence (i.e., one to ten sequences). After all ten sequences, the final result for the target character was estimated by the averaging of epochs accumulatively through the sequences. The estimated target character was colored with green and displayed on the top left area of the screen as online feedback (Figure 2-A).

F. Online Test Phase: RC-VF Spelling

During RC-VF sessions the conventional RC-speller was combined with sequential visual feedback. After the end of the first sequence (and the end of all the following ones), the four letters with the highest selection probability were highlighted with a small number (see Figure 2-B). If the desired letter was labeled with a '1', the subject was able to intervene and select the desired letter by winking. Therefore, the RC-VF system is divided into three submodules of 1) EEG classification, 2) visual feedback, and, 3) EOG detection modules (see Algorithm I).

Given EEG epochs, x_i^j where the i ($i = 1, \ldots, 12$) is number of trials corresponding to 12 flashes (6 row+6 column) in the sequence j ($j = 1, \ldots, 10$), the ERP feature (v_i^j) sets were also calculated.

$$\mathbf{x}_i^j = \left[x_1^j, x_2^j, \ldots, x_i^j\right] \tag{4}$$

After the first sequence (i.e., $j = 2$), the averaged ERP features $\bar{\mathbf{v}}_i$ were calculated by accumulatively averaging the current feature vector with the correct set of previous trials. Please note that the RC and RC-VF system followed the same feature averaging approach where the ERP features for target and non-target trials were averaged accumulatively over time to achieve a reliable result.

$$\bar{\mathbf{v}}_i = \frac{1}{n}\sum_{j=1}^{n} \mathbf{v}_i^j \tag{5}$$

Where n represents currently available number of sequences during spelling and $\bar{\mathbf{v}}_i$ indicates averaged ERP features. The LDA classifier output $Y_i = f(\bar{\mathbf{v}}_i)$ is then calculated from the input feature vectors $\bar{\mathbf{v}}_i$ for each trial i.

$$Y_i = \mathbf{w}^{\mathrm{T}} \cdot \bar{\mathbf{v}}_i + b \tag{6}$$

Where $Y_{i,i=(1,\ldots 12)}$ denotes the classification outputs for 6×6 matrix positions; 6 rows $Y_{(1,\ldots,6)}$ and 6 columns $Y_{(7,\ldots,12)}$. The classification scores for the 36 characters were calculated by summation of Y_{row} and Y_{column} according to each character's row/column positions. For instance, the classification score of the character 'K' was calculated by summation of Y_2 and Y_5.

The 36 letters were then ranked, based on their classification scores (rank_p, and $p = 1, \ldots, 36$). In other words, the letter with the highest classification score appeared in the first position of the rank variable ($= \mathrm{rank}_1$). The classification scores were updated on a single trial basis, i.e. after every row/column flash.

Once the first sequence was completed, the four letters with highest classification outputs (i.e., $= \mathrm{rank}_{(1,2,3,4)}$) were chosen and visualized on the screen. The ranking was indicated by the superscripts 1/2/3/4 above the characters (Figure 2-B). While the rank of all 36 letters was calculated on a single trial basis, the ranking was visibly updated twice per sequence (i.e. every 6 flashes, which corresponds to 1.5 s).

If the classifier denoted the target letter as the top candidate ($= \mathrm{rank}_1$) at any given time, subjects were instructed to actively wink once. This action would be recognized as a selection of the candidate letter; the selection command was generated if the EOG signal satisfied the predefined stopping criterion (i.e., equation 3), and the spelling of the next character would begin (See Algorithm 1 for the detailed procedures).

Similar to the standard RC speller, the RC-VF speller was limited to a maximum of 10 sequences to select each character. In the event that the classifier was unable to mark the intended character as number '1' after 10 sequences, the highest ranked character would be automatically chosen, as is done in the conventional RC speller. Subjects were instructed to spell 'KOREA_UNIVERSITY' throughout the session. The spelling result of each character was displayed on the top left of the screen in the same fashion as in the RC speller paradigm.

The multi-score visual feedback was dynamically presented above the characters. The decision to display the rankings of the characters with the four highest scores was made based upon the high level of competition observed in the selection procedure (see Table IV).

Specifically, high levels of competition mostly occurred between the first few highly ranked characters (typically ranked 1 to 6 as seen in pseudo-online emulation) in the early stage of the sequential procedure. For instance, Table IV indicated that most of the target characters were ranked in the top 3 (except 'A', and 'R') at the end of second sequence. Our decision to display the top 4 was due to the need to minimize visual complexity while maintaining a high probability that the target character would be displayed with some ranking in the early stages of the sequential progression. We displayed four visual scores as this number satisfied those conditions, but this parameter could be flexibly adjusted in consideration of the experimental environment or user's preference.

Algorithm 1 RC-VF Speller Procedure

Input: EEG epochs (x), LDA parameters $(\mathbf{w}, b)$

- $X=\{x_i\}_j$, $i = (1, ..., 12)$, $j = (1, ..., 10)$, EEG epochs (x) where the i and j denote the number of flashes (6 rows+6 columns) and sequences, respectively.
- $\mathbf{mean}\,[E]_{t1}^{t2}$, Mean amplitude within the time intervals of $t1$ and $t2$ from the EOG channel.

Output: Selected target character: $rank_1$

- $rank_p$, p is ranking index $(p = 1, 2, ..., 36)$ by sorting the classifications scores for all 36 characters .

Procedures:

```
while speller on until stopping criterion is satisfied do
    EEG/EOG signal acquiring
    if EOG signal is not generated then
        v̄_i = feature(x_i^j),                               ▷ Accumulatively averaging ERP features
        Y_i = w^T · v̄_i + b,                                 ▷ Classifier output for input feature vector v
        𝕐_k = [(Y_1 + Y_7), (Y_1 + Y_8), ..., (Y_6 + Y_12)], k = (1, ..., 32)   ▷ Classification scores
        rank_p = sort(𝕐), p = (1, ..., 32)                   ▷ Ranking index p for all 36 characters
        rank_p^old = rank_p^new                              ▷ Update rank scores on every iteration
    end
    if half or end of sequence then
        visualfeedback(rank_(1,2,3,4));                      ▷ Present visual feedback
    end
    if EOG buffer full then
        if th_min ≤ mean[E]_t1^t2 ≤ th_max then
            interrupt;                                       ▷ Stopping criterion satisfied
            selection(rank_1)                                ▷ Target character is selected
        end
    end
end
```

G. Online Test Phase: Free RC-VF Spelling

In the free spelling RC-VF session, subjects were given the liberty to decide what they wanted to spell. For instance, subject C spelled the sentence 'WELCOME_TO_SEOUL'. In this session, the sentence and target characters were not presented at the top of the monitor, only the characters selected by the user were displayed.

III. Performance Evaluation

For evaluation of the two types of spelling systems, the classification accuracies, as well as information transfer rates (ITRs), were computed across the 20 subjects. ITRs are commonly used as an evaluation measurement for BCIs. The unit of ITRs is given as bits per unit time [bits/min] and can be calculated as:

$$\mathrm{ITR} = M\left\{\log_2 N + P\log_2 P + (1-P)\log_2\left(\frac{1-P}{N-1}\right)\right\} \quad (7)$$

where M denotes the number of commands per minute and N indicates the number of possible choices. Each choice N has an equal probability of being selected by the user. P is the accuracy of the speller (i.e. the probability that the speller selects the desired letter). In summary, the ITR corresponds to the amount of information received by the system per unit time [44]. A paired t-test (significance level: $p < 0.05$) was computed to examine the performance improvement between the RC speller and the RC-VF system for accuracy and ITR. The statistical tests were performed with the null hypothesis of equal means. For the performance evaluation of the free RC-VF spelling system, accuracy, ITR, and spelling time of each letter were calculated.

IV. Results

Figure 4 depicts the classification accuracy for individual subjects along with their averages. Averaged ITRs are shown across all subjects. The decoding accuracies (y-axis) are calculated with respect to the number of sequences (x-axis). The number of sequences varied from one to ten. Results for the two conditions, namely the RC speller (blue square boxes) and the RC-VF speller (red square box), are shown. The performances of both RC-VF sessions were averaged. It is important to note that the proposed RC-VF system selects the target character by the generated EOG signal; each target character was chosen at a different time-point within the sequential procedure. The red square box and the red line indicate the mean of the sequential time-point and its standard deviation, respectively.

Detailed results for both conditions can be obtained from Table II. The accuracy and ITRs for each subject as well as their averages are reported for the RC condition and

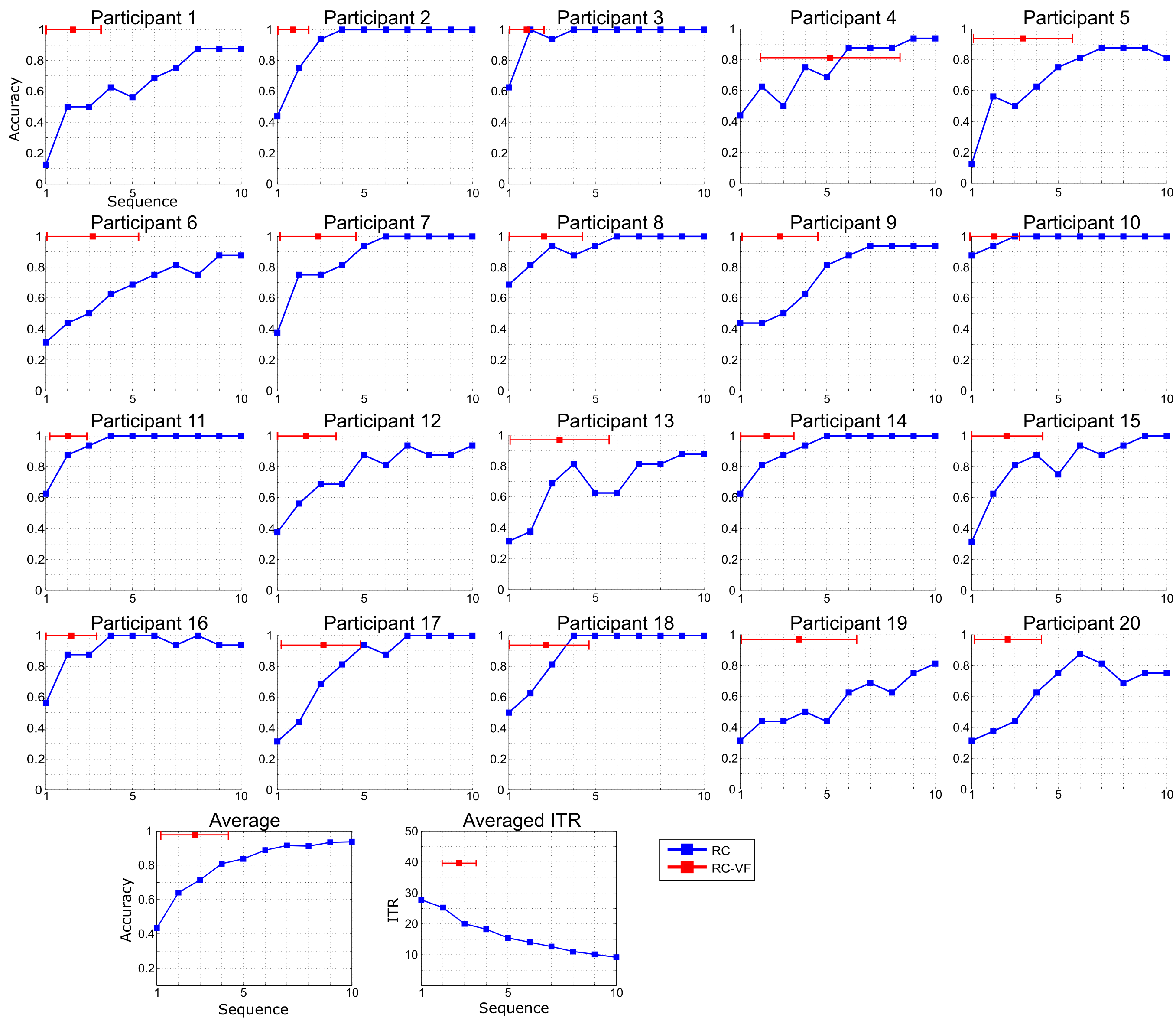

Fig. 4. Classification accuracy curve for each subject. Two conditions (RC, RC-VF) were evaluated. Average spelling performance (y-axis) was calculated for one to ten sequences (x-axis). Averaged classification accuracy and ITR is shown in the lower right.

both RC-VF sessions. Previous BCI studies have argued that a minimum accuracy level of 70% is needed for efficient communication [45]. To reach an accuracy level of 70% in the RC condition, 3 sequences would need to be averaged, which would take 9 s per letter. For an accuracy level of 90%, 7 sequences would need to be averaged (21 s per letter). In the two sessions of the RC-VF condition, the average accuracies were 97.5% and 97.8%. An average spelling time of 8.5 s ($\pm$3.0) (or 2.9 sequences) per character and 7.8 s ($\pm$2.0) (2.5 sequences) per character were required. In this study, subject B exhibited the best performance: only 5.1 s were needed to spell each character while maintaining perfect, 100% accuracy.

The Average ITRs of the two RC-VF speller sessions were 38.6 ($\pm$12.5) bits/min and 40.6 ($\pm$12.8) bits/min. The average ITR for the RC speller was 25.5 ($\pm$18.9) bits/min, if only one sequence is averaged. If more sequences are averaged, the ITR decreases (see Figure 4). Paired t-tests did not reach a statistical significant level for the differences of ITR between RC-VF sessions one and two (p=0.077).

In order to simplify the comparison of results between the RC and RC-VF conditions, averaged performances of the two copy-spelling RC-VF sessions were calculated and subsequently used for statistical comparison. For the RC-VF speller, the subject was able to actively stop the presentation process at any time, so the sequence numbers varied across trials. It is therefore not possible to match the number of sequence in the RC-VF condition to those of the RC condition. An average of 2.9 sequences are required in the RC-VF condition; therefore, the decoding accuracy and ITR in the sequences of 2, 3, and 5 which are close to the sequence numbers in the RC-VF (i.e., 2.9 sequence) were chosen to compare system performance. The RC-VF speller required 2.9 sequences to reach an average accuracy of 97.6%, while

TABLE II

hl CLASSIFICATION ACCURACY AND ITR OF RC AND TWO SESSIONS OF RC-VF SPELLER. SEQ. STANDS FOR THE NUMBER OF AVERAGED SEQUENCES (1 SEQUENCE=3S). THE AVERAGE OF 2, 3, 5 AND 10 SEQUENCES ARE CONSIDERED FOR THE RC SPELLER. THE CALCULATION OF THE ITR FOR THE RC SPELLER WAS PERFORMED BY AVERAGING 1 OR 2 SEQUENCES (LAST TWO COLUMNS)

Subject	Accuracy [%]						ITR [bits/min]			
	RC				RC-VF01	RC-VF02	RC		RC-VF01	RC-VF02
Seq.	2	3	5	10	2.9	2.5	1	2	2.9	2.5
A	50.0	50.0	56.3	87.5	100	100	16.1	7.7	43.4	48.6
B	75.0	93.7	100	100	100	100	30.8	20.7	60.0	60.3
C	100	93.7	100	100	100	100	51.7	20.7	46.2	71.9
D	62.5	50.0	68.8	93.8	75.0	87.5	22.9	10.7	10.3	18.5
E	56.3	50.0	75.0	81.3	87.5	100	19.4	12.3	19.3	39.6
F	43.8	50.0	68.8	87.5	100	100	13.0	10.7	39.2	28.0
G	75.0	75.0	93.8	100	100	100	30.8	18.0	34.4	37.5
H	81.3	93.7	93.8	100	100	100	35.1	18.0	36.2	43.0
I	43.8	50.0	81.3	93.8	100	100	13.0	14.0	40.7	33.0
J	93.8	100	100	100	100	100	45.1	20.7	51.8	48.8
K	87.5	93.7	100	100	100	100	39.9	20.7	44.4	58.9
L	56.3	68.7	87.5	93.8	100	100	19.4	15.9	42.4	46.7
M	37.5	68.7	62.5	87.5	100	93.8	10.1	9.2	27.5	30.9
N	81.3	87.5	100	100	100	100	35.1	20.7	55.4	40.0
O	62.5	81.2	75.0	100	100	100	22.9	12.3	39.5	39.6
P	87.5	87.5	100	93.8	100	100	39.9	20.7	50.3	45.2
Q	43.8	68.7	93.8	100	93.8	93.8	13.0	18.0	28.0	29.6
R	62.5	81.2	100	100	93.8	93.8	22.9	20.7	31.2	35.2
S	43.8	43.7	43.8	81.3	100	93.8	13.0	5.2	24.1	28.8
Y	37.5	43.7	75.0	75.0	100	93.8	10.1	12.3	47.4	28.5
Mean	64.1	71.5	83.8	93.8	97.5	97.8	25.2	15.5	38.6	40.6

TABLE III

CLASSIFICATION ACCURACY AND ITR FOR FREE SPELLING WITH RC-VF. EACH SUBJECT SELECTED THEIR OWN SENTENCE TO SPELL. THE CORRECTLY SPELLED LETTERS ARE COLORED BLUE WHILE MISSPELLED LETTERS ARE COLORED RED

Sub.	Target Sentence	Estimated Text	Acc. [%]	ITR [bits/min]	Time [s]/letter
A	'PEACE_AND_LOVE'	'PEACE_AND_LOVE'	100	47.2	6.6
B	'WE_ARE_THE_WORLD'	'WE_ARE_THE_WORLD'	100	47.4	6.5
C	'WELCOME_TO_SEOUL'	'WELCOME_TO_SEOUL'	100	66.0	4.7
D	'MACHINE_LEARNING'	'MA**O**HINE_**X**EARNING'	87.5	17.7	13.5
E	'LOVE_AND_WAR'	'LOVE_AND_WAR'	100	29.2	10.6
F	'HAPPY_HALLOWEEN'	'HAPPY_HALLOWEEN'	100	29.9	10.4
G	'HELLO_WORLD'	'HELLO_WORLD'	100	46.7	6.6
H	'GOOD_MORNING'	'GOOD_MORNING'	100	49.8	6.2
I	'MERRY_CHRISTMAS'	'MERRY_CHRISTMAS'	100	36.6	8.5
J	'CAT_IS_CUTE'	'CAT_IS_CUTE'	100	39.1	7.9
K	'RECYCLE_BIN'	'RECYCLE_BIN'	100	49.1	6.3
L	'FLASHLIGHT'	'FLASHLIGHT'	100	45.4	6.8
M	'KIM_SEON_MIN'	'KIM_SEON_MIN'	100	31.5	9.9
N	'WE_WANNA_PEACE'	'WE_WANNA**G**PEACE'	92.8	29.3	9.1
O	'HAPPY_NEW_YEARS'	'HAPPY_NEW_YEARS'	100	28.7	10.8
P	'SEOUL_WELCOMES_YOU'	'SEOUL_WELCOMES_YOU'	100	32.3	9.6
Q	'SUNG_SOO_KIM'	'SUNG**4**SOO_KIM'	91.6	23.3	11.2
R	'NEUROSCIENCE'	'NEUROSCIENCE'	100	49.7	6.2
S	'VERY_TIRED'	'VERY_TIRED'	100	26.9	11.5
Y	'MENTAL_SPELLER'	'MENT**M**L_SPELLER'	92.8	26.6	10.0
Mean			98.2	37.6	8.7

the RC speller showed accuracies of 64.1%, 71.5%, 83.8% and 93.8% for 2, 3, 5, and 10 sequences, respectively.

RC-VF accuracy was significantly higher than RC accuracy at 2, 3, and 5 sequences ($p < 0.001$) but only marginally if 10 sequences were averaged in the RC condition ($p = 0.0502$). The RC-VF condition showed significant higher ITRs for all sequences (i.e., one to ten) compared to the RC condition ($p < 0.001$).

Table III depicts the spelling performance of the RC-VF system during free spelling. Most of the subjects were able to successfully spell their sentence without making any mistakes. Four subjects (D, N, Q, Y) did not reach 100% accuracy. The average accuracies across subjects were 98.2%, and the ITR was 37.6 (±12.2) bits/min. The average typing speed was 8.7 (±2.5)s/letter. Subject C showed the best performance. He required 75.2 s (4.7 s per character)for writing the sentence 'WELCOME_TO_ SEOUL' (16 characters including space) with 100% accuracy.

Figure 5 shows the relation of single-trial classification performance (x-axis) to the spelling accuracy (y-axis) for real-time RC (black dots) and RC-VF (red dots). Each dot represents the results of a single subject. To calculate single-trial classification performance, 8-fold chronological cross-validation was performed using offline training data.

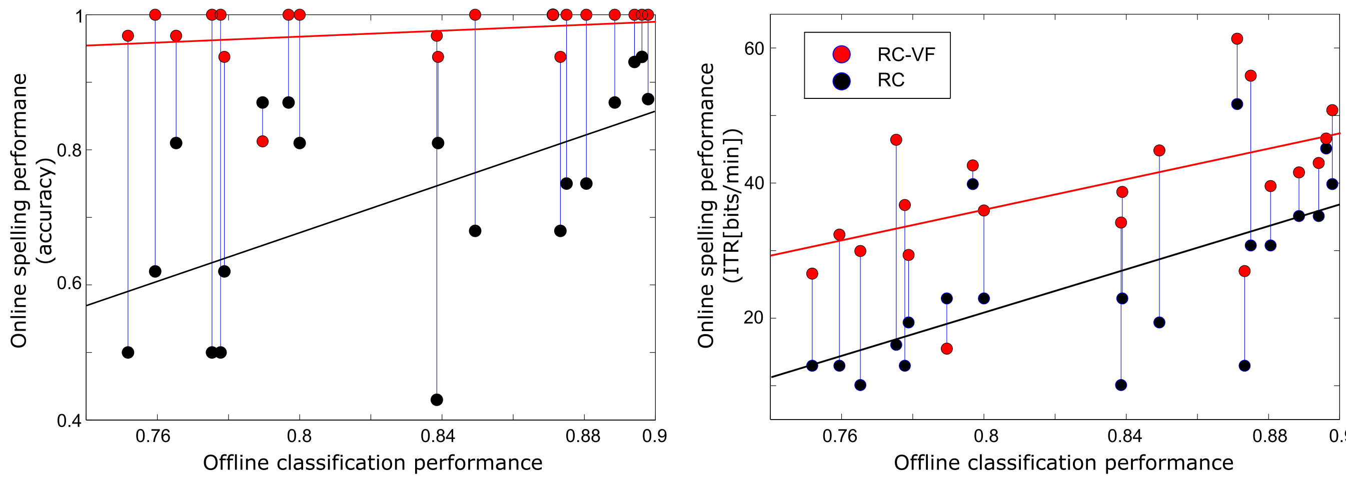

Fig. 5. The relationship between offline single-trial classification performance (x-axis) on the training data using 8-fold cross-validation and online spelling performance (y-axis). Accuracy (left) and ITR (right) are shown for each of the 20 subjects and the two conditions (red dots RC-VF; black dots RC). Red and black lines depict a linear fit for each paradigm (RC-VF and RC).

The spelling accuracy and ITR were calculated from the online RC and RC-VF data. To ensure maximal comparability, the spelling accuracy and ITR were calculated with the same number of sequences for RC and RC-VF data. For instance, if subject B required 1.8 sequences on average to spell in the RC-VF condition, 2 sequences were used for the calculation of the accuracy and ITR in the RC condition (accuracy: 75%, ITR: 20.7 bits/min; see Table II). A linear regression model was calculated for each condition (red and black lines). In both plots, the RC speller showed strong positive correlations between the offline classifier performance and online spelling performance (accuracy: RC=0.87; ITR: RC=0.68). The correlation coefficient for the RC-VF condition was considerably lower for the left plot (accuracy: RC-VF=0.34; ITR: RC-VF=0.61).

V. Discussion

A. EOG Performance

In this study, all of the subjects were able to generate an EOG-based selection task intentionally in the RC-VF spelling paradigm. The EOG signals are a reflection of muscle activity, which has a comparatively high signal-to-noise ratio (see Figure 3). Therefore, it is not very surprising to observe perfect discrimination between the binary classes of 'winking' and 'resting'. Similarly, previous studies were able to reach accuracy rates of more than 90% in a four-class paradigm based on EOG data [7], [8].

We considered two types of EOG errors in the RC-VF experiment. The first one is that the subject unintentionally generates the EOG movement at the wrong time (i.e., idle state); therefore, the system selects the unintended character. In the RC-VF system, the EOG detector was constantly activated (i.e. asynchronous system). A type-one error should be prevented as it would result in the selection of an undesired letter. It could happen when the EOG threshold is too low– a malfunction of EOG discrimination between eye-winking and a normal eye-blink (or other body movements). However, type-one errors were not observed in our experiment most likely due to three reasons. Firstly, the winking task is a very distinct eye movement; the task is not duplicated with normal eye movements such as blinking or gazing. Secondly, the EOG thresholds were carefully chosen by considering the averaged EOG amplitude and the standard deviation for discriminating the winking and natural eye blinking task (see figure 3). Lastly, all participants had a practice session to get used to the EOG detector and for fine-tuning the EOG threshold values.

In the proposed RC-VF system, the EOG detector was active all the time using a sliding window (see Figure 3). Such an asynchronous system has to secure a high level of reliability for idle-state detection rate (i.e., preventing type-one errors). One way is to allow only for the generation of the EOG command at a certain time interval. For instance, the EOG detector is active once after each end sequence. This system concept would be a bit more time-consuming compared to the fully asynchronous system; however, it could successfully block unintentional EOG commands during the idle-state.

The second type of EOG error would be that the EOG signal amplitude does not exceed the threshold. This type of error mostly occurs in the practice session (described in II-D); the setting of a subject-specific threshold and the generation of a uniform EOG signal are important. These errors rarely occurred in the RC-VF spelling session, but in cases where they do, the subject has to generate the wink with greater care and concentration to exceed the threshold.

While the binary EOG tasks (i.e., winking and eye-up) worked robustly throughout our experiment, a failure of the EOG system could possibly occur from unintended/unexpected movements of the eyes or body regardless of the system complexity. In this study, we used a single EOG channel for the user's convenience; most EOG-based HCI systems have used more than two channels for estimating the horizontal and vertical information from the eyes [12], [46]. Therefore, one solution to prevent the interference by some other movement's noise lies in using multi-channel EOG to acquire more specific information about the eye's movements. For instance, if the system uses two EOG channels for detecting the winking task, the system could more robustly block the unintended movement as the winking task has highly discriminant patterns between the horizontal EOG channels.

While in this study, we have used a simple threshold-based EOG detection method, similar to earlier approaches [7], [8], [12], [46], a wide variety of other eye movement algorithms and feature extraction methods are available. Lv *et al.* [47] proposed an eye movement pattern matching and spectral entropy algorithm to detect the endpoints of EOG pulses, which can improve the performance in a noisy background. Bulling *et al.* [48] devised 90 different features and validated them to find optimal subsets using minimum redundancy maximum relevance (mRMR) feature selection, which can recognize more complex eye-movement patterns. He and Li [13] proposed an EOG-based spelling system by combining the support vector machine (SVM) classification and waveform detection. These EOG detection algorithms have been used for more complex classification problems; therefore, these types of algorithms could be successfully incorporated into our RC-VF system for a more robust detection of the EOG task.

For our spelling framework, we have chosen the EOG detector as the confirmation for character selection. However, it should be noted that, in principle, movement features other than eye movements could be used to operate the system. If a given user (or patient) prefers a different action other than eye movements, our system can adapt to detect a different body movement to fulfill the role of the EOG.

In principle, movement detection of any body part, such as tongue [49], [50], hand [51], feet [52], etc. [21], is possible. The detection of such movements is easily accomplished due to the high SNR and will therefore maintain the advantage and reliability of our system. In that regard, patients with movement disabling conditions, such as ALS, paralysis, etc., can benefit greatly from our system, as the target muscle can be chosen with considerable flexibility.

Accurate performance in EOG detection is required in the proposed RC-VF system. The simple threshold-based EOG detector was used in this study, as we judged that the difficulty level of EOG classification is low (i.e., binary class) and this method has already shown remarkable performance in more complex classification problems [7], [8], [46]. Nevertheless, we suggest some approaches to enhance the performance of the EOG detector 1) by using multi-channel EOG [12], [46], or 2) by using more complex classification algorithms [47], [48], or even 3) by using the other available body parts [21], [49], [51], [52]. These methods have trade-off relationships between decoding accuracy, system complexity, and user convenience. Therefore, all these settings would be carefully designed according to the user/patient conditions.

B. ERP Performance and System Flexibility

In our results, the conventional RC-based ERP speller has a performance that is very similar to previously reported findings [28]–[31], [36]–[38]. Incidentally, there have been quite a number of approaches to enhance performance in ERP spellers [27], [29], [33], [34], [53], [54]. In regard to their methodological philosophy, these approaches follow one of three routes: 1) boosting the ERP deflections by altering target stimulus parameters, or 2) combining ERP components with other brain signal components, or 3) applying advanced feature selection methods. A prime example for boosting the strength of the ERP response is a study by Kaufmann *et al.* [33] where images of faces were used as stimuli to evoke stronger ERP responses. Li *et al.* [34] and Jin *et al.* [27] have also found significant performance increases by altering the interface design, screen size, and colors of the stimuli. The second method of combining different components is termed hybrid-BCI. Previous studies have shown performance increases from combinations of ERP and other components, such as SSVEP [55]–[57], motor-imagery (MI) [58], [59], error-related potentials (ERRP) [53], and others [54]. The advanced feature extraction techniques, such as kernel PCA (Principal Component Analysis) [60] and non-linear machine learning techniques, such as logistic regression or support vector machines [61]–[63] would be considered for further performance improvement. The advantage of our proposed framework is that the novel concept of our feedback-based switch can be applied to these approaches as an additional factor for performance improvement. Therefore, our framework has the potential to yield additional performance increases to previous speller systems.

C. RC-VF System

Table IV explains the operating principles of our framework by examining the spelling results of one particular subject (subject ‘F’). The left column depicts the target letters, and P1, P2, P3 represent the top 3 rankings of the classifier outputs from all 36 letters. In this study, we found high variability in the number of sequences that were needed by the system to properly identify a target character as the most likely candidate. For instance, the letter ‘K’ was identified as the most likely candidate, even when only one sequence was averaged. However, for the letter ‘A’ (row 5), four sequences needed to be averaged, so that it was ranked as the most probable letter. Therefore, predefining a fixed number of sequences as the threshold [29], [64], [65] will inevitably lead to suboptimal results.

Our proposed system creates a direct path to selecting a target letter through active user feedback. This is a surefire way to select the proper character in circumstances where the system gives a rather unreliable choice. To see a clear difference between our proposed framework and the conventional threshold-based approach, we have employed a pseudo-online technique [59], [66] to make grounds for equal comparison. Offline data acquired from the RC experiment (subject ‘F’) was fed to the pseudo-online emulator. In our proposed approach, the 36 characters were ranked after every sequence by the classifier output. If the target character was assigned as rank one, we assume that it would be selected by the user through the use of the EOG.

We compared our approach to a fixed threshold of 4 and 7 sequences. For these two thresholds, the average accuracy for this particular subject resulted in 62.5% and 81.2%,

TABLE IV

SELECTION PROCEDURE OF INDIVIDUAL LETTERS FOR RC (THRESHOLD-BASED) AND RC-VF (EOG-BASED) FOR SUBJECT F WITH DETAILED RANKING OF INDIVIDUAL LETTERS. THE TABLE SHOWS THE RANKING OF LETTERS WITH RESPECT TO THE NUMBER OF SEQUENCES AVERAGED. P1, P2 AND P3 ARE THE THREE HIGHEST RANKING SCORES. ACTUAL CLASSIFICATION OUTPUTS ARE GIVEN NEXT TO THE LETTER. THE LAST 3 COLUMNS INDICATE IF THE TARGET WAS SELECTED, AND IN THE CASE OF EOG, THE SEQUENCE NUMBER IS INDICATED. IN THE RC PARADIGMS, A CHECK DENOTES IF THE TARGET WAS DETECTED AFTER 4 OR 7 SEQUENCES.

Target Letter		Num. of Sequences							Correct (Seq.)		
		1	2	3	$4^{(Thrs1)}$	5	6	$7^{(Thrs2)}$	EOG	(Thrs1)	(Thrs2)
	P1	**K (4.0)**	K (4.4)	K (2.5)	K (2.9)	K (2.4)	K (2.7)	K (2.3)			
'K'	*P2*	L (2.9)	L (3.5)	L (2.2)	G (2.4)	I (1.9)	G (1.7)	I (1.3)	√ (1)	√	√
	P3	I (2.7)	G (2.6)	I (1.8)	I (2.2)	G (1.8)	I (1.4)	G (1.2)			
		4 (1.5)	**O (2.6)**	O (2.7)	O (2.3)	O (1.5)	O (1.3)	O (1.3)			
'O'		1 (1.4)	R (2.2)	P (2.2)	P (1.9)	P (0.9)	N (0.5)	R (0.4)	√ (2)	√	√
		O (1.0)	Q (1.7)	R (1.8)	R (1.2)	N (0.6)	P (0.5)	N (0.4)			
		Q (2.2)	R (3.3)	**R (2.7)**	R (2.9)	R (2.3)	R (2.0)	R (1.8)			
'R'		R (2.2)	_ (2.3)	Q (1.3)	Q (1.4)	Q (1.2)	Q (1.0)	Q (0.8)	√ (3)	√	√
		E (1.4)	L (2.1)	M (1.2)	_ (1.1)	L (0.8)	_ (0.5)	_ (0.5)			
		Q (0.4)	**E (2.1)**	E (1.7)	E (1.4)	E (0.8)	B (0.7)	E (0.6)			
'E'		E (0.3)	C (1.2)	C (1.4)	C (1.1)	C (0.8)	E (0.6)	B (0.6)	√ (2)	√	√
		9 (0.1)	B (1.0)	9 (0.8)	B (1.0)	B (0.7)	C (0.5)	C (0.3)			
		B (3.2)	D (1.6)	B (0.9)	**A (1.1)**	A (0.9)	A (1.2)	A (1.3)			
'A'		D (2.4)	B (1.2)	E (0.5)	B (1.0)	B (0.7)	B (0.7)	B (0.6)	√ (4)	√	√
		N (2.3)	E (1.0)	D (0.2)	D (0.6)	D (0.4)	C (0.4)	D (0.5)			
		7 (3.7)	7 (2.5)	**_ (1.9)**	_ (1.7)	_ (1.3)	_ (1.4)	_ (1.1)			
'_'		_ (2.1)	_ (2.2)	7 (1.5)	5 (0.9)	5 (0.8)	5 (0.9)	5 (0.5)	√ (3)	√	√
		9 (2.0)	5 (1.5)	5 (1.2)	7 (0.8)	6 (0.5)	6 (0.5)	6 (0.5)			
		W (3.4)	X (2.3)	X (2.6)	X (2.8)	X (2.7)	X (2.0)	X (2.2)			
'U'		T (3.3)	U (2.2)	U (2.2)	U (2.3)	U (2.5)	U (1.9)	U (1.8)	√ (9)	-	-
		U (2.6)	W(1.7)	T (1.7)	S (2.1)	W (2.3)	S (1.7)	S (1.6)			
		R (2.1)	**N (1.8)**	Q (2.7)	Q (2.5)	Q (2.0)	Q (1.8)	Q (1.8)			
'N'		P (1.8)	O (1.7)	R (1.4)	P (1.5)	N (0.9)	N (1.0)	P (1.3)	√ (2)	-	-
		Q (1.4)	P (1.5)	O (1.4)	O (1.3)	O (0.9)	O (1.0)	N (1.1)			
		C (1.9)	O (1.6)	O (1.3)	O (1.1)	O (0.7)	**I (0.3)**	O (0.2)			
'I'		I (1.7)	Q (1.3)	Q (1.2)	Q (0.9)	I (0.7)	O (0.1)	I (0.1)	√ (6)	-	-
		E (1.7)	I (0.8)	I (0.4)	I (0.3)	Q (0.3)	K (0.1)	N (-0.1)			
		V (3.1)	V (3.0)	V (3.8)	V (3.2)	V (2.4)	V (1.9)	V (1.7)			
'V'		D (1.9)	X (1.6)	X (2.6)	T (1.7)	D (1.2)	8 (1.1)	8 (0.9)	√ (1)	√	√
		8 (1.5)	D (1.5)	S (2.0)	X (1.6)	P (0.9)	P (0.8)	P (0.9)			
		D (2.6)	D (1.5)	**E (1.4)**	E (1.8)	E (2.3)	E (3.1)	E (2.9)			
'E'		B (2.4)	E (1.1)	D (1.2)	D (1.2)	D (1.3)	D (1.4)	A (1)	√ (3)	√	√
		F (1.7)	A (0.7)	A (0.5)	A (0.6)	9 (1.1)	A (1.4)	9 (1)			
		I (7.0)	O (6.1)	O (4.5)	O (3.3)	**R (2.2)**	R (2.0)	R (1.6)			
'R'		O (6.7)	Q (4.6)	Q (3.8)	Q (2.7)	O (2.1)	O (1.8)	O (1.6)	√ (5)	-	√
		7 (6.4)	M (3.9)	M (3.4)	M (2.5)	Q (1.6)	Q (1.6)	M (1.2)			
		S (4.5)	X (2.5)	X (2.1)	S (1.6)	S (1.5)	S (1.8)	S (1.6)			
'S'		A (3.4)	S (2.3)	S (1.9)	X (1.4)	X (1.5)	X (1.5)	X (1.3)	√ (1)	√	√
		X (3.2)	W (1.5)	F (1.5)	A (1.1)	U (0.9)	U (1.0)	U (0.8)			
		I (2.6)	7 (3.1)	1 (2.1)	1 (1.1)	C (1.0)	I (1.1)	I (0.8)			
'I'		7 (2.4)	1 (2.9)	7 (2.0)	7 (1.0)	1 (0.9)	C (0.9)	C (0.5)	√ (1)	-	√
		1 (1.8)	C (1.5)	C (1.4)	I (1.0)	I (0.9)	1 (0.7)	7 (0.3)			
		1 (1.4)	Z (0.5)	Z (1.7)	Z (1.5)	Z (0.8)	Z (1.1)	**T (0.6)**			
'T'		Y (1.3)	B (0.4)	T (1.0)	T (0.6)	T (0.8)	B (1.0)	Z (0.6)	√ (7)	-	√
		3 (0.5)	T (0.2)	B (1.0)	B (0.5)	B (0.6)	T (0.8)	B (0.6)			
		Y (3.9)	Y (2.8)	Y (1.7)	Y (2.1)	Y (2.4)	Y (2.3)	Y (2.1)			
'Y'		4 (3.1)	M (2.1)	M (1.7)	M (1.3)	3 (1.4)	3 (1.0)	3 (1.1)	√ (1)	√	√
		3 (2.4)	3 (1.8)	G (0.5)	A (0.8)	Z (0.8)	M (0.6)	M (0.6)			
Acc. [%]									**100**	**62.5**	**81.2**
ITR [bits/min]									**33.1**	**11.5**	**10**
Total Time [s]									**150 (±7.3)**	**192**	**336**

respectively. Allowing a user-based confirmation through EOG translates into a dynamic threshold, which leads to a significant performance increase, as seen in Table IV. For instance, the target letters ‘E’, ‘A’ and ‘_ ’ are robustly detected after a few sequences (stable case). However, for some cases (letters ‘U’, ‘N’, ‘I’), two (or more) letters are competing for the top rank position and may change position depending on the number of sequences averaged (unstable case).

In both these cases, our system provides an optimal solution with the help of user-based feedback. In the stable case, the user can select the correct letter at an early stage, which leads to increased ITR. Additionally, in the unstable case, the user can also speed things up by stopping the selection process when the target letter is ranked in first place after the first sequence while improving overall accuracy.

The online spelling performance is highly dependent on the classifier's parameters derived from the calibration data. In Figure 5, the black line indicates a strong positive relationship between the single-trial classification performance of offline data and online spelling performance (accuracy and ITRs). However, the proposed system shows that even with relatively low single-trial classification accuracy for the offline data, subjects are still able to reach high spelling accuracy during the online spelling phase.

Low ITR, as well as performance variability between the subjects, are the key obstacles currently preventing BCI spellers from real-world applications. Our study shows that the integration of EEG signals with EOG leads to significant increases in accuracy. Furthermore, it also reduces performance variability between subjects, particularly for subjects who would normally exhibit low performance (please refer to Figure 5).

D. Visual Feedback of the ERP Speller

Conventional ERP-based spellers typically only give visual feedback to users once a letter has been selected by the system. The visual feedback given to the user is generally a final decision for that specific letter, which cannot be changed easily (except if the user utilizes a delete option as the next letter). In our approach, the user is given feedback at an early stage, and the user is therefore able to participate in the decision by confirming the proposed selection. This novel type of interaction on the subject's side not only increases the ITR and robustness of the system, but it also helps the user maintain concentration while spelling.

After completing the experiment, all subjects completed a simple survey to evaluate their different impressions of the RC and RC-VF spelling. The questionnaire includes the following three questions:

1. preference: system preference level between RC and RC-VF speller (1 point: very low, 5 points: very high).
2. concentration: level of concentration during spelling (1 point: very low, 5 points: very high).
3. difficulty of generating a wink: level of difficulty/discomfort during RC-VF system (1 point: very easy, 5 points: very hard).

Figure 7 shows the averaged rating scores corresponding to preference, concentration, and difficulty. Eighty percent of the participants preferred the RC-VF mode (four participants gave the same rating scores to both spellers), with 90% finding visual feedback helpful in concentrating on the spelling. The participants reported that they felt interested in watching their brain response as visual feedback, and it helped with concentration on the spelling task. The averaged rating score

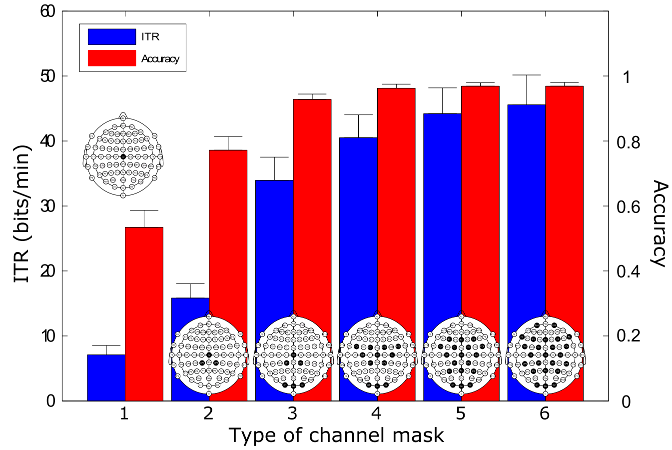

Fig. 6. Performance changes with respect to 6 types of channel configurations (Mask 1: Cz (1 ch), Mask 2: Cz-CP1-CP2 (3 ch.), Mask 3: Cz-CP1-CP2-O1-Oz-O2 (6 ch.), Mask 4: FC5-FC1-FC2-FC6-C2-Cz-C4-CP1-CP2-O1-Oz-O2 (12 ch.), Mask 5: F3-Fz-F4-FC5-FC1-FC2-FC6-C2-Cz-C4-CP1-CP2-P3-Pz-P4-O1-Oz-O2 (18 ch.), Mask 6: all 24 ch.). The left y-axis depicts ITR (corresponding to blue bar) and right y-axis depicts accuracy, corresponding to red bar.

for difficulty was 1.9 in which four subjects felt discomfort (rating scores >2) for generating an EOG signal.

Please note that all participants in this study were in healthy condition and the age spectrum is small (24-32). The rating scores could be different for patients or larger age ranges.

E. Usability in Real-World Applications

To increase the usability of BCI systems so that they become applicable in the real world, it is necessary to reduce the required setup time. Reducing the number of channels can minimize the setup time of an EEG-based BCI. Our experiment was recorded and analyzed with 24 EEG channels; however, we were interested to find out the minimum number of channels that were required for stable system performance.

To this end, we designed 6 types of channel configurations (1 ch. in Mask 1: Cz, 3 ch. in Mask 2: Cz-CP1-CP2, 6 ch. in Mask 3: Cz-CP1-CP2-O1-Oz-O2, 12 ch. in Mask 4: FC5-FC1-FC2-FC6-C2-Cz-C4-CP1-CP2-O1-Oz-O2, 18 ch. in Mask 5: F3-Fz-F4-FC5-FC1-FC2-FC6-C2-Cz-C4-CP1-CP2-P3-Pz-P4-O1-Oz-O2, all 24 ch. in Mask 6). The classification accuracy validated along with the channel configurations in offline condition. Figure 6 shows the channel configurations as well as performance changes. For validation, offline data from RC sessions were fed into a pseudo-online emulator. As expected, the full channel setup showed the highest performance and the lowest performance was obtained if only one channel was used. However, our results indicate that two channel configurations with 6 and 12 channels (Mask 3 and Mask 4, respectively) also yielded reasonable performance when compared to a full channel setup (Mask 6).

F. Enhancing the Performance of RC-VF Speller

We already discussed how our proposed system could easily be combined with previous speller approaches for additional performance improvements. In order to demonstrate our system flexibility and its possible use as a real-world

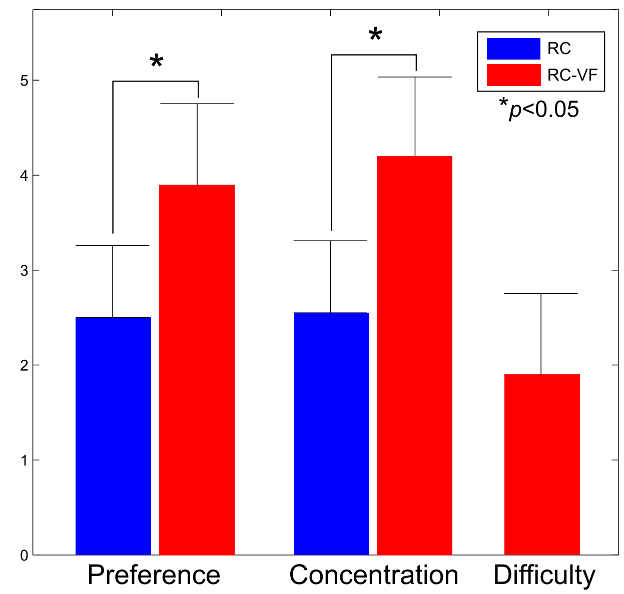

Fig. 7. The averaged rating scores from a questionnaire for all participants (paired t-test, p). Preference (1 point: very low, 5 points: very high), concentration (1 point: very low, 5 points: very high), and difficulty (1 point: very easy, 5 points: very hard).

application, an extended experiment was designed to maximize both spelling performance and usability. Six factors were manipulated from the basic RC-VF system: 1) The stimulus-time interval (STI) was set to 150 ms, followed by an inter-stimulus interval (ISI) of 50 ms. 2) In order to enhance the ERP signal, random set-based representation pattern (RASP) [29] with a face stimulus [33] was used. 3) In order to reduce set-up time, only five channels were used in the experiment (Cz, CP3, CP4, Oz, and FP1 for EOG detection). 4) Following participant feedback, instead of presenting a ranking score (i.g 1/2/3/4) above the character, only the highest ranked character (i.e. ‘1’) was highlighted in a different color. 5) Following participant feedback for greater comfort, the winking task for the selection of the target character was replaced by eye movement (eye-up) [11]. 6) The speller system was set to allow an unlimited number of sequences and the weight parameter w_p, $(p = 1, \ldots, 32)$ was added to the classification procedure.

The decision to change factor four from a presentation of the top four ranking candidates to just the top ranked candidate was made due to user feedback and a lack of evidence that this extra information improved performance. We had believed that the dynamic changes of the visual feedback for scores could increase (or adjust) the user's concentration level so that the user could focus more on the target stimuli in effort to acquire a higher score, but there was little evidence to support this. Therefore, we simply present the single most probable candidate character with the highest classification score as a different color.

Similarly to factor four, the eye winking task (factor five) was reported to be uncomfortable by some participants. The eye-up task is also a very distinct eye movement task. The user was instructed to generate maximum eye movement from the current position to the upper position when they wanted to

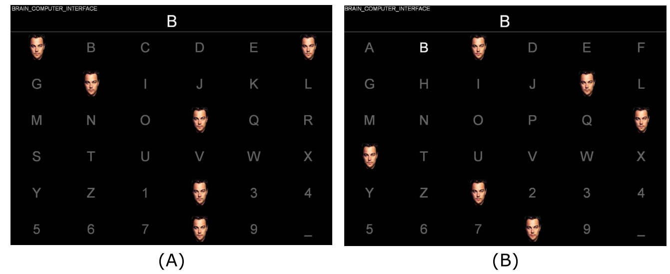

Fig. 8. Extended experimental setup. (A) Random set stimulus presentation with famous face image (L. W. DiCaprio, Actor). (B) The target character ‘B’ which had the highest classification score is highlighted as a different color.

choose the target character. Also, the FP1 channel is used for EOG detection to reduce the discomfort of attaching the electrode near their eye. Figure 3-right plot shows the EOG data for the eye-up task. In this figure, we calibrated data for the two conditions of eye-up and natural eye-blink tasks. The eye-up task showed a strong characteristic response compared to the baseline interval (−500 to 0 ms) as well as to natural eye-blinking. Thus, the eye-up task would be suitable to use as the EOG task. Both the wink and eye-up movements usually do not occur in normal routines and circumstances, so their specificity to signify an intentional event is very high. In summary, the system components such as the visual feedback, EOG task, and the channel set highly depend on user preference and the external environment (e.g. system architecture, EEG device, etc.). The additional experiment showed the extendibility of our proposed system; the system components could be easily changed/modified to individual needs.

The final modification is the most noteworthy, and it is important to understand its implementation. First, the w_p is initialized as 1.0 at the beginning of a sequence. This weight parameter is updated when the user doesn't select the currently highlighted character as we assume that the character is not the target character. For instance, if the character ‘K’ is visually highlighted (=rank_1) by the classifier output (Y) and the user doesn't select ‘K’ by generating an EOG signal, then we give lesser weight to the w_k. For the next sequence, the classifier output is $w \cdot f(v)$ which means the character ‘K’ has a relatively smaller probability to be chosen as a target feedback. That makes the system more efficient by reducing feedback redundancy.

The proposed classification strategy has two important implications for enhancing spelling performance. First, all candidate letters have an opportunity to be chosen as a target character. If a user is able to generate an EOG command properly, the use of the weighted classifier will have 100 % accuracy with the unlimited number of sequences, but necessarily the spelling speed will be slower. Second, it makes the system more efficient by reducing feedback redundancy. We have already shown that the primary reason for low spelling performance comes from close competition between the candidate characters which have a high probability classifier score (see Table IV). For instance, in Table IV, the selection of the characters of ~~‘_’~~, ‘U’, ‘I’(first), ~~‘E’~~, ‘T’

TABLE V
SPELLING ACCURACY AND ITR OF RC-VF SYSTEM WITH ENHANCED METHODS. SIX SUBJECTS (A, F, G, M, P, AND, S) ARE CHOSEN FROM THE PREVIOUS EXPERIMENT (SEE TABLE II)

Sub.	Acc.[%]	ITR[bits/min]	Time[s]/letter
A	100	83.5	3.71
F	100	82.6	3.75
G	100	69.3	4.47
M	100	37.2	8.21
P	100	44.6	6.95
S	100	29.8	10.3
Mean	100	57.8	6.27

was delayed by its unreliable classifier output of the target letter. If the high ranked non-target characters are eliminated in the early stages, the spelling performance is fast and accurate. Rather than eliminate the unselected character from the candidate bin in cases where a user might miss choosing target character in time, we applied a weighted classifier strategy so that another selection opportunity would be possible. And lastly, conventional speller systems require two steps for error correction where the user must first negate the previous input character and then next input the correct target letter. Our system, on the other hand, eliminates the need for this error correction procedure.

Six subjects (A, F, G, M, P and S) participated in the extended experiment. The training and test procedures were the same as in the previous RC-VF experiment. The experimental results showed 100% spelling accuracy with an averaged ITR of 57.8 ($\pm$23.6) [bits/min] (see Table V). All participants showed significantly improved performance compared to the basic RC-VF experiment.

G. Limitation of the Proposed RC-VF Speller

While we believe that our proposed RC-VF speller system can be used by most patients in need of a BCI system, there are limits to its appropriate application. In most cases of spinal cord injury such as quadriplegia, motor function is preserved above the neck. Additionally, our system can work in cases of locked-in-syndrome (LIS), which is a condition where a patient is aware but cannot move due to complete paralysis of nearly all voluntary muscles in the body except for vertical eye movement and blinking [67]. However, in cases such as total locked-in syndrome or total immobility where the patient is not able to control any voluntary movement including all eye movements [67], our system is obviously not appropriate. Given that eye based spellers (EOG or eye-tracker) and matrix-type EEG spellers require voluntary eye movements as the user needs to gaze at the targeting position, these systems, like ours, are less useful for patients with severe oculo-motor impairment. In these cases, like in total locked-in syndrome, gaze-independent speller systems have been proposed, and we would point your attention to Treder *et al.* [68] who developed a two-stage visual speller based on covert spatial attention and non-spatial feature attention. Additionally, Acqualagna *et al.* [69] proposed a novel BCI paradigm consisting of a central rapid serial visual presentation (RSVP) of the stimuli that shows promising results. It is important to be clear that our proposed RC-VF system belongs to the gaze-dependent speller category and mainly focuses on the target user who can control at least one part of their body's movement, including their eyes.

An ERP-based BCI system is a time-consuming process that needs to collect enough EEG trials until the classification output is clear. To reduce the consumption of time, previous approaches have changed the experimental setting [27], [29], [33], [34], [53], [54] to enhance the ERP components, or they combined physiological signals such as EMG/EOG [19]–[22] or another brain component such as SSVEP/MI [70].

Recently, dynamic stopping methods have been proposed to enhance the speller's performance. The dynamic stopping approach updates the probabilities for all possible symbols based on the classifier response and chooses the one character when the probability exceeds some stopping criterion (i.e. a threshold value). Various stopping algorithms have been proposed like ones that use a simple cumulated classification score. Lenhardt *et al.* [71] proposed the P300 speller system which dynamically limits the number of subtrial presentations according to the user's current online performance. The study defined two scoring functions namely sum-threshold and ratio-threshold. The system selects a target symbol when the overall classifier outputs for all candidate characters is lower than the predefined threshold (sum-threshold) or the score ratio between the highest scored and second highest scored character is greater than the predefined threshold (ratio-threshold). The results showed an accuracy of 87.5% and transfer rate of a 29.3 bits/min. Furthermore, a Bayes classifier with a threshold probability [72], or others that use language models [31], [38] have also been proposed for increasing the ITR. Such language models also showed significantly enhance performance (90.36% accuracy with 54.42 bits/min [38] and 92.3% accuracy with 39.6 bits/min [31] accuracy).

The reason for low ITRs in spelling systems (as well as in most other BCI systems) is primarily due to the competition between high-ranked candidate characters (i.e. the one true target vs. non-targets). Thereby, the adaptive and the conventional method have the common characteristic that the classifier should wait until the classification score for the highest scored character is clear. This delay is basically unavoidable in the conventional as well as the adaptive system because the decision making is highly reliant on the classifier's output.

Our study also exhibits instances of this interfering competition that leads to delay in classification (see Table IV, character '_') in the spelling procedure. This is particularly problematic for lower performance users of a BCI. To enhance performance, previous approaches only focused on eliciting a stronger brain component or improving the system based classification strategy to discriminate between the target and non-target data. The dynamic stopping method is an example of this latter technique, and classification still remains dependent upon achieving a certain threshold.

Our hybrid EEG-EOG spelling system, however, has an entirely different approach for confirming the target compared to these other methods. The main idea of the proposed system is to share the decision process between the system and the user so that the user is more actively involved in the selection procedure. Therefore, the RC-VF system can circumvent this unnecessary competition between the high-ranked candidate characters; the system doesn't need to wait until the classifier's output is clear. Furthermore, the highest ranked non-target character would be eliminated early when the user fails to select it via an EOG command. This more interactive system between the user and computer can effectively reduce the number of trials by avoiding prolonged competition between candidate target characters, so in cases of users who still retain some mobility, application of our system's technique would certainly improve performance. In cases of total immobility, the previous methods that improve signal processing and/or enhanced the ERP component are still the best options to date.

VI. Conclusion

In this study, the outstanding online performance with an accuracy of 97.6% and an ITR of 39.6 [bits/min] indicates that the proposed RC-VF paradigm is a very promising system as a real-world application. Our extended experiment further showed real world applicability by reducing the number of channels, changing the input EOG signal for easier use (eye-up versus blinking), and modifying the stimulus presentation for better ERP signal generation and comfort for the user. The extended experiment also introduced the weighted classifier paradigm that eliminates the need for two-step error correction and enables 100% spelling accuracy.[1] Not only did the extended experiment achieve 100% accuracy, but it also improved the ITR to 57.8[bits/min]. In both of our experimental setups, the EEG-EOG hybrid design of the RC-VF speller system improves the spelling speed and accuracy of a conventional BCI system, while only requiring one EOG task. And most importantly we have demonstrated that the concept of this system can be applied to most existing gaze-dependent systems with the benefit of increased accuracy and ITR. In our future work, we will focus on testing (and possibly adapting) this system with the help of a clinical subject population. We hope that it will be possible soon to bring this technology from the test bench to the clinical subjects for whom the design was intended.

References

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

[2] P. Majaranta and K.-J. Räihä, "Twenty years of eye typing: Systems and design issues," in *Proc. Symp. Eye Tracking Res. Appl.*, 2002, pp. 15–22.

[3] A. T. Duchowski, *Eye Tracking Methodology Theory and Practice*, vol. 373. Springer, 2007.

[4] D. W. Hansen, J. P. Hansen, M. Nielsen, A. S. Johansen, and M. B. Stegmann, "Eye typing using Markov and active appearance models," in *Proc. 6th IEEE Workshop Appl. Comput. Vis. (WACV)*, Dec. 2002, pp. 132–136.

[5] A. T. Duchowski, "A breadth-first survey of eye-tracking applications," *Behav. Res. Methods, Instrum., Comput.*, vol. 34, no. 4, pp. 455–470, 2002.

[1] A demo video of the proposed system (participant F) is available on YouTube (https://youtu.be/pznGIeuIzYA).

[6] L. A. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalograph. Clin. Neurophysiol.*, vol. 70, no. 6, pp. 510–523, Dec. 1988.

[7] R. Barea, L. Boquete, M. Mazo, and E. López, "System for assisted mobility using eye movements based on electrooculography," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 10, no. 4, pp. 209–218, Dec. 2002.

[8] L. Y. Deng, C.-L. Hsu, T.-C. Lin, J.-S. Tuan, and S.-M. Chang, "EOG-based human–computer interface system development," *Expert Syst. Appl.*, vol. 37, no. 4, pp. 3337–3343, 2010.

[9] T. W. Picton, "The P300 wave of the human event-related potential," *J. Clin. Neurophysiol.*, vol. 9, no. 4, pp. 456–479, 1992.

[10] G. Pfurtscheller and F. L. Da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, 1999.

[11] A. B. Usakli and S. Gurkan, "Design of a novel efficient human–computer interface: An electrooculagram based virtual keyboard," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 8, pp. 2099–2108, Aug. 2010.

[12] Q. Huang *et al.*, "An EOG-based human-machine interface for wheelchair control," *IEEE Trans. Biomed. Eng.*, 2017.

[13] S. He and Y. Li, "A single-channel EOG-based speller," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 11, pp. 1978–1987, Nov. 2017.

[14] M. Arns, S. de Ridder, U. Strehl, M. Breteler, and A. Coenen, "Efficacy of neurofeedback treatment in ADHD: The effects on inattention, impulsivity and hyperactivity: A meta-analysis," *Clin. EEG Neurosci.*, vol. 40, no. 3, pp. 180–189, 2009.

[15] K. Shindo *et al.*, "Effects of neurofeedback training with an electroencephalogram-based brain–computer interface for hand paralysis in patients with chronic stroke: A preliminary case series study," *J. Rehabil. Med.*, vol. 43, no. 10, pp. 951–957, 2011.

[16] M.-H. Lee, B.-J. Kim, and S.-W. Lee, "Quantifying movement intentions with multimodal neuroimaging for functional electrical stimulation-based rehabilitation," *NeuroReport*, vol. 27, no. 2, pp. 61–66, 2016.

[17] P. Majaranta, I. S. MacKenzie, A. Aula, and K.-J. Räihä, "Effects of feedback and dwell time on eye typing speed and accuracy," *Univ. Access Inf. Soc.*, vol. 5, no. 2, pp. 199–208, 2006.

[18] A. Usakli, S. Gurkan, F. Aloise, G. Vecchiato, and F. Babiloni, "A hybrid platform based on EOG and EEG signals to restore communication for patients afflicted with progressive motor neuron diseases," in *Proc. Annu. Int. Conf. Eng. Med. Biol. Soc. (EMBC)*, Sep. 2009, pp. 543–546.

[19] D. Puanhvuan, S. Khemmachotikun, P. Wechakarn, B. Wijarn, and Y. Wongsawat, "Navigation-synchronized multimodal control wheelchair from brain to alternative assistive technologies for persons with severe disabilities," *Cogn. Neurodyn.*, vol. 11, no. 2, pp. 117–134, 2017.

[20] H. Wang, Y. Li, J. Long, T. Yu, and Z. Gu, "An asynchronous wheelchair control by hybrid EEG–EOG brain–computer interface," *Cogn. Neurodyn.*, vol. 8, no. 5, pp. 399–409, 2014.

[21] J. Ma, Y. Zhang, A. Cichocki, and F. Matsuno, "A novel EOG/EEG hybrid human–machine interface adopting eye movements and ERPs: Application to robot control," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 3, pp. 876–889, Mar. 2015.

[22] S. Soekadar *et al.*, "Hybrid EEG/EOG-based brain/neural hand exoskeleton restores fully independent daily living activities after quadriplegia," *Sci. Robot.*, vol. 1, no. 1, p. eaag3296, 2016.

[23] C.-C. Postelnicu and D. Talaba, "P300-based brain-neuronal computer interaction for spelling applications," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 2, pp. 534–543, Feb. 2013.

[24] J. Jiang, Z. Zhou, E. Yin, Y. Yu, and D. Hu, "Hybrid brain-computer interface (BCI) based on the EEG and EOG signals," *Bio-Med. Mater. Eng.*, vol. 24, no. 6, pp. 2919–2925, 2014.

[25] B. Koo, Y. Nam, and S. Choi, "A hybrid EOG-P300 BCI with dual monitors," in *Proc. Int. Winter Workshop Brain-Comput. Interface (BCI)*, Feb. 2014, pp. 1–4.

[26] K. Lin, A. Cinetto, Y. Wang, X. Chen, S. Gao, and X. Gao, "An online hybrid BCI system based on SSVEP and EMG," *J. Neural Eng.*, vol. 13, no. 2, p. 026020, 2016.

[27] J. Jin *et al.*, "The changing face of P300 BCIs: A comparison of stimulus changes in a P300 BCI involving faces, emotion, and movement," *PLoS ONE*, vol. 7, no. 11, p. e49688, 2012.

[28] J. Jin, E. W. Sellers, and X. Wang, "Targeting an efficient target-to-target interval for P300 speller brain–computer interfaces," *Med. Biol. Eng. Comput.*, vol. 50, no. 3, pp. 289–296, 2012.

[29] S.-K. Yeom, S. Fazli, K.-R. Müller, and S.-W. Lee, "An efficient ERP-based brain-computer interface using random set presentation and face familiarity," *PLoS ONE*, vol. 9, no. 11, p. e111157, 2014.

[30] Q. Li, S. Liu, J. Li, and O. Bai, "Use of a green familiar faces paradigm improves P300-speller brain-computer interface performance," *PLoS ONE*, vol. 10, no. 6, p. e0130325, 2015.

[31] W. Speier, C. Arnold, J. Lu, A. Deshpande, and N. Pouratian, "Integrating language information with a hidden Markov model to improve communication rate in the P300 speller," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 3, pp. 678–684, May 2014.

[32] J.-H. Lee, J.-H. Lim, H.-J. Hwang, and C.-H. Im, "Development of a hybrid mental speller combining EEG-based brain-computer interface and webcam-based eye-tracking," in *Proc. 35th Annu. Int. Conf. Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 2240–2242.

[33] T. Kaufmann, S. M. Schulz, A. Köblitz, G. Renner, C. Wessig, and A. Kübler, "Face stimuli effectively prevent brain–computer interface inefficiency in patients with neurodegenerative disease," *Clin. Neurophysiol.*, vol. 124, no. 5, pp. 893–900, 2013.

[34] Y. Li, C. S. Nam, B. B. Shadden, and S. L. Johnson, "A P300-based brain–computer interface: Effects of interface type and screen size," *Int. J. Hum.-Comput. Interaction*, vol. 27, no. 1, pp. 52–68, 2010.

[35] B. Z. Allison and J. A. Pineda, "Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: Implications for a BCI system," *Int. J. Psychophysiol.*, vol. 59, no. 2, pp. 127–140, 2006.

[36] G. Townsend, J. Shanahan, D. B. Ryan, and E. W. Sellers, "A general P300 brain–computer interface presentation paradigm based on performance guided constraints," *Neurosci. Lett.*, vol. 531, no. 2, pp. 63–68, 2012.

[37] J. Shen, J. Liang, J. Shi, and Y. Wang, "A dynamic submatrix-based P300 online brain–computer interface," *Biomed. Signal Process. Control*, vol. 15, pp. 27–32, Jan. 2015.

[38] B. O. Mainsah, K. A. Colwell, L. M. Collins, and C. S. Throckmorton, "Utilizing a language model to improve online dynamic data collection in P300 spellers," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 837–846, Jul. 2014.

[39] M.-H. Lee, S. Fazli, K.-T. Kim, and S.-W. Lee, "Development of an open source platform for brain-machine interface: OpenBMI," in *Proc. 4th Int. Winter Conf. Brain-Comput. Interface (BCI)*, Feb. 2016, pp. 1–2.

[40] M.-H. Lee *et al.*, "OpenBMI: A real-time data analysis toolbox for brain-machine interfaces," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2016, pp. 1884–1887.

[41] S. Lemm, B. Blankertz, T. Dickhaus, and K.-R. Müller, "Introduction to machine learning for brain imaging," *NeuroImage*, vol. 56, no. 2, pp. 387–399, 2011.

[42] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, "Single-trial analysis and classification of ERP components—A tutorial," *NeuroImage*, vol. 56, no. 2, pp. 814–825, 2011.

[43] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Ann. Eugenics*, vol. 7, no. 2, pp. 179–188, 1936.

[44] D. J. McFarland, W. A. Sarnacki, and J. R. Wolpaw, "Brain–computer interface (BCI) operation: Optimizing information transfer rates," *Biol. Psychol.*, vol. 63, no. 3, pp. 237–251, 2003.

[45] A. Kübler and N. Birbaumer, "Brain–computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients?" *Clin. Neurophysiol.*, vol. 119, no. 11, pp. 2658–2666, 2008.

[46] R. Barea, L. Boquete, M. Mazo, and E. López, "Wheelchair guidance strategies using EOG," *J. Intell. Robot. Syst.*, vol. 34, no. 3, pp. 279–299, 2002.

[47] Z. Lv, X.-P. Wu, M. Li, and D. Zhang, "A novel eye movement detection algorithm for EOG driven human computer interface," *Pattern Recognit. Lett.*, vol. 31, no. 9, pp. 1041–1047, 2010.

[48] A. Bulling, J. A. Ward, H. Gellersen, and G. Troster, "Eye movement analysis for activity recognition using electrooculography," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 4, pp. 741–753, Apr. 2011.

[49] P. J. Alfonso, K. Honda, T. Baer, and K. S. Harris, "Multi-channel study of tongue EMG during vowel production," *J. Acoust. Soc. Amer.*, vol. 71, no. S1, p. S54, 1982.

[50] A. B. Barreto, S. D. Scargle, and M. Adjouadi, "A practical EMG-based human-computer interface for users with motor disabilities," *J. Rehabil. Res. Develop.*, vol. 37, no. 1, p. 53, 2000.

[51] J. Kim, S. Mastnik, and E. André, "EMG-based hand gesture recognition for realtime biosignal interfacing," in *Proc. 13th Int. Conf. Intell. User Interfaces*, 2008, pp. 30–39.

[52] A. Crossan, S. Brewster, and A. Ng, "Foot tapping for mobile interaction," in *Proc. 24th BCS Interaction Spec. Group Conf.*, 2010, pp. 418–422.

[53] B. Dal Seno, M. Matteucci, and L. Mainardi, "Online detection of P300 and error potentials in a BCI speller," *Comput. Intell. Neurosci.*, vol. 2010, Jan. 2010, Art. no. 11.

[54] T. D. Lalitharatne, K. Teramoto, Y. Hayashi, and K. Kiguchi, "Towards hybrid EEG-EMG-based control approaches to be used in bio-robotics applications: Current status, challenges and future directions," *Paladyn, J. Behav. Robot.*, vol. 4, no. 2, pp. 147–154, 2013.

[55] E. Yin, T. Zeyl, R. Saab, T. Chau, D. Hu, and Z. Zhou, "A hybrid brain–computer interface based on the fusion of P300 and SSVEP scores," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 4, pp. 693–701, Jul. 2015.

[56] 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, no. 2, p. 026001, 2013.

[57] J.-Y. Hwang, M.-H. Lee, and S.-W. Lee, "A brain-computer interface speller using peripheral stimulus-based SSVEP and P300," in *Proc. 5th Int. Winter Conf. Brain-Comput. Interface (BCI)*, Jan. 2017, pp. 77–78.

[58] 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.

[59] M.-H. Lee, S. Fazli, J. Mehnert, and S.-W. Lee, "Subject-dependent classification for robust idle state detection using multi-modal neuroimaging and data-fusion techniques in BCI," *Pattern Recognit.*, vol. 48, no. 8, pp. 2725–2737, 2015.

[60] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Comput.*, vol. 10, no. 5, pp. 1299–1319, 1998.

[61] K.-R. Müller, C. W. Anderson, and G. E. Birch, "Linear and nonlinear methods for brain-computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 165–169, Jun. 2003.

[62] Y. Anzai, *Pattern Recognition & Machine Learning*. Amsterdam, The Netherlands: Elsevier, 2012.

[63] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 2013.

[64] C. Wang, C. Guan, and H. Zhang, "P300 brain-computer interface design for communication and control applications," in *Proc. 27th Annu. Conf. IEEE Eng. Med. Biol.*, Shanghai, China, Jan. 2006, pp. 5400–5403.

[65] H. Serby, E. Yom-Tov, and G. F. Inbar, "An improved P300-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, no. 1, pp. 89–98, Mar. 2005.

[66] B. Blankertz, G. Curio, and K.-R. Müller, "Classifying single trial EEG: Towards brain computer interfacing," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 1, 2002, pp. 157–164.

[67] G. Bauer, F. Gerstenbrand, and E. Rumpl, "Varieties of the locked-in syndrome," *J. Neurol.*, vol. 221, no. 2, pp. 77–91, 1979.

[68] 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, no. 6, p. 066003, 2011.

[69] L. Acqualagna, M. S. Treder, M. Schreuder, and B. Blankertz, "A novel brain-computer interface based on the rapid serial visual presentation paradigm," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, Aug. 2010, pp. 2686–2689.

[70] G. Pfurtscheller *et al.*, "The hybrid BCI," *Frontiers Neurosci.*, vol. 4, p. 3, Apr. 2010.

[71] 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.

[72] C. S. Throckmorton, K. A. Colwell, D. B. Ryan, E. W. Sellers, and L. M. Collins, "Bayesian approach to dynamically controlling data collection in P300 spellers," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 3, pp. 508–517, May 2013.