ELSEVIER



Contents lists available at ScienceDirect

Computers in Biology and Medicine



journal homepage: www.elsevier.com/locate/cbm

Quadcopter flight control using a low-cost hybrid interface with EEG-based classification and eye tracking



Byung Hyung Kim¹, Minho Kim, Sungho Jo^{*,1}

Department of Computer Science, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon, Republic of Korea

ARTICLE INFO

ABSTRACT

Article history: Received 15 January 2014 Accepted 29 April 2014

Keywords: Hybrid interface Brain-computer interface Mental concentration Eye tracking Quadcopter flight control We propose a wearable hybrid interface where eye movements and mental concentration directly influence the control of a quadcopter in three-dimensional space. This noninvasive and low-cost interface addresses limitations of previous work by supporting users to complete their complicated tasks in a constrained environment in which only visual feedback is provided. The combination of the two inputs augments the number of control commands to enable the flying robot to travel in eight different directions within the physical environment. Five human subjects participated in the experiments to test the feasibility of the hybrid interface. A front view camera on the hull of the quadcopter provided the only visual feedback to each remote subject on a laptop display. Based on the visual feedback, the subjects used the interface to navigate along pre-set target locations in the air. The flight performance was evaluated by comparing with a keyboard-based interface. We demonstrate the applicability of the hybrid interface to explore and interact with a three-dimensional physical space through a flying robot. © 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Over the past decade, various interface systems have been explored that enable not only healthy but also physically impaired people to immerse themselves in images or videos and interact through voluntary movements such as moving their eyes in various directions. Among the diverse approaches, eye tracking techniques and brain-computer interfaces (BCIs) are those in the spotlight. The two paradigms enable intuitive and natural interaction. This paper proposes a hybrid interface where users' eyes and mental states directly influence control systems. In particular, we present a noninvasive and low-cost wearable interface to control a quadcopter in three-dimensional (3D) space.

When a human being is able to move their eyes voluntarily, eye tracking is a promising interface scheme [1]. Eye movement recording can be interpreted as control commands without the need of a keyboard or mouse input. In contrast to conventional inputs, eye tracking is wearable. Hence, it can be advantageous for not only physically impaired people but also healthy people in the sense that both hands are still available. It has a high information transmission rate, short preparation time, and omnidirectional controllability [2]. In most of the previous eye movement based tracking work, they usually

http://dx.doi.org/10.1016/j.compbiomed.2014.04.020 0010-4825/© 2014 Elsevier Ltd. All rights reserved. track pupil contour and estimate eye gaze with a monocular camera. Though such approaches may have been shown to be effective in controllability in the two-dimensional (2D) space, they are not necessarily able to be used in 3D space because depth information is not extractable.

Among BCI approaches, electroencephalogram (EEG)-based BCIs enjoy many advantages such as inexpensive equipment, less environmental constraints, and noninvasive measurement [3]. In terms of classification accuracy, recent studies have demonstrated EEG-based BCIs allow users to control machines with multi-states classification [4–6]. Nevertheless, from the perspective of practical applications, consistently classifying multi-class BCI is not easy because of the difficulty of human subjects in maintaining mental concentration and the intensive training required [7]. In this sense, two-class BCI classification is favorable due to its relative simplicity, but has a limit of controllability. To overcome the problem and improve practical applicability, the hybrid BCI paradigm, which uses a brain activity-based interface together with any other multimodal control device, has been a popular topic of investigation [8]. The hybrid BCI paradigm is expected to be a good choice for practical implementation with feasible hardware [9]. Among various choices of control devices for hybridization with BCI, eye-tracking devices are of particular interest because of their omnidirectional controllability in the 2D space.

There were attempts [10,11] to attain enhanced interfaces by hybridizing eye tracking and EEG-based classification. The two schemes could be compensatory in function. These attempts were made based on an idea that EEG-based classification can be a way

^{*} Corresponding author.

E-mail address: shjo@kaist.ac.kr (S. Jo).

¹ Tel.: +82 42 350 3540; fax: +82 42 350 3510.

of commanding selections while eye tracking is used to point at target spots. However, its applications to date are limited to the 2D physical space such as cursor control on a computer screen.

To further extend the hybridization's potential for asynchronous interaction of the exploration of a subject's surroundings, considering applications in a 3D physical environment are valuable. Furthermore, applications to challenging tasks such as controlling the flight of an airborne object [12], which is not an easy task even with a nominal interface, would be a more valid study. In this study, fresh look at a hybrid interface combining eve movement and brain activity is considered as way of extending control commands for 3D physical applications: specifically in this study. quadcopter flight control. The combination of eve movement and mental concentration is mapped to all possible flight actuations of a quadcopter. Hence, a human subject can control the quadcopter remotely while simply looking at the computer monitor keeping at a safe distance from its air space. The subject also obtains realtime environmental information from a camera mounted on the guadcopter. Therefore, he or she can make decisions on the flight's direction based on this information.

The other consideration of this study is the interface's practical applicability [13,14]. Currently, standard eye trackers and EEG acquisition systems are very expensive and bulky. They are still mainly used in laboratory environments. To enhance real-life usability, we prepare a low-cost hybrid interface for the present study. In addition, the hybrid interface aims to be easily wearable. Hence, the aim of this study is the development of a low-cost and

easy-to-use hybrid interface that interprets eye movements and brain mental activity to allow real-time control of applications in 3D environment. Therefore, the contribution of the proposed system, as an alternative to existing work, is a new interface that addresses the limitations of the previous systems in a single system. The main contributions of this work are two-fold:

- 1. Easy-to-learn and easy-to-use system: hybridizing eye tracking and EEG-based classification allowing users to complete their tasks easily with various commands in 3D physical space.
- 2. Not only low-cost but also a convenient wearable device: people can control their flight naturally in everyday life.

The rest of this paper is organized as follows: Section 2 presents a detailed method of this proposed work. Sections 3 and 4 present experimental setup and results. Section 5 presents some conclusions and discussion.

2. Methods

2.1. Hybrid interface system

Fig. 1 shows the overview of our proposed hybrid interface. It consists of an EEG acquisition headset and a custom-built eye tracker based on previous studies [14,15]. The eye tracker's total cost was less than \$40USD. The eye tracker consists of two



Fig. 1. The overview of our proposed hybrid interface.



Fig. 2. Layout of the software architecture.

components: an infrared camera, and light-emitting diodes (LEDs). Five LEDs are fixed around the lens of the infrared camera that is connected to the frame of a pair of glasses about 8 cm away from the left eye. When a subject wears the glasses, the camera is pointed toward the left eye to capture its image. LEDs illuminate the eye to enhance the contrast between the pupil and the iris. The camera captures eye images sequentially with a spatial resolution of 640×320 pixels at sampling rate of 33 Hz and the image stream is transmitted to a computer via a USB. From the sequential eye images, eye gaze is estimated on the computer.

A low-cost commercial EEG acquisition headset (Emotiv Epoc, US), which is also shown in Fig. 1, is combined with the eye tracker. The headset is significantly less expensive in comparison with the state-of-art EEG acquisition systems. Its reliability as an EEG acquisition system has been studied in some previous papers [9,11,16,17]. The EEG acquisition headset consists of 14 electrode channels plus CMS/DRL references around the sensorimotor cortex. In this study, EEG signals from the headset are used to classify a subject's mental concentration.

The combination of the low-cost eye tracking and EEG acquisition systems comprises the hybrid interface, which takes advantages from the two systems. Fig. 2 shows the layout of the software architecture in this study. The lowest layer is formed by the hardware layer, which generates raw data in the form of a video stream, image sequences, and EEG signals from devices. This information is then processed by the kernel space layer. Its task is to generate a stream of eye position from the raw image data and filtered brain signals from EEG signal data. Note that this layer includes transformation from device to screen coordinates, but a perspective transformation such as a calibration procedure is excluded. The final part of our framework is the user space layer. In this laver, in order to interpret eve movement and mental states. calibration phases and classification procedures are conducted. The user space layer also has the task of generating visible output for the user via a Graphical User Interface (GUI). It receives events from Drone SDK, Emotiv SDK, and Gaze Tracker as well as displays its output on a monitor. Implementation of this layer has been programmed in C# under the .NET framework.

2.2. Algorithms

Once the camera initially grabs an image, our eye-tracking algorithm proceeds based on previous studies [14,15]. First, the image is binarized, and then eye features such as points inside detected contour between the pupil and iris are extracted. Applying a RANSAC procedure to eliminate outliers, the center of the pupil is determined at the center of an ellipse fitted into the feature points. These points are then mapped to eye gaze through

a calibration process. In the calibration process the coefficients for interpolating gaze points on horizontal and vertical axes are calculated using a second-order polynomial [18]. Nine points on the top-left, top, top-right, left, center, right, bottom-left, bottom, and bottom-right are displayed on the screen. The user is instructed to fixate his/her gaze on each of the 9 target points in sequence. Each time, the location of the center of the pupil is recorded. Given the 9 corresponding points from the center of the pupil, the pupil vector $p = (x_p, y_p)^t$ is transformed to a screen coordinate $s = (x_s, y_s)^t$ as follows:

$$x_p = (a_0 + a_1 x_s + a_2 y_s + a_3 x_s y_s + a_4 x_s^2 + a_5 y_s^2)$$

$$y_p = (a_6 + a_7 x_s + a_8 y_s + a_9 x_s y_s + a_{10} x_s^2 + a_{11} y_s^2)$$

where a_i are the coefficients of this second order polynomial. From the above equation, an over-determined linear system with 18 equations is obtained. The coefficients of the system can be solved using a least squares method.

Collected brain activity signals are processed in real-time to classify two mental states: intentional concentration and nonconcentration states. To discriminate the two states, we use the Common Spatial Pattern (CSP) algorithm [17]. CSP is a technique to analyze multichannel data based on recordings from two classes. Let $X \in \mathbb{R}^{L \times M}$ be a segment of an EEG signal where *L* is the number of channels and *M* is the number of samples in a trial and $x(t) \in \mathbb{R}^{L}$ be EEG signal at a specific time *t*. Hence *X* is a column concatenation of *x*(*t*). Given two sets of EEG signals from intentional mental concentration and non-concentration states, CSP yields decomposition of the signal parameterized by a matrix $W \in \mathbb{R}^{L \times L}$. The matrix *W* projects the signal x(t) to $\hat{x}(t) = W^T x(t)$. CSP filters maximize the variance of the spatially filtered signal under one condition while minimizing it for the other condition. The mathematical procedures to find the spatial patterns are well formulated in previous studies [17]. We use the Burg method based on an autoregressive (AR) model to estimate the power spectrum of the projected signals. Based on previous reports, we fix the model order of the AR model to be 16 and set the power values of a feature vector between 11 and 19 Hz. According to the literature [19,20], brain activity during mental concentration or alertness comprises distinct rhythms over frontal areas in the beta wave band whereas it has much weaker amplitude in the alpha wave band over the occipital lobe with eyes closed. For optimally classifying feature vectors acquired from training session, we use a Support Vector Machine (SVM) algorithm, which is a supervised learning method used for classification and regression. It is known to be good at generalization and has popularly been used for EEG-based classification [21]. In this work, the linear kernel-based SVM was applied.

For robust classification during real-time execution, data points acquired from each 1 s time window with 125 ms increment are



Fig. 3. Quadcopter and its control windows. (a) Quadcopter's directions indicated by arrows. (b) Control window design in GUI. In mode A the central area of the GUI is colored green and it carries out four movements: front, left, right, and back hover. On the other hand, in mode B, the area is colored yellow carrying out four movements: ascend, descend, left turn, and right turn. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

used to classify a mental state. Furthermore, the final confirmation of an intended state is made when an identical state is repeatedly selected from two sequential windows after classification. Therefore, a state selection confirmation takes 250 ms if the selection is successful and the confirmation is used to update the selection of the modes A and B.

2.3. Quadcopter flight control

We set the eight directions of the quadcopter as shown in Fig. 3(a): up and down, left and right, forward and backward translation, and left and right turns. Fig. 3(b) shows the GUI of our system on the monitor. In the GUI window, visual feedback is displayed in the central rectangular region. The top, bottom, left and right regions are partitioned. A subject can cause the quadcopter to take off by concentrating mentally with their eyes closed. Once the quadcopter takes off, two control modes, A and B, can be only altered by concentrating mentally while gazing at the central area. The two modes are distinguished by the color of a selected region in the GUI window. Its colors are green and vellow respectively while the control modes A and B are selected. In either control mode, a user can select one of the directions through placing his/her eyes on the desired area on screen. In control mode A, looking at the top and bottom areas in the GUI window moves the quadcopter forward and backward respectively while looking at the left and right areas moves it in the left and right directions respectively without rotation. In control mode B, looking at the top and bottom areas moves the guadcopter up and down respectively while looking at the left and right areas makes it turn to the left and right respectively. In each mode, our system keeps the selected movement unless the user is looking at different areas. A subject can land the quadcopter on the ground by concentrating mentally with eyes closed while the quadcopter floats up in the air. The quadcopter includes a three-axis accelerometer. Using the sensor, flight trajectory can be estimated.

3. Experiments

Our entire experimental process for controlling a quadcopter in 3D space is as follows: after initial calibration of the eye tracker, subjects are asked to help train two modules separately. Then, the subjects are tested whether they are ready to fly a quadcopter or not using the proposed hybrid interface under specific criterion in virtual environment. If they pass this test, they finally perform the actual test trials.

3.1. Data acquisition and processing

Five healthy subjects (age 24.44 ± 3.02 (mean \pm SD) years) without any prior experience with eye tracking and EEG-based BCI were participated in experiments. They were informed about the entire protocol and aims of the study, and they all signed written informed consent forms. The KAIST Institutional Review Board approved the proposed experimental protocol of this study. Each participant was seated comfortably in a chair facing a computer display, which was placed about 1 m in front of the subject on a table. Each subject wore the hybrid interface. The head mounted eye tracker captured images of eye movements with a spatial resolution of 640×320 pixels at sampling rate of 33 Hz. The EEG acquisition headset records data at sampling frequency of 128 Hz from a 14-channel layout (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4), which is shown in Fig. 4.

Fig. 4 shows the paradigm and the statistical analysis of the BCI control. The subjects sat comfortably in an armchair, facing a computer screen. The duration of each trial was 8 s. As shown in

Fig. 4(a), during the first 4 s, while the screen was blank, the subject was in the non-concentrated state. Immediately after these first 4 s, a visual cue (two crosses) was presented on the left and right side of the screen, indicating the concentration task to be performed during 4 s. Once the visual cue disappeared, one trial was completed. As indicated in Fig. 4, 0–4 s and 4–8 s segments in a trial were selected to represent intentional non-concentration and concentration states, respectively. During the two states, the subjects are required not to move their eyes, so that the electrooculogram (EOG) artifacts are minimized as much as possible. For each subject. CSP was performed on the data under the two states separately. For each state, data of all trials were concatenated to a 40-s (10 trials \times 4 s) long data segment. CSP resulted in two sets of spatial filters corresponding to the concentration and nonconcentration states. Fig. 4(c) visualizes brain activity between concentration and non-concentration states over frequencies with respect to power spectrum density (PSD) per subject at a frequency band between 11 and 19 Hz, in which feature vectors were assigned. Each plot was obtained by averaging brain activity of each state. Each density difference at a particular electrode spot is visualized. It explains how brain activity is different at different frequencies on different lobes of the brain as described in Section 2. Brain signals related to the non-concentration state from most subjects were activated in the alpha wave frequency band between 11 and 13 Hz displaying colors between blue and green over the occipital lobe. In contrast, the concentration state-related signals are activated in the beta wave frequency band between 13 and 19 Hz displaying colors between red and orange over the frontal lobe. For example, brain signals of subjects S1, S2, S3, and S4 have relatively activated at low frequencies between 11 and 13 Hz over the occipital lobe during the non-concentration state while the signals increase over the frontal area at relatively high frequencies between 16 and 19 Hz. However, these brain phenomena related to the two states are not always common among all subjects. In the case of subject S5, it is not easy to distinguish the difference visually in the frequency band over the frontal lobe. We overcome the inter- and intra-subject variability of brain signals through analyzing them statistically. Fig. 4(b) shows the statistical analysis for a subject. It explains how CSP works in 2D and SVM discriminates the two mental states. In the figure, the distribution of samples after the filtering is also shown and the solid line separates two mental states with maximum margin.

The eye tracker was calibrated to each participant according to the standard procedures. This requires the user to look at 9 points on the screen, and it takes less than one minute to perform. Before conducting experiments, the eye tracker and mental concentration parts of our hybrid system were examined to estimate their accuracy individually. Table 1 summarizes the evaluation results. Over five subjects, an average error of 0.512 cm with the angular error of 0.284° is reported. These results are on par with commercially available eye tracking systems such as the binocular eye-tracker EyeLink II(\$25,000USD) which was reported to be less than 0.5° [22]. In the case of mental concentration accuracy, the classification success rate was above 91.67%. At the center area of the screen, periods of concentration and non-concentration were repeated 12 times randomly with green and yellow colors respectively.

Finally, prior to controlling a quadcopter in the real world, subjects were asked to familiarize themselves to the proposed hybrid interface by way of 3D control tasks until they achieved 80% or more accuracy. In the 3D control tasks, one of the eight tasks was randomly chosen and appeared in the center of the control window. The subject was instructed to perform the task within 5 s. One session consists of 10 trials, which lasted up to 45 s. Subjects were attended for three days and they were given a maximum of five sessions each day. They should complete at least eight out of 10 trials in a session in order to show their readiness. If they failed



Fig. 4. Experiment paradigm and statistical analysis for the brain-computer interface component. (a) BCI experimental paradigm for single trial. (b) An example of SVM classification after CSP filtering in 2D. (c) Visualization of brain activity difference between the non-concentration and concentration state per electrode over the frequency band from 5 subjects.

 Table 1

 Accuracy of the eye tracking and BCI systems of the hybrid interface.

Subject	Eye calibration	Angular eye calibration	BCI classification
	error (cm)	error (°)	accuracy (%)
1	0.66	0.36	91.67
2	0.58	0.32	83.33
3	0.49	0.27	100
4	0.31	0.18	91.67
5	0.52	0.29	91.67

to satisfy this criterion, the users were required to repeat their training of mental concentration, calibration of their eye movements, and testing of 3D control tasks. EMG contamination arising from cranial muscles was often present early in BCI training and early target-related EMG contamination might be the biggest reason for unsuccessful trials [23]. However, this EEG contamination of all participants gradually waned over the 5 sessions in these 3D control tasks. Note that we also have some discussion about this EMG contamination in the discussion session. As well as EMG, and EOG artifacts in the frontal electrodes were also considered due to eye movements. In our experiments, instead of using filtering algorithms, we let subjects fail trials if EOG artifacts were significantly present in the given trials since we already required the subjects not to move their eyes.

3.2. Experimental set-up and task

Two sets of experiments were conducted navigating a commercial quadcopter (Parrot AR drone 2.0, Parrot SA., France). In the first set of experiments, the keyboard-based interface system was tested; in the second set of experiments, the focus of the tests was the hybrid component. The flying robot is equipped with a forward facing HD camera (720p, 30ftps) that provides visual feedback to a subject. The real-time video stream from the camera is transmitted to our system via Wi-Fi and actuation control signals from our system are sent to the quadcopter wirelessly so that it moves following the user's commands.

In order to conduct air vehicle flight control experiments, an adequately large physical space was required. We set up the experimental environment in a school gymnasium. Fig. 5 illustrates the layout of the physical environment. Balloons were used because they are safe against collisions and easy to handle. In the



Fig. 5. Experimental set-up. (a) The layout of the experimental set-up. (b) Real experimental set-up in a school gymnasium.

air, seven helium-filled balloons were hung at different heights, which were aimed to test vertical movements of the quadcopter. As shown in Fig. 5(a), the balloons were located properly so that users could draw a figure-eight trajectory flight path. Drawing a figure-eight trajectory flight path was requested in order to test horizontal movements of the quadcopter including its turns. Its entire flight plan is set as follows. First, the quadcopter took off at one meter away from the starting point (indicated as 0 in Fig. 5) of the gymnasium. It passed the first three balloons (indicated as 1, 2, and 3 respectively in Fig. 5) that were hung at 3, 5, and 3 m off the ground respectively. Then, it headed back to its starting point. Second, it passed the next three balloons (indicated as 4, 5, and 6 respectively in Fig. 5) with the same heights as the first three balloons. Finally, the journey was finished once the guadcopter landed on the initial departing spot (indicated as 0 in Fig. 5). One practical problem came from the fact we only provide visual feedback to each user. Due to limited field of view, it was difficult for users to drive their quadcopter toward a target naturally with vertical and horizontal movements at the same time. To overcome this problem, users should climb in straight vertical flight so as to secure a clear view toward their next target, and then move forward. Relative distances of the balloons are indicated in Fig. 5 (a). Human subjects were seated in front of a computer, facing away from the quadcopter, wearing the hybrid interface one meter away from the flight environment. They could not directly observe the flight, instead, they had to control the flight based only on the video stream displayed on the monitor. The front view HD camera mounted on the quadcopter provided the visual information on its environment as mentioned previously.

While considering their current environment, human subjects generated desired commands by looking at the designated regions on the monitor and concentrating mentally through the hybrid interface. Corresponding to confirmed commands, the quadcopter was actuated via wireless communication in real-time.

3.3. Performance evaluation

To evaluate the flight trials taken using the hybrid interface, we not only referred to the experimental control and performance analysis in [4] as a baseline comparison, but also we adopted six evaluation methods: *travelled distance*, *total time taken*, *normalized path length*, *Area Under the Curve*, *Speed of control*, *and Success rate* for evaluation purposes. These are used to measure users' sensitivity of controllability in terms of agility and smoothness respectively. Our goal was to test whether the proposed system sufficiently would track eye movements and classify mental states simultaneously as a user carried out complicated tasks to control



Fig. 6. Normalized path length.

their flight. Note that we also provide some discussion between the previous study [4] and us in the discussion section. We first calculated *traveled distance* (TD) of the flight trajectory and its *total time taken* (TT) until it arrived at the final destination. Two measurements were used to assess the performance of how fast each user finished the whole voyage in each experiment. This reflects the capability of users to identify targets properly and move the robot toward the targets quickly.

The normalized path length (NPL) was introduced to measure how users can control their flight smoothly [24]. It indicates how straight a trajectory is. In other words, we calculated how different a trajectory was from the line segment connecting its extremes. As indicated in Fig. 6, this difference can be measured by simply comparing the path lengths of the trajectory l_1 and l_2 with the straight-line segments d_1 and d_2 .

$$NPL = \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{l_1^{i,j} + l_2^{i,j}}{d_1^{i,j} + d_2^{i,j}},$$

where i and j are the indices of the starting and target positions respectively and n is the number of targets. The normalized path length was defined as the ratio of the two. The closer the similarity to 1, the better.

We used the *Area Under the Curve* (AUC) of the Receiver Operating Characteristic (ROC) to assess the performance of the machine-learning component of our system. Based on the trade off between false positive and false negative the AUC essentially characterizes the behavior of a classifier. An AUC score is 1 when a perfect classifier is one that produces zero false positive and false negative. On the other hand, a random classifier would have an AUC of 0.5. So, the closer the AUC is to 1, the better classifier it is. In our work, prior to controlling the quadcopter in physical 3D space, subjects were instructed to conduct one test in a virtual environment. Control commands were randomly chosen at every four seconds from the eight commands that are available in our system. Subjects should complete the command within four seconds and then our system classified their behavior. A total of 15 control commands were given during 1 min. The results of classification were used to calculate the AUC using the efficient method proposed in [25].

Success rate (SR) was assessed by calculating the number of target passed within the limited time divided by the total number of targets. Limited time was given by dividing the distance between two points by the given constant forward velocity of the quadcopter. If the subject successfully controlled the guadcopter to pass a target within the limited time, one target arrival was awarded. The SR was evaluated to demonstrate how many tasks were completely perfect during one flight. The SR was calculated by

SR =	Number of targets passed within the limited time				
	Total number of targets				

4. Results

In this section, we describe the experimental validation of our interface comparing with the results obtained by a keyboard-based interface. Table 2 summarizes the results over the five subjects.

4.1. Sensitivity of controllability in virtual environment

Table 2 reports the values of the AUC score resulted from one test in a virtual environment using the hybrid interface.

Table 2

The summarized performance of the keyboard-based interface and the hybrid interface.

subject	Keyboard-based interface			Our hybrid interface					
	TD	TT	NPL	SR	TD	TT	NPL	AUC	SR
1	1.17	1.23	1.14	0.95	1.54	1.71	1.44	0.93	0.90
2	1.46	1.53	1.31	0.90	2.12	2.11	1.89	0.78	0.85
3	1.89	2.79	1.52	0.71	2.54	2.76	2.11	0.81	0.56
4	1.52	1.67	1.48	0.90	2.33	2.32	2.04	0.79	0.66
5	1.57	1.53	1.53	0.90	2.07	2.31	2.09	0.83	0.76



The achieved results ranged from 0.78 to 0.93 and they showed the performance of our proposed interface would be an efficient method of control in 3D space. Generally, subjects who had achieved high AUC demonstrated a high sense of controllability in terms of other metrics. However, some users such as subject 3 who could not control the guadcopter in the physical space as much as what he/she had done in a virtual environment. Compared with subjects 2 or 4 who had a similar AUC as subject 3, subject 3 reported a much lower SR than the other two subjects. The ability to adapt to new environments would be one reason why this occurs. Fig. 7 reports experimental results over three consecutive sessions. Subject 3 initially suffered from an inability to adapt to new environments in order to control the quadcopter. but his/her sense of controllability in 3D physical space increased over consecutive experimental sessions. Hence, in the last experimental session, subject 3 achieved a similar SR as subjects 2 or 4.

4.2. Subject operational agility

The SR in Table 1 reports that all five subjects were able to control the quadcopter quickly and smoothly compared to the keyboard-based interface. On average, subjects showed 85.55% of the SR that was acquired via the keyboard-based interface system. Therefore, the SR also shows how our system can be an alternative solution with respect to usability. For the purpose of this study, the velocity of the flying robot for forward movement was set to a constant. However, rotating and adjusting the altitude of the quadcopter changes its velocity [4]. For example, the velocity increased when making large turns in a single direction but decreased when temporarily changing directions. In the case of subjects 2 and 5, the differences between two SRs are small, but TTs are relatively large. These differences indicate that subject 2 carried his/her flight with large turns with less hesitancy than subject 5.

4.3. Subject operational smoothness

Subjects were only given visual feedback for controlling the quadcopter in 3D space. Therefore, in order to navigate their flight in a physical space, users' flight smoothness is important. From the results of NPL, our proposed interface demonstrated a 26.75%

TT



Fig. 7. Experimental performance metrics of the subjects over three consecutive sessions. The closer the similarity to 1, the better.



Fig. 8. Flight trajectories of five subjects using keyboard-based interface and our hybrid interface. (a) Flight trajectories in the horizontal plane (Top view). (b) Flight trajectories in the vertical plane (Side view). For each subject, red solid lines show his/her trajectories using our hybrid interface over three sessions. Blue lines indicate trajectories using the keyboard interface over three sessions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

deviation averaged over the five subjects from the trajectories that were achieved via keyboard control, which indicate good performance in terms of reliability for an alternative interface. Fig. 8 shows flight trajectories for five participants with different views. As shown in Fig. 8(b), most users were able to smoothly change their altitude whereas as some of them could not navigate as smoothly in the horizontal plane as shown in Fig. 8(a). This seems to be obvious because navigating the quadcopter vertically requires only two movements in mode B, but it requires six movements in both modes A and B in order to control the quadcopter within a horizontal plane. In particular, most users turned the guadcopter in a stuttering fashion when they needed to face their quadcopter towards the next target. This is in contrast to simply going forward towards the target which was relatively smoother. This is more than likely due to multiplexing the use of gaze as both a control mechanism as well as a searching mechanism. The subject, while turning the quadcopter would be also actively searching for the next target within the displayed video. This would turn their gaze away from the control area, stopping the turn in progress. After unsuccessfully finding the next target within the video, they would resume the turn by looking at the control area again. This cycle would cause turns to stutter.

4.4. Experimental control results

Fig. 9 shows the ratios of average performance metrics for our proposed hybrid interface system compared to the keyboardbased interface system. This figure shows our proposed hybrid interface system would, overall, be a good alternative interface compared with keyboard-based interfaces. In particular, our system demonstrated its robustness with respect to completeness of tasks (SR) by achieving 85.55% of the success rate that was achieved via the keyboard-based system. Fig. 8 indicates the there is a greater difference between the ratios of TT. TD. and NPL than SR. TT, TD, and NPL increased when the subjects became disoriented and lost their bearings while going to the next target. In the experiments when they made mistakes during the flight, it took some time to reorient the quadcopter toward the next target again. Moreover, with respect to mind control, when the subjects missed their next target, they usually started to suffer from anxiety, so it became difficult for them to concentrate on the hybrid interface. As shown in Fig. 7, most subjects were able to



Fig. 9. The ratios of average performance metrics for our hybrid interface system compared with the keyboard-based interface system.

increase their flight accuracy from three consecutive sessions. This raises some questions as to whether additional training could result in our hybrid interface having timings that are more similar to those of the keyboard-based interface.

4.5. Real-time performance

Both the quadcopter and the eye tracker captured images at 33 Hz. On the other hand, EEG signals were retrieved every 125 ms, while retaining 1 s worth of previous data. The mental concentration requires an identical classification result from two sequential data windows. Therefore, a state selection confirmation takes 250 ms if the selection is successful. We stress that our system is capable of real-time operation: the average processing needed in each of the eye tracker and mental concentration is 117 ms and 84 ms on a dual-core 2.80 GHz CPU, which is less than the average EEG signal period.

5. Discussion

The results from this study demonstrated that only using eye movements and mental concentration without formal body movements was still good enough to execute a challenging task: flight control of a quadcopter within a 3D physical space. Self-developed software and hardware except the EEG acquisition headset found that an inexpensive interface was possible. It was observed that the hybrid interface achieved compatible performances to nominal keyboard controlled cases. It is obvious that the present interface is easily applicable to other 3D environmental applications.

We want to remark that eye movements and concentration are natural and intuitive. Because the hybrid interface is based on these natural actions, its usage can feel just as second nature as people get used to wearing the device. In particular, the present scheme is beneficial for physically disable people. EEG- based binary classification is simple and trainable with few complications.

Since our hybrid system fuses eye movements and mental concentrating, signals from eye movements (EOG) and some cranial muscle movements (EMG) can be orders of magnitude larger than brain-generated electrical potentials. Furthermore it can be one of the main sources of artifacts in EEG data. In our study, as already described in the experimental section, these artifacts are dealt with by discarding contaminated trials. On the other hand, there are some other alternatives to remove the artifacts such as blind source separation, which is a signalprocessing methodology that includes independent component analysis (ICA) for removing electroocular artifacts from EEG data. The idea behind this approach analysis is to generate distinct components between neural and non-neural activity such as externally generated noise or an ocular artifact. These produced components are not related to signal propagation or head models, which means the data are blindly processed. The advantage of this algorithm is that it is not affected by errors in head propagation models. The disadvantage is that there is no guarantee that any algorithm of this method can capture the individual signals in its components. So far, the algorithm cannot yet separate individual signal components perfectly. In other words, there are still some leakages between the ocular and non-ocular components even though the algorithm separates the ocular activity cleanly. Some eye motions may be reflected in several components representing the ocular sources. Therefore, a distribution of sources falls in the vicinity of the eyes, rather than on the eyes themselves when the components are mapped geometrically.

In each of the two different standalone EEG-based and eye tracking-based systems, a lot of research has been done to achieve multiple classification [5,6]. Despite their own advantages, both

systems are not yet ready to work in everyday life with respect to usability. In [4], authors demonstrated the possibility of controlling a quadcopter in 3D space using an EEG-based standalone system. The major advantage of their system compared with ours is they have less possibility to have contaminated EEG data by EOG and EMG since they only used EEG signals retrieved from a high cost system. However, it requires intensive training in order to classify multiple states accurately. In particular, this could become more a serious problem when current wearable BCI devices that have a relatively small number of channels are used for practical applications. Furthermore, classifying multiple states using sensorimotor rhythms is not always available for everyone [26]. Therefore, EEG standalone systems are less suitable than ours with respect to usability.

In the case of eye tracking-based systems, it has a distinctive advantage that it can quickly determine where the user's interest is focused automatically. However, this approach cannot be used for depth directional navigation in 3D. There might be also an alternative way through replacing bi-state classification/discrimination (concentrated and non-concentrated) by other eye movements, such as blinking. However, in this case, safety issues should be considered for controlling a quadcopter.

While EMG and EOG act as artifacts in EEG data physically, unstable mental states due to anxiety, fatigue, and frustration produce psychological artifacts causing inconsistent EEG patterns [27,28]. In our experiments, users who suffered from anxiety also had difficulty concentrating on the hybrid interface. One of the main reasons was because the viewing angle through the camera was too narrow to give the user good situational awareness. One possible way to overcome the problem is obviously to provide a wide-angle view to users from multiple cameras. The multiple cameras would help participants find their missing target quickly and this would reduce anxiety. The other way is to analyze eve movements for detecting excessive anxiety. People tend to move their eyes rapidly and abnormally when they feel anxious. Therefore, modelling abnormal eye movements would be a potential solution to detect users' anxiety. Researchers have been trying to apply different signal processing techniques for BCIs in an attempt to improve the signal-to-noise ratio of the input signal [29]. Other studies train users to control their EEG patterns through extensive and resource demanding neuro-biofeedback training [30,31].

6. Conclusion

In this paper, we proposed a wearable hybrid interface which tracks eye movements and interprets mental concentration at the same time. The proposed interface enables both disabled and fully capable people to complete their tasks with various commands in 3D physical space. Through this low-cost and easily wearable device, people can control their flight naturally and easily in everyday life. From the results of our study, we have also successfully demonstrated the potential of hybrid control with potential applications for hands-free control in both disabled and fully capable people. For further application of the hybrid interface with eye tracking and brain activity, integrating various neural signal responses such as P300-based potentials [32], sensorimotor rhythms [33], or steady-state visual evoked potentials (SSVEP) [34]. Sawith eye tracking will be attempted in the future.

7. Summary

The proposed hybrid interface enables users to navigate their quadcopter in three-dimensional space with movement of their eyes and mental concentration. The hybrid system is low cost and easily wearable, hybridizing eye tracking and EEG-based classification. It sufficiently tracks and interprets two inputs to augment the number of control commands to enable the quadcopter to travel in eight different directions within the physical environment. In experiments, we demonstrate the feasibility of the hybrid interface compared with a keyboard-based interface. The results from this study show our proposed interface would, overall, be a good alternative interface compared with keyboard-based interfaces.

Conflict of interest statement

None declared.

Acknowledgments

This work was supported by Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education (2013R1A1A2009378) and by the MOTIE (The Ministry of Trade, Industry and Energy), Korea, under the Technology Innovation Program supervised by KEIT (Korea Evaluation Institute of Industrial Technology), 10045252, Development of robot task intelligence technology. The authors appreciate anonymous reviewers and editors for their valuable comments.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.compbiomed.2014.04.020.

References

- A.T. Duchowski, Eye Tracking Methodology: Theory and Practice, 2nd Ed., Springer-Verlag, London, 2007.
- [2] R.J.K. Jacob, K.S. Karn, Eye tracking in human-computer interaction and usability research: ready to deliver the promises 4 (2003) 573–607Mind's Eye: Cogn. Appl. Asp. Eye Mov. Res. 4 (2003) 573–607.
- [3] J.R. Wolpaw, et al., Brain-computer interface technology: a review of the first international meeting, IEEE Trans. Rehabil. Eng. 8 (2) (2000) 164–173.
- [4] K. LaFleur, K. Cassady, A. Doud, K. Shades, E. Rogin, B. He, Quadcopter control in three-dimensional space using a noninvasive motor imagery-based braincomputer interface, J. Neural Eng. 10 (2013).
- [5] A.S. Royer, A.J. Doud, M.L. Rose, B. He, EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies, Neural Systems and Rehabilitation Engineering, IEEE Trans. 18 (6) (2010) 581–589.
 [6] A.J. Doud, J.P. Lucas, M.T. Pisansky, B. He, Continuous three-dimensional
- [6] A.J. Doud, J.P. Lucas, M.T. Pisansky, B. He, Continuous three-dimensional control of a virtual helicopter using a motor imagery based brain-computer interface, PIOS One 6 (10) (2011) e26322.
- [7] Y. Chae, J. Jeong, S. Jo, Toward brain-actuated humanoid robots: asynchronous direct-control using an EEG-based BCI, IEEE Trans. Robot. 28 (5) (2012) 1131–1144.
- [8] B.Z. Allison, R. Leeb, C. Brunner, G.R. Müller-Putz, G. Bauernfeind, J.W. Kelly, C. Neuper, Toward smarter BCIs: extending BCIs through hybridization and intelligent control, J. Neural Eng. 9 (1) (2012) 1–7.
- [9] B. Choi, B.S. Jo, A low-cost EEG system-based hybrid brain-computer interface for humanoid robot navigation and recognition, PLOS One 8 (9) (2013).
- [10] T.O. Zander, M. Gaertner, C. Kothe, R. Vilimek, Combining eye gaze input with a brain-computer interface for touchless human-computer interaction, Int. J. Hum.-Comput. Interact. 27 (1) (2010) 38–51.
- [11] M. Kim, S. Jo, Hybrid EEG and eye movement interface to multi-directional target selection, in: Proceedings of the IEEE EMBS, (2013) 763–766.
- [12] A. Sanna, F. Lamberti, G. Paravati, E.A. Henao Ramirez, F. Manuri, A kinectbased natural interface quadcoptor control, Enterain. Comput. 4 (3) (2013) 179–186.
- [13] P. Brunner, L. Bianchi, C. Guger, F. Cincotti, G. Schalk, Current trends in hardware and software for brain-computer interfaces (BCIs), J. Neural Eng. 8 (2011) 025001.
- [14] D. Li, J. Babcock, D.J. Parkhurst, OpenEyes: a low-cost head-mounted eye-tracking solution, in: Proceedings of the Symposium on Eye Tracking Research & Application (ETRA), (2006) 95–100.
- [15] D. Borghetti, A. Bruni, M. Fabbrini, L. Murri, F. Sartucci, A low-cost interface for control of computer functions by means of eye movements, Comput. Biol. Med. 37 (2007) 1765–1770.

- [16] A.T. Campbell, et al., NeuroPhone: brain-mobile phone interface using a wireless EEG headset, in: Proceedings of the ACM SIGCOM Workshop on Networking, Systems, & Applications on Mobile Handhelds, (2010) 3–8.
- [17] P. Bobrov, et al., Brain-computer interface based on generation of visual images, PLOS One 6 (6) (2011) e20674.
- [18] C.H. Morimoto, D. Koons, A. Amit, M. Flickner, S. Zhai, Keeping an eye for HCI, in: Proceedings of the XII Brazilian Symposium on Computer Graphics and Image Processing, 1999171–176.
- [19] H. Ramoser, J. Muller-Gerking, G. Pfurtscheller, Optimal spatial filtering of single trial EEG during imagined hand movement, IEEE Trans. Rehabil. Eng. 8 (4) (2000) 441–446.
- [20] E.R. Kandel, J.H. Schwartz, T.M. Jessell, Principles of Neural Science, 4th Ed., McGraw-Hill, New York, US (2001) 916.
- [21] D. Purves, et al., Neuroscience, 3rd Ed., Sinauer Associates, Inc., Sunderland, Massachusetts, USA (2004) 668–670.
- [22] W.W. Abbott, A.A. Faisal, Ultra-low-cost 3D gaze estimation: an intuitive high information throughput compliment to direct brain-machine interfaces, J. Neural Eng. 9 (4) (2012) 046016.
- [23] D.J. McFarland, W.A. Sarnacki, T.M. Vaughan, J.R. Wolpaw, Brain-computer interface (BCI) operation: signal and noise during early training sessions, Clin. Neurophysiol. 116 (1) (2005) 56–62.
- [24] D.J. Hand, R.J. Till, A simple generalisation of the area under the ROC curve for multiple class classification problems, Mach. Learn. 45 (2001) 171–186.
- [25] R. Poli, C. Cinel, A. Matran-Fernandez, F. Sepulveda, A. Stoica, Towards cooperative brain-computer interfaces for space navigation. in: Proceedings of the 2013 International Conference on Intelligent User Interfaces. ACM, 2013.

- [26] B. Graimann, B. Allison, G. Pfurtscheller (Eds.), Brain–Computer Interfaces: Revolutionizing Human–Computer Interaction, Springer, Heidelberg, Germany, 2010.
- [27] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, G. Pfurtscheller, How many people are able to operate an EEG-based brain-computer interface (BCI)? Neural Syst. Rehabil. Eng. IEEE Trans. 11 (2) (2003) 145–147.
- [28] G. Pfurtscheller, C. Neuper, Motor imagery and direct brain-computer communication, Proc. IEEE 89 (7) (2001) 1123–1134.
- [29] A. Bashashati, M. Fatourechi, R.K. Ward, G.E. Birch, A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals, J. Neural Eng. 4 (2) (2007) R32.
- [30] H.J. Hwang, K. Kwon, C.H. Im, Neurofeedback-based motor imagery training for brain-computer interface (BCI), J. Neurosci. Methods 179.1 (2009) 150–156.
- [31] C. Neuper, A. Schlogl, G. Pfurtscheller, Enhancement of left-right sensorimotor EEG differences during feedback-regulated motor imagery, J. Clin. Neurophysiol. 16 (4) (1999) 373–382.
- [32] J. Park, K.E. Kim, S. Jo, A POMDP approach to P300-based brain-computer interfaces, in: Proceedings of the International Conference on Intelligent User Interfaces (IUI), (2010) 1–10.
- [33] Y. Chae, S. Jo, J. Jeong, Brain-actuated humanoid robot navigation control using asynchronous brain-computer interface, in: Proceedings of the International IEEE EMBS Conference on Neural Engineering, (2011) 519–524.
- [34] G.R. Müller-Putz, R. Scherer, C. Brauneis, G. Pfurtscheller, Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components, J. Neural Eng. 2 (4) (2004).