SICE Annual Conference 2015
July 28-30, 2015, Hangzhou, China

Hybrid-BCI Smart Glasses for Controlling Electrical Devices

YONGWON Kim [1], NETIWIT Kaongoen [1], SUNGHO Jo [1]

1. Intelligent Systems & Neurobotics Lab, Department of Computer Science, KAIST, Daejeon, Korea, Republic of, 305-701
E-mail: shjo@cs.kaist.ac.kr

Abstract: We are surrounded by diverse objects such as televisions, light switches, and computers that we must interact with in our everyday life. Often, this interaction should be done simultaneously with other work. To accomplish this without restricting the user's action, we propose a wearable glasses-type device that enables this interaction using eye movements and brain activity. The wearable design minimizes the behavioral restrictions in a daily life scenario. Our device uses a natural and intuitive method to determine the user's interest by tracking his gaze, and triggering operations by analyzing his thinking and natural response to visual stimulation using binary classification and steady state visually evoked potential (SSVEP) of Brain-Computer Interface (BCI) technique, respectively. The combination of these features enables our device to provide hands-free and remote interaction between a human and an object (in this specific case, a television). Participants in our user evaluation reported that our device and its application strategy were intuitive, natural, and easy to understand when interacting with objects. The average performance time for the system to switch on the interest object (television) was 4.43 ± 9.93 [mean (s) ± SD] seconds and the overall average accuracy for remote control function was 70.83%.

Key Words: Wearable computing, Smart glasses, Gaze tracking, Brain-computer interface, Electroencephalography (EEG)

1 Introduction

In our daily life, we interact with many objects such as TVs, air conditioners, and lights. In recent years, the Internet of Things (IoT) has huge interest as a method of interacting with these objects. In a broad sense, IoT refers to enabling communications between all kinds of objects through internet [5]. It can help the interaction between humans and objects by providing a direct path to convey the user's intention. However, because the human cannot connect to the internet directly, an end device is needed. Designing such an end device is a challenging issue in human-computer interface areas

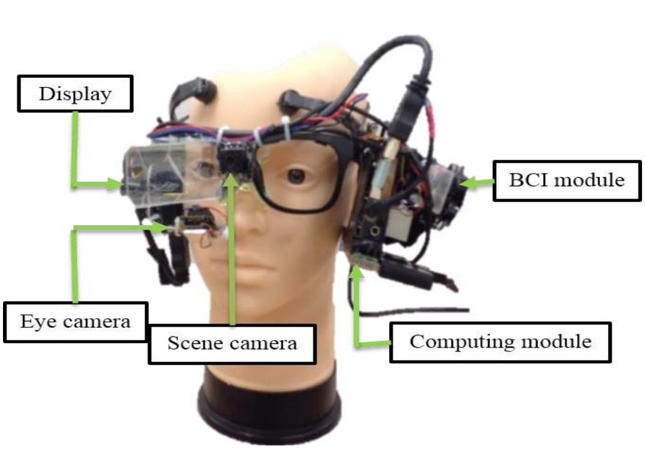
It seems that existing devices such as smartphone and Google Glass perform such a role well. However, all of these devices restrict the actions a user can perform while using them. For example, a smartphone occupies one or both hands and that possibly makes it hard for users to carry a box and manipulate their smartphone at the same time. Similarly, Google Glass requires one hand to operate the touchpad built into the side of the device. These models of interaction for using these devices limits their practical efficiency with respect to their applications to daily life. Therefore, in order to satisfy the requirements for the interaction between humans and objects in daily life, the most important thing is to design a device which does not restricted the user's behavior.

We present a new glasses-type wearable device that enables interaction by simply looking at an object and imagining an operation of interest. The proposed device is designed to be hands-free, combining eye- gaze tracking with a brain-computer interface (BCI). Two cameras built onto our device track the user's' gaze to enable the system to capture wherever the user looks. Furthermore, the device is hybridized with a BCI. A BCI is an interaction technology which measures and translates the brain activity to control the external device. In the proposed system, it allows a user to interact with his object of interest. In this paper, we discuss the design decisions regarding the proposed device and describe a suitable application scenario. In addition, we examine the viability of the proposed device and present our conclusions.

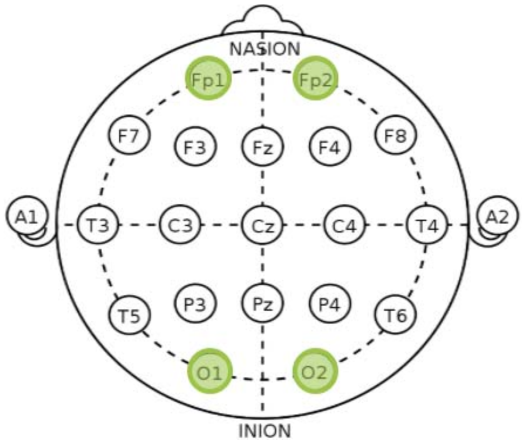
2 Hybridized Eye-Gaze Tracking And BCI

Various systems have been explored to allow users to conduct their tasks without a loss of efficiency. With respect to hands-free operation, eye-gaze tracking technologies are quite advanced [10, 11], since they directly detect the user's attention/interest in a target. Furthermore, they do not require intensive training. Despite these advantages, eye-gaze tracking systems are limited with regard to the selection operation. In order to make a selection, eye-tracking solutions rely on the amount of time the eyes remain on an object, or how frequently the eyes blink, and try to find the optimal time. Due to the complexity of various tasks, it is almost impossible to find optimally fixed dwell time to cover every conditions [12].

One possible solution is hybridization with another modality. In many combinations, we focus on hybridization with a BCI. Due to its direct communication pathway between the brain and a device, it is cited as one of the most intuitive and natural interfaces [6]. Therefore, a properly designed BCI protocol, such as binary classification, has been suggested for the selection operation for eye-gaze tracking [1, 2, and 12]. However, previous studies only considered its use with regard to control preselected object. These studies did not allow the user to select the object to interact with. In this study, we proposed the combination of eye-gaze tracking and BCI binary classification to build an interaction line between user and object of interest in 3 dimensional (3D) space by looking at the object and focusing user`s attention. While using the combination of eye-gaze tacking and BCI binary classification to allow the



(a)



(b)

Figure 1. (a) The proposed device components and (b) the EEG acquisition sensors location.

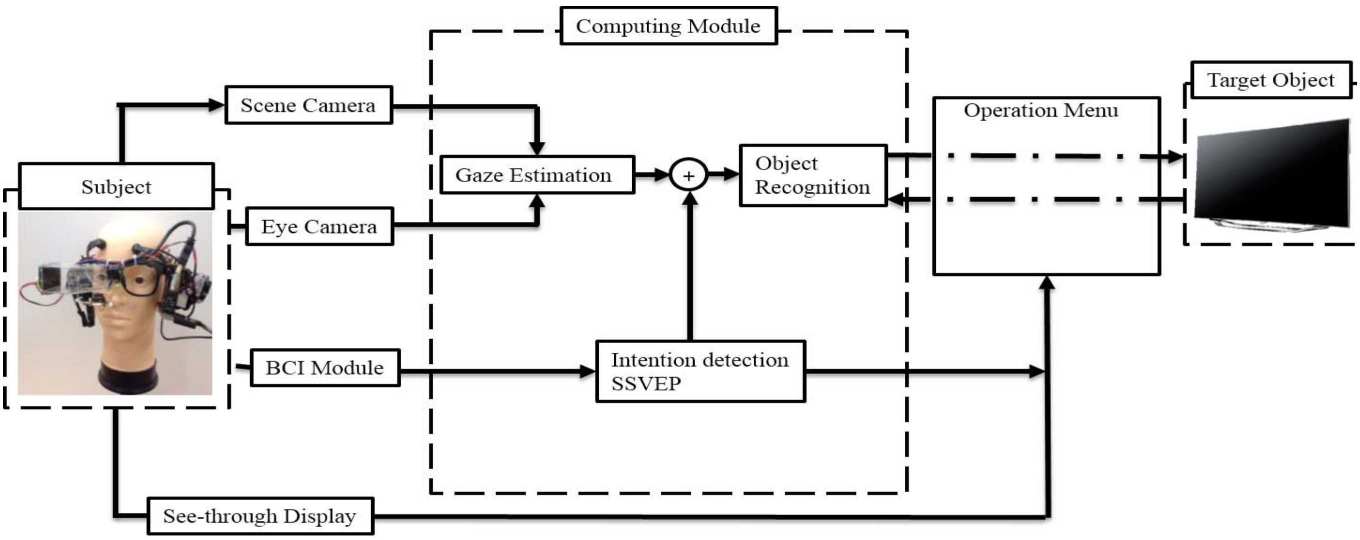


Figure 2. Hybrid interface system overview.

user to select an object to control, steady state visually evoked potential (SSVEP) method are used to control after opening interaction line. In summary, we propose a prototype wearable and hands-free hybrid BCI system which overcomes the limitations and provides a solid basis for daily life application.

3 Method

3.1 The Proposed Device

Figure 1 (a) shows our proposed device. It consists of four modules (display, computing, eye-gaze tracking, BCI). Eye-gaze tracking module has two web cameras for capturing eye images (Eye camera) and forward-facing scene images (Scene camera) respectively. We removed infrared (IR) cut-off filter from eye camera and installed one IR illuminator to enhance the contrast between pupil and iris.

BCI module consists of four electrode channels (Fp1, Fp2, O1, O2) plus CMS/DRL references (see figure 1 (b) for the location) and analog-to-digital converter (ADC). Modified version of Emotive EEG Neuroheadset using wet saline electrodes was used as EEG recording device. It records EEG signals at 128Hz and send the recorded signals to computing module wirelessly.

For computing module, we employed an ODROID-U2 single-board computer (Hard kernel Inc.) with a Samsung Exynos4412 Prime chip (1.7Ghz ARM Cortex-A9 Quad Cores), 2 GB of RAM, and the Linux environment.

3.2 Algorithm

Figure 2 shows our system overview. Once the eye camera and scene camera start to capture sequential eye and scene images, image processing procedure is performed to extract pupil center location. For the robust pupil center tracking, we applied Starburst algorithm [3]. The algorithm first reduce the line noise using Gaussian filter. Then, it chooses a point that represents the best guess of pupil center. As next step, the derivatives along 18 rays which start from the best guess point are calculated against pixel by pixel. If the derivative value exceeds a threshold it stop and considers that pixel as a candidate pupil edge pixel. For the best guess point, previous extracted pupil center are used in each image. After extracting candidate edge points is done, the algorithm applies Random Sample Consensus (RANSAC) paradigm for ellipse fitting [7]. The ellipse that includes the largest number of candidate points considered as estimated pupil. For the scene images, we use SURF (Speeded Up Robust Features) [13] as the feature detector and classify whether the scene being seen now matches with the picture of the target object in the database by using FLANN (Fast Library for Approximate Nearest Neighbors) that is the library containing implementations of the fast approximate nearest neighbor search in high dimensional spaces [14].

In our system, two type of EEG signal analyses are performed. First one is binary classification which discriminate user's mental state: concentration vs non-concentration. Together with eye gazing direction and scene image recognition which make the system knows that the user is currently looking at the target object, if the user is in concentration state, the system will open the network connection between the system and the target object.

For the binary classification, we uses alpha (8 -13Hz) and beta (14-29Hz) band power and the band-power difference as features [9]. The band-power difference calculated following

$$\text{Power}_{diff} = \left[\frac{(P_1 - P_2)}{(P_1 + P_2)}\right]$$

where P_1 is the power in one channel and P_2 is the power in another channel in the same spectral band. Then, an optimal classifier to discriminate EEG signal into two classes was generated based on acquired features by the Support Vector Machines (SVM) algorithm [8] with a linear kernel to the features.

The second BCI module is SSVEP-typed BCI that serves as TV-remote control that control TV after it is turned on. There are three commands that the user can send to the TV: volume down, volume up, and turn off the television. Each button on the menu screen flickers with different frequency: 12 Hz for volume down button, 15 Hz for volume up button and 20 Hz for power button.

Canonical correlation analysis (CCA) method is used to classify the SSVEP signal. CCA is a method that finds a pair of linear combinations, for two sets, such that the correlation between the two canonical variables is maximized [4]. In our case, we use CCA method to find the maximum canonical correlation between our EEG signal and the three references signal defined as the matrix of sine and cosine signal with the frequencies corresponding to the frequencies (and its harmonics) of the stimulus. The frequency that yields the highest value of maximum canonical correlation will be selected as the output command to the target devices.

For both binary and SSVEP classification, the sliding window of 2 s length with 0.25 s increments was applied to EEG signals and three windows were used to perform the classification. To reduce the sensitivity of the "turning off" function, we don't send "turning off" command to the television unless the system give two consecutive results of "turning off" command.

4 Experiments

4.1 Subjects

Four healthy male subjects (age 25 ±2 years) voluntarily participated in our experiment. All of the subjects were of the same laterality (right-handed), were free of any neurological disorders and eyes diseases, and had never experienced any glasses-type wearable device.

4.2 Experimental Setup & Tasks

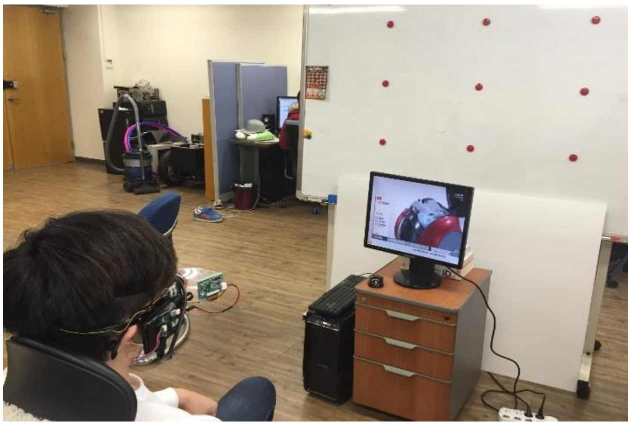


Figure 3. Experimental environment setup

To measure the performance of our devices, we conducted an experiment consisting of four experiment tasks for the subjects to perform. Figure 3 shows our experimental setup. After each subject performed eye-gaze tracking calibration and trained EEG classifier for the binary classification, the participants comfortably sit in a chair facing the TV at distance of 2 m away from it. Once experiment starts the system notify task 1 (turn on the TV) through speaker. Then subject looked at the TV and focused his attention to turn on the TV. The time starting from the point that the task notification was announced to the point that the television turned on was measured as the performance time of the experiment task 1.

The objectives of task 2 to task 4 was to measure the performance for the SSVEP remote control function of our devices. After the TV was turned on, the SSVEP remote-control menu screen appeared on the see-through display of our device. Task 2 (volume TV down for 3 levels from initial) and task 3 (volume TV up for 3 levels from initial level) were assigned to be performed by the subjects. The order of each task was randomly shuffled and each task performed three times. Task 4 (turn off the TV) was given as the last mission to finish the experiment. Once each task was given, the subjects looked at corresponding menu image on the see-through display which blinked at different frequency to perform the task.

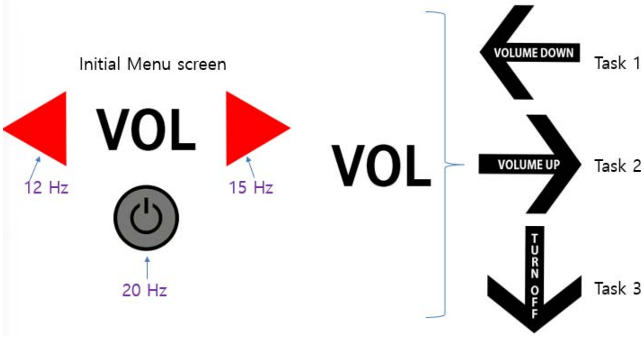


Figure 4: (Left) The initial menu screen for SSVEP remote control function with frequency labelling on each button. (Right) The 'VOL' button change to sign directing the user which button to look while performing experiment

Each task was separated by the initial menu screen for 5 seconds time interval. The accuracy for SSVEP classification and the performance time to finish each task was then measured. It should be noticed that we did not measure the success/failure rate for the experiment. In other words, we did not let the system turn off the television when the subject supposes to perform task 2 and task 3.

4.3 Experimental Results

The system calibration and training results of each subject are shown in Table 1. The eye calibration error is the average distance between the target points and the point estimated by eye-gaze tracking system. For the EEG binary classification accuracy, we reports the result of 10-fold cross-validation. The average eye-gaze tracking calibration error was $0.59 \pm 0.05\ (mean(cm) \pm SD)$ (mean ± standard deviation) and binary classification accuracy was $70.94\ \pm 1.29\ (mean(\%) \pm SD)$

Table 1: Eye calibration error, binary classification accuracy of each subject

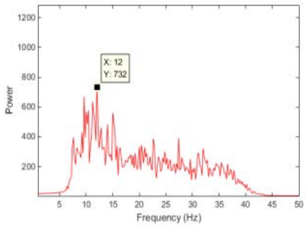
Subjects	Eye calibration Error(SD) (cm)	Binary classification result (SD) (%)
Subject1	0.59(0.44)	71.28(1.55)
Subject2	0.61(0.34)	70.05(3.28)
Subject3	0.54(0.29)	72.62(2.04)
Subject4	0.65(0.28)	69.81(1.64)
Mean	0.60	70.94
SD	0.05	1.29

Table 2 shows performance time for each task and the accuracy for SSVEP classification in details. The average performance times to perform each task were 4.43 ± 9.93 [mean (s) ± SD] for task1, 4.96 ± 1.13 for task 2, 5.75 ± 1.40 for task3, and 3.42 ± 0.55 for task 4. The SSVEP classification accuracies were 73.50 ± 6.86 [mean (%) ± SD], 73.00 ± 6.48, and 66.00 ± 12.30 for stimulus frequencies 12 Hz, 15 Hz, and 20 Hz, respectively.

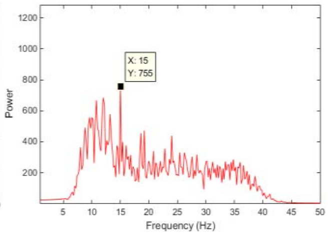
.

Table 2: Performance time and SSVEP classification result

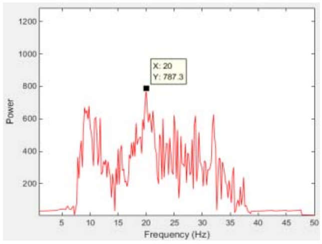
Subjects		Time (s)		Task Type	Classification Accuracy (%)	Time (s)
Subject 1	**Task 1: Turn on TV**	3.62	**Task 2-4: SSVEP TV Remote Control**	**2**	73.00	5.07
				3	76.00	4.07
				4	50.00	3.86
Subject 2		4.79		**2**	64.00	6.28
				3	64.00	7.48
				4	67.00	3.86
Subject 3		5.58		**2**	78.00	3.53
				3	79.00	5.82
				4	67.00	3.24
Subject 4		3.72		**2**	79.00	4.91
				3	73.00	5.62
				4	80.00	2.74



(a)



(b)



(c)

Figure 5: The average raw EEG data obtaining from channel O1 and O2 when the subject perform SSVEP experiment. We can see that there is a peak in the frequency corresponding to the frequency of the stimulus (a) 12 Hz (b) 15 Hz and (c) 20 Hz.

5 Discussion

There are several topics needed to be discussion about this project. The first thing is about the design of device. Our device use many component and thus resulting in the size and weight that is not appropriate for the user to use it in daily life. However this could be improved by using better and more expensive components that could give the same performance but smaller in size and weight. Since each users have different size of head, the design that can be changed in size also recommended for the improvement of the design.

Second thing that need to be mentioned here is that we performed the experiment in our lab-environment condition where the subject needs to sit still with no movement, and there is no other stimulus that can distract the subject to perform the experiment. However, in real life, users may want to move themselves around while using the device. This could cause a lot of noises and affect the accuracy of the device. To solve this problem, accelerometer and gyro meter might help to cancel the noise generated from the motion of the user.

The SSVEP signal using in this design is also needed to be concerned. Generally, SSVEP signal are gathered from the signal generated by both of two eyes watching at the stimulus with specific frequency. Nevertheless, our design show the stimulus to the user only on the right eye. This may cause the signal to be weaker comparing to the traditional way and left eye might detect other stimulus in the environment that have different frequency and affect with the SSVEP signal. However, the use of EEG from both frontal cortex (Fp1, Fp2) and occipital lobe (O1, O2) allows us to combine two BCI techniques to use in one system hence providing more interfacing function and make the interface system more flexible.

6 Conclusion

In this paper, we presented a novel glasses-type wearable device and its application strategy. The proposed device provided the user with a point-and-select ability through the hybridization of eye-gaze tracking and a brain-computer interface. During experiments, participants successfully manipulated a TV through our device. Our work is a step towards making the wearable device a more practical possibility for interacting with objects in daily life.

References

[1] A., Frisoli, C., Loconsole, D., Leonardis, F., Banno, M., Barsotti, C., Chisari, and M. Bergamasco, "A new gaze-bci-driven control of an upper limb exoskeleton for rehabilitation in real-world tasks.," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 42, 6, pp. 1169–1179. 2012

[2] B. H., Kim, M., Kim, and S. Jo, "Quadcopter flight control using a low-cost hybrid interface with eeg-based classification and eye tracking.," Computers in biology and medicine , 2014.

[3] D. Li, D. Winfield, and D. Parkhurst, "Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches," IEEE CVPR Workshop Vision for Human-Computer Interaction (V4HCI), pp. 79–79, Jun. 2005

[4] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao. "An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method." Journal of neural engineering 6, no. 4: 046002.2009

[5] K., O'Hara, A., Sellen, and R. Harper, "Embodiment in brain-computer interaction." In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, pp. 353–362. 2011

[6] H., Sundmaeker, P., Guillemin, P., Friess, and S., Woelffle, Vision and challenges for realising the Internet of Things., Eur-Op, 2010

[7] M. Fischler and R. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," Communications of the ACM, vol. 24, no. 6, pp. 381–395, 1981

[8] O., Chapelle, "Training a Support Vector Machine in the Primal", Neural Computation., Vol.19, No. 5, pp.1155-1178, 2007

[9] P., Ramaswamy. "Brain computer interface design using band powers extracted during mental tasks." In Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on, pp. 321-324. IEEE, 2005.

[10] R. J. Jacob and K. S. Karn, "Eye tracking in human-computer interaction and usability research: Ready to deliver the promises.," Mind 2, 3, 4, 2003.

[11] R. Vertegaal, "A fitts law comparison of eye tracking and manual input in the selection of visual targets.," In Proceedings of the 10th international conference on Multimodal interfaces, ACM, pp. 241–248..2008

[12] T. O., Zander, M., Gaertner, C., Kothe, and R. Vilimek, "Combining eye gaze input with a brain–computer interface for touchless human–computer interaction." Intl. Journal of Human–Computer Interaction 27, 1, pp. 38–51. 2010

[13] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: Speeded Up Robust Features" Katholieke Universiteit Leuven

[14] "FLANN - Fast Library for Approximate Nearest Neighbors." FLANN - Fast Library for Approximate Nearest Neighbors : FLANN - FLANN Browse. N.p., n.d. Web. Fall 2014. <http://www.cs.ubc.ca/research/flann/>.