Exploring Gaze-Motor Imagery Hybrid Brain-Computer Interface design

Darren O'Doherty, Yogesh Kumar Meena, Haider Raza, Hubert Cecotti, Girijesh Prasad
School of Computing and Intelligent Systems, Magee Campus, University of Ulster, Londonderry, Northern Ireland, UK
d.doherty1@ulster.ac.uk, meena-y@email.ulster.ac.uk, raza-h@email.ulster.ac.uk, h.cecotti@ulster.ac.uk, g.prasad@ulster.ac.uk

***Abstract*—Non-invasive Brain-Computer Interface (BCI) has appeared as a new hope for a large population of disabled people, who were waiting for a new communication means that would translate some brain responses into actions. After several decades of research in fields such as neuroscience and machine learning, the performance remains too low due to the low signal to noise ratio of the EEG signal, and the time that has to be dedicated to the recording of the brain responses. Hybrid BCIs consider the combination of several modalities, including brain responses, for new communication systems. The creation of a Hybrid BCI requires particular care as it possesses the constraints from several modalities. We propose to investigate the performance that could be achieved in a paradigm, where gaze control is used for the selection of an item on a computer screen and motor imagery is used to enable the selected item on the screen. Based on the results obtained from gaze detection with an eye tracker, and motor imagery detection with non-invasive EEG recording, we show that the performance of a parallel Hybrid BCI is only beneficial if the accuracy of each modality reaches a particular limit, and if the number of commands from each modality is carefully chosen.**

Keywords—Brain-Computer Interfaces (BCIs), Motor Imagery (MI), Electroencephalography (EEG), Information Transfer Rate (ITR), Eye Tracker, Hybrid BCI

I. INTRODUCTION

Despite the interest and research in non-invasive Brain-Computer Interface (BCI) based on the recording of electroencehalogram (EEG) signal, the performance remains too low to allow disabled people to fully exploit BCI systems [1]. Whereas BCI systems were initially designed as the single communication pathway for some severely disabled people such as locked-in patients, BCI can also be used by other groups of patients, offering a greater speed and accuracy [2]–[4]. Non-invasive Brain-Computer Interfaces (BCIs) typically take advantage of the detection of a single type of brain response and they can be successfully used by a large group of users [5]–[7]. It can be event-related potentials (ERPs) such as the P300 speller [8], the detection of event related synchronization/desynchornization (ERD/ERS) i.e. motor imagery detection (MI), and steady-state visual evoked potentials (SSVEP) [9]. To fully integrate and consider a BCI system for a clinical and rehabilitation purpose, it is important to determine the constraints that are given by the conditions of the patients. For completely locked-in patients, gaze control may not be available, and a particular BCI solution has to be provided. However, when a patient does have a control of his/her gaze, then more appropriate BCIs can be considered, i.e., brain responses that are gaze dependent. In addition to the type of BCI that can be used by a patient, a recent trend is the creation of Hybrid BCI that allows the combination of several modalities [10]–[12].

A Hybrid BCI can either use two or more BCI systems or at least one BCI system with one other common system. The Hybrid BCI can be designed in two different ways, sequentially or in parallel [13], [14]. A sequential Hybrid BCI can be the sequence of two BCIs, a BCI based on motor imagery to enable/disable a second BCI such as a P300 speller. A parallel BCI can correspond to the simultaneous detection of two brain responses such as SSVEP and ERPs. The design of a Hybrid BCI must respect the constraints that are related to the presentation of stimuli that should evoke the particular brain responses to detect.

The vividness of MI, which is highly dependent on the visual perception and imagery, is an important factor that influences the feature extraction from the currently used EEG based BCI technique [15]. Additionally, it is well understood that imagery is a process of visual synthesis, and hence it is considered as a coordinate link between visual memory and eye movement patterns. Past research has confirmed significant correlations between the eyes scan path, perceptual processes, and higher cognitive functions [16], [17]. Further, it is suggested that in normal circumstances, motor movements, as the hand or the eyes and the associated attentional shifts, are linked together and the motor cortex of the brain is engaged during this mental transformation. Apart from the above facts, the idea for the development of an eye tracker based BCI system has also been supported by the fMRI based BOLD examination, in which eye-hand coordination and its alliance with cerebellum is demonstrated [18].

In this paper, we propose to evaluate the performance of a parallel Hybrid BCI based on the combination of motor imagery (MI) responses and gaze control. The system is based on the following principle: the eye tracker allows determining the position of the gaze on a computer screen, whereas the detection of a motor imagery response in the EEG allows to determine the selection of the target of interest (see Figure 1). By combining eyetracking and BCI, we want to determine if this solution can be a means to increase the number of commands with a predefined number of items on the screen, and if the motor imagery could be reliably used in this type of scenario. The goal of the proposed Hybrid BCI paradigm is to provide a directional search and select command for people who are able to control their gaze and who can use this Hybrid BCI system at home for rehabilitation and controlling other augmentative devices.

In this study, we intend to quantify the gain of performance that can be obtained by combining gaze and motor imagery

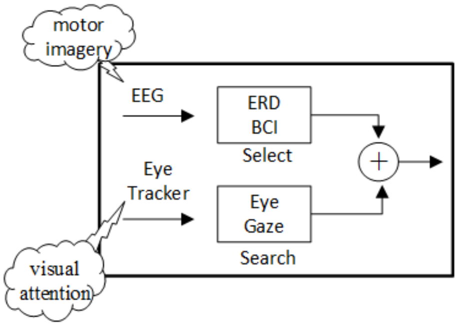

Fig. 1. A parallel Hybrid BCI that combines an ERD BCI and an eye tracker.

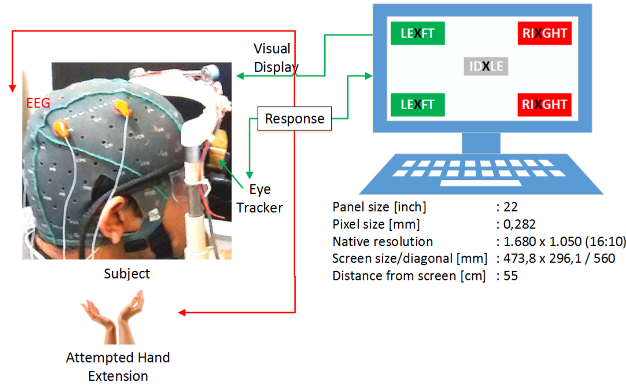

Fig. 2. Proposed Hybrid BCI paradigm.

Fig. 3. Gaze estimation results for a single calibration test run.

detection, and the constraints that are related to the accuracy of the detection from both modalities, gaze and EEG. To analyze the possible gain that could be obtained, we consider the results from two different experiments with similar experimental designs. The first experiment corresponds to gaze detection. The second experiment corresponds to motor imagery detection, and uses EEG data from a BCI competition.

The paper is organized as follows. First, Section II presents the experimental protocol. Section III is dedicated to the methods, the signal preprocessing and classification techniques. The results are presented in Section IV. Finally, the results are discussed in Section V.

II. Experimental Protocol

A. Subjects

A total of five healthy male subjects participated in the study, with an age range 26-35 years, mean age of 30.6 and standard deviation 3.5. All subjects had no prior experience with an eye tracker. None of the subjects had any visual or neurological conditions. Prior to experiments, each subject were advised of the nature and purpose of the study. No financial reward was received by the subjects for their participation in the study.

B. Design and Procedure

Figure 2 illustrates the proposed Hybrid BCI paradigm. The experiment evaluates a search and select task in a two dimensional environment. The eye tracker is used as a searching device which has a direct mapping to the mouse cursor, providing a real-time feedback response. The BCI was used to provide an additional select command to the user, using left or right hand MI. A visual cue paradigm was used and consisted of four visual stimuli, each containing a fixation cross. Two visual stimuli were located on the left side of the screen and two located on the right side of the screen with an additional visual stimuli located in the centre of the screen, representing idle state. Each visual stimuli contained one of two commands, left (green) or right (red). Subjects had to use the eye tracker controlled search command to direct the cursor to the target visual stimuli, whilst attempting the prompted select command, left or right hand MI. A rigid and stable eye tracker head lock was used for positioning the subjects head in the center of the screen. Each subject was seated in a comfortable chair located in front of the eye tracker head lock and approximately 55 cm from a 22 inch LCD monitor. Lighting conditions remained steady during experiments.

Prior to an experiment, the eye tracker was calibrated using a sixteen-point calibration. This provided each subject with an accurate position of gaze. Figure 3 shows the gaze estimation results for a single calibration test run. The blue squares represent 16 target points and the red squares represent the gaze estimations. Once a suitable calibration is achieved, the calibration data was stored and reused each time a subject returned. Each experiment consisted of one session, comprising of 120 trials with a trial duration of 7 s. In each trial the target of interest appears in a random order but they are evenly distributed over a session, appearing 30 times each. During each trial subjects were prompted to search for a single target of interest on the visual display using the eye tracker controlled mouse cursor, whilst attempting a signaled left or right hand MI command. Figure 4 shows the timing activity for a single trial. A trigger signal is sent at 0 s to indicate the start of each trial. A beep command is issued at 1 s to indicate a get ready signal, with a visual cue appearing at 2 s. The visual cue, a coloured rectangle with a fixation cross, appears for a period of 3.5 s, and during this activity period subjects would perform the search and select assessment task. Each experiments lasted 14 minutes. At the start of each session, a short demonstration was performed to introduce each subject to the experimental procedure and provide an overview of the task.

C. Data Acquisition

Eye tracking signals were recorded from an Arrington Research Eye Tracker system, comprising of a monocular

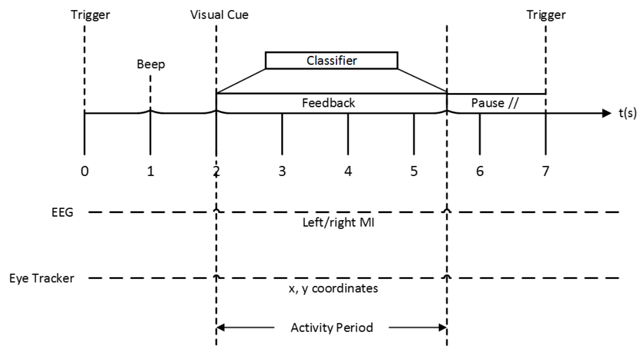

Fig. 4. Timing activity for a single trial.

high resolution infrared camera, with a sample rate of 128 Hz. The eye tracker was used for detecting the horizontal (x) and vertical (y) coordinates of gaze position, using four classes, each representing a single visual stimuli rectangle with a fixation cross.

III. METHODS

A. Gaze Detection

Eye tracker signals (gaze x and gaze y coordinates) were separated according to label indexes, which identified one of the four classes. These indexes specified the start of the activity period in each trial. A median filter was applied to the eye tracker data to remove blinks and the mean gaze position in horizontal and vertical directions were recorded for each trial. The distance of the x and y coordinates from each visual stimuli was calculated and the minimum distance from the x and y coordinates was assigned as the actual output result. The eye tracker model contained four labels, namely 1, 2, 3, and 4, representing each expected result. A confusion matrix was used to represent the expected and actual conditions, and used for computing the accuracy.

B. EEG Data

In this study, we consider the EEG signal from the BCI competition 2008-Graz dataset B to evaluate the performance of the proposed method. This dataset consists of EEG data from nine subjects. Three bipolar recording (C3, Cz and C4) were recorded with a sampling frequency of 250 Hz. All signals were recorded monopolarly with the left mastoid serving as reference and the right mastoid as ground. For each subject, five sessions are provided. Each trial is a complete paradigm of 8 seconds. In the result section, we only consider the results from the five best subjects as an approximation of the performance that can be obtained with subjects that could use a motor imagery based BCI.

C. Feature Extraction

1) Temporal filtering: Band power features for all the channels are extracted for mu (μ) (8-12 Hz) and beta (β) (12-30 Hz) bands. During each trial, the time over the imagery period are selected as the temporal features for the classification. The information obtained from event-related desynchronization (ERD) at certain frequency bands are useful for classification. ERD can be seen as a correlate of an activated cortical area, and presents itself as a decrease in power in certain frequency bands in the EEG, and is observable in the μ and central β rhythms.

2) Spatial filtering: Raw EEG scalp potentials are known to have poor spatial resolution due to volume conduction. If the signal of interest is weak while other sources produce strong signals in the same frequency band, it creates a problem. Common Spatial Patterns (CSP) is a data-driven technique to analyse multichannel data based on the recording from two classes, i.e., it corresponds to two different brain responses. CSP is used as a spatial filtering method in discriminating different classes in brain response detection problems [19]–[21]. CSP can be estimated and interpreted using the framework of Rayleigh coefficient maximization. In a nutshell, CSP filter maximizes the variance of spatially filtered signal under one condition while minimizing it for the other condition. CSP analysis is applied in order to obtain an effective discrimination of mental states that are characterised by ERD/ERS effects.

D. Classification

For the binary classification of the brain evoked response of left and right imagination, we consider support vector machines (SVM) [22] and linear discriminant analysis (LDA) with input features obtained after temporal and spatial filtering. The evaluation was performed with classifier training on the first session, and test on sessions 3, 4, and 5. In the next section, we report the mean accuracy, and the standard deviation across sessions.

In addition to the accuracy, we provide the information transfer rate (ITR) [23] in bits per minute (bpm) defined by $ITR = \frac{60}{T} \cdot \vartheta$ where ϑ, the information transfer rate, in bits per symbol, is defined as follows:

$$\vartheta \quad = log_2(N_{out}) + Plog_2(P) + (1-P)log_2(\frac{1-P}{N_{out}-1}) \quad (1)$$

and P being the probability of the good detection, *i.e.* the accuracy, N_{out} being the number of possible different outputs, and T being the time in seconds of recorded EEG signal that is required to take the decision among the N_{out} outputs.

IV. RESULTS

A. Gaze Detection

The performance of the eye tracker was evaluated using a single gaze data set collected from one experiment per subject. This resulted in 120 trials of gaze data per subject. The mean of the data elements and the distance of each x, y coordinate from each stimuli were calculated for the activty period of each trial (see Figure 4). The minimum distance was used for assigning each trial to a class, with a confusion matrix for visualizing performance and computing the accuracies. The offline experimental results for 5 subjects is summarized in Table I. The accuracy across subjects is 92.22±9.79. The target selections obtained from the eye tracker are shown in Figure 5.

B. Single-Trial Detection

Tables II and III present the accuracy and the ITR respectively, in bits per symbol, obtained with motor imagery with two classes, N_{out}=2 (left and right). In addition, we provide the ITR for the combination of gaze and EEG when the system

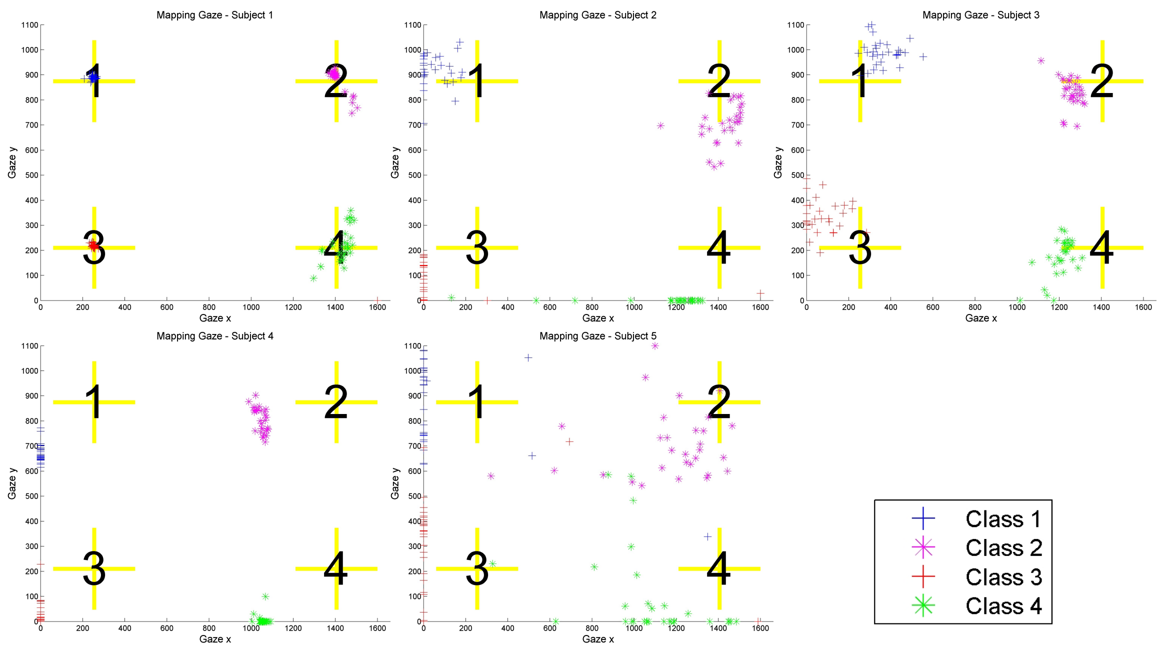

Fig. 5. Example of target selections obtained with the eye tracker. Yellow crosses represent the center of each target.

TABLE I. OFFLINE EXPERIMENTAL RESULTS.

	Accuracy (%)	Level of Expertise
Subject 1	99.10	Experienced
Subject 2	91.60	Moderate
Subject 3	100.00	Moderate
Subject 4	97.50	Novice
Subject 5	76.40	Novice
Mean± std	92.92±9.79	

is used for 4 and 8 commands, by combining the selected item, and motor imagery detection.

TABLE II. PERFORMANCE FOR SINGLE-TRIAL DETECTION WITH SVM.

	Accuracy (%)	ITR (bits per symbol)		
Subject		N_{out}=2	N_{out}=4	N_{out}=8
1	94.58 ± 5.09	0.70	1.61	2.54
2	62.50 ± 4.10	0.05	0.45	0.99
3	59.79 ± 7.71	0.03	0.39	0.90
4	83.54 ± 6.74	0.35	1.09	1.89
5	61.67 ± 2.37	0.04	0.43	0.96

TABLE III. PERFORMANCE FOR SINGLE-TRIAL DETECTION WITH LDA.

	Accuracy (%)	ITR (bits per symbol)		
Subject		N_{out}=2	N_{out}=4	N_{out}=8
1	93.13 ± 3.48	0.64	1.53	2.45
2	59.79 ± 0.72	0.03	0.39	0.90
3	61.25 ± 3.48	0.04	0.42	0.95
4	79.58 ± 7.96	0.27	0.95	1.70
5	72.50 ± 1.65	0.15	0.72	1.38

Figure 6 depicts the minimum accuracy that has to be achieved for motor imagery with different number of commands, from 2 to 4, when considering a perfect accuracy for the eye tracker, in order to increase the ITR. Particularly, a minimum accuracy of 86% should be achieved in order to increase the ITR with the proposed Hybrid BCI, when each selected item by the gaze can lean two different commands

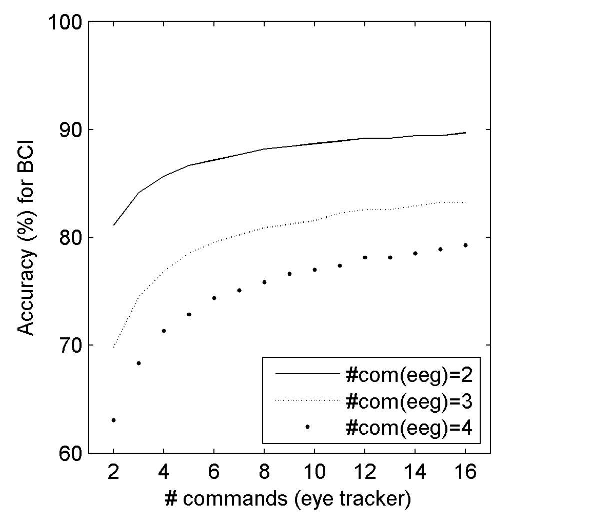

Fig. 6. Minimum accuracy to achieve to have an improvement of the ITR thanks to the addition of commands.

with motor imagery. With the current subset of subjects, only Subject 1 would obtain some gain in performance from a combination of BCI and eye tracker to increase of the number of commands, compared to the use of an eye tracker only.

V. DISCUSSION

For patients who require an alternative communication mean, the combination of eye tracker and BCI appears as a judicious solution to avoid issues related to the dwell time, and false positives. BCI can be used as a switch i.e., just to enable a single item observed by the user, or it can increase the potential commands, thanks to the detection of different brain responses. However, the use of several modalities can increase the difficulty to use the system, mostly for disabled people. In this study, we have evaluated the performance that could be achieved with a Hybrid BCI based on simultaneous detection of gaze and eye tracker. Despite the advantage of the combination, the accuracy of motor imagery detection can be a bottleneck for the overall performance of the system. We have shown that as the performance of a Hybrid BCI depends on the accuracy of each modality, some users may not benefit from a Hybrid BCI when it is used to increase the number of commands, and in this case it is probably better to only consider an eye tracker or an SSVEP based BCI.

A common observation given to BCIs that rely on gaze control, such as SSVEP based BCI, is the possibility to use an eye tracker instead of a BCI, as it would provide a better accuracy and a more convenient interface and setup. Yet, for a large number of commands i.e., with a high precision of the gaze, an eye tracker requires a particular care for the calibration. Ideally, the eye tracker should be coupled with a camera that analyzes the scene and determine potential points of interest.

With a Hybrid BCI, the dwell time of the eye tracker is not suppressed but the selection is reinforced by the detection of a brain response. As the brain response requires a certain amount of signal to be detected, e.g., about 2s for motor imagery, then the possible reduction of the dwell time is limited by the BCI. In addition, the user has to master the lag that occurs between the stimulus onset and the detection of a command. An issue with motor imagery for the detection of two commands with left and right motor imagery is the Simon effect. For instance, it may be more likely to perform errors for the selection of items on the left side of the screen which requires right motor imagination. This effect should be taken into account for the design of an efficient Hybrid BCI that uses motor imagery and visual stimuli.

VI. Conclusion

In this paper, we have evaluated the type of performance that can be obtained when motor imagery and gaze control can be combined. The success and the improvements that can be obtained from this combination are largely dependent on the accuracy of both modalities. Despite the advantage of Hybrid BCI, and the possibility of increasing the number of possible commands thanks to the combination of different modalities, the system should be set in relation to the performance of each modality in order to optimize the final user's experience. Experimental data from the simultaneous acquisition of EEG and eye tracker signals, using all 5 subjects, has been recorded for analysis. We hypothesize that we can optimize the number of commands from each modality in order to improve the information transfer rate, as well as achieving more distinctly classifiable patterns, through the combination of modalities. The resulting Hybrid interface will also assist in generating commands with directional properties and thus support a BCI assistive device with greater accuracy and functionality.

References

[1] J. R. Wolpaw, "Brain-computer interface research comes of age: Traditional assumptions meet emerging realities," *Journal of Motor Behavior*, vol. 42, no. 6, pp. 351–353, 2010.

[2] 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, pp. 767–791, 2002.

[3] J. d. R. Millán, R. Rupp, G. R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, C. Neuper, K.-R. Müller, and D. Mattia, "Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges," *Frontiers in Neuroscience*, vol. 4, no. 161, pp. 1–15, 2010.

[4] H. Cecotti, "Spelling with non-invasive brain-computer interfaces - current and future trends," *J. Physiology-Paris*, vol. 105, no. 1-3, pp. 106–114, 2011.

[5] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller, "How many people are able to operate an EEG-based brain-computer interface (BCI)?" *IEEE Trans. Neural Syst Rehabil Eng.*, vol. 11, no. 2, pp. 145–147, 2003.

[6] C. Guger, S. Daban, E. Sellers, C. Holznera, G. Krausza, R. Carabalonac, F. Gramaticac, and G. Edlinger, "How many people are able to control a P300-based brain.computer interface (BCI)?" *Neuroscience Letters*, vol. 462, pp. 94–98, 2009.

[7] I. Volosyak, H. Cecotti, D. Valbuena, and A. Gräser, "Evaluation of the bremen SSVEP based BCI in real world conditions," in *Proc of the 11th International IEEE Conference on Rehabilitation Robotics*, 2009, pp. 322–331.

[8] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, pp. 510–523, 1988.

[9] H. Cecotti, "A self-paced and calibration-less SSVEP based brain-computer interface speller," *IEEE Trans. on Neural Systems and Rehab. Eng.*, vol. 18, pp. 127–133, 2010.

[10] B. Z. Allison, R. Leeb, C. Brunner, G. R. Müller-Putz, G. Bauernfeind, J. W. Kelly, and C. Neuper, "Toward smarter BCIs: extending BCIs through hybridization and intelligent control," *J. Neural Eng.*, vol. 9, pp. 1–7, 2012.

[11] S. Amiri, R. Fazel-Rezai, and V. Asadpour, "A review of hybrid brain-computer interface systems," *Adv. Human-Computer Interact*, vol. 2013, p. 18, 2013.

[12] G. Pfurtscheller, B. Z. Allison, G. Bauernfeind, C. Brunner, T. S. Escalante, R. Scherer, T. O. Zander, G. Mueller-Putz, C. Neuper, and N. Birbaumer, "The hybrid BCI," *Frontiers in Neuroscience*, vol. 4, no. 42, pp. 1–11, 2010.

[13] Y. Li, J. Pan, F. Wang, and Z. Yu, "A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 11, pp. 3156–66, Nov. 2013.

[14] E. Yin, Z. Zhou, J. Jiang, F. Chen, Y. Liu, and D. Hu, "A speedy hybrid BCI spelling approach combining P300 and SSVEP," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 2, pp. 473–83, Feb. 2014.

[15] V. Kaiser, C. Brunner, R. Leeb, C. Neuper, and G. Pfurtscheller, "Investigation of cue-based vertical and horizontal eye movements with electroencephalographic and eye-tracking data," *Clinical Neurophysiology*, vol. 120, pp. 1988–1993, 2009.

[16] A. T. Duchowski, *Eye tracking methodology: theory and practice*. London, UK, Springer-Verlag (London), Inc., 2002.

[17] S. A. Brandt and L. W. Stark, "Spontaneous eye movements during visual imagery reflect the content of the visual scene," *Journal of Cognitive Neuroscience*, vol. 9, pp. 27–38, 1997.

[18] R. C. Miall, G. Z. Reckess, and H. Imamizu, "The cerebellum coordinates the eye and hand tracking movement," *Nature Neuroscience*, vol. 4, pp. 638–644, 2001.

[19] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Müller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Proc Magazine*, vol. 25, no. 1, pp. 41–56, 2008.

[20] M. Krauledat, M. Tangermann, B. Blankertz, and K.-R. Müller, "Towards zero training for brain-computer interfacing," *PLoS ONE*, vol. 3, no. 8, p. e2967, 2008.

[21] H. Zhang, H. Yang, and C. Guan, "Bayesian learning for spatial filtering in an eeg-based braincomputer interface," *IEEE trans. Neural Networks and Learning Systems*, vol. 24, pp. 1049–1060, 2013.

[22] V. Vapnik, *Statistical Learning Theory*. Wiley, New York, NY, 1998.

[23] C. E. Shannon and W. Weaver, "The mathematical theory of communication," *Urbana, IL: University of Illinois Press*, 1964.