2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

EOG/ERP Hybrid Human-Machine Interface for Robot Control

Jiaxin Ma[1], Yu Zhang[2], Yunjun Nam[3], Andrzej Cichocki[4], and Fumitoshi Matsuno[5]

Abstract— **Electrooculogram (EOG) signals are potential responses generated by eye movements, and event related potential (ERP) is a special electroencephalogram (EEG) pattern which evoked by external stimuli. Both EOG and ERP have been used separately for implementing human-machine interfaces which can assist disabled patients in performing daily tasks. In this paper, we present a novel EOG/ERP hybrid human-machine interface which integrates the traditional EOG and ERP interfaces together. Eye movements like the blink, wink, gaze, and frown are detected from EOG signals using double threshold algorithm. Multiple ERP components, i.e., N170, VPP and P300 are evoked by inverted face stimuli and classified by linear discriminant analysis (LDA). Based on this hybrid interface, we also design a control scheme for the humanoid robot NAO (Aldebaran robotics, Inc). On-line experiment results show that the proposed hybrid interface can effectively control the robot's basic movements and order it to make various behaviors. While normally operating the robot by hands takes 49.1 s to complete the experiment sessions, using the proposed EOG/ERP interface, the subject is able to finish the sessions in 54.1 s.**

I. Introduction

Brain-machine interface (BMI), also called brain computer interface (BCI), can translate brain signals into control signals without using muscles [1]. It is mainly designed for assisting people with severe motor disabilities, helping them re-establish communicative and environmental control abilities [2]. It may also apply to able-bodied people in some special situations where the other means of communication are unavailable or occupied. Among a variety of noninvasive BMI methods, electroencephalogram (EEG) method has high time resolution, less environmental limits, and requires inexpensive equipment [3]. Although nowadays there are many researches about EEG-based BMI, they are more theoretical than practical. One reason is because its information transfer rate (ITR) is usually very low, and thus the response time of system is unsatisfactory for most daily tasks. Moreover, the patients may have various kinds of needs, like using wheelchair to move around, and requesting food/water/etc. from carers. Usually a BMI system can only manage one certain kind of task. It is rather difficult to have a universally robust BMI system dealing with different kinds of situations. Against these disadvantages of BMI, we propose a new hybrid human-machine interface which combines EOG (using eye movements) and EEG (using ERPs).

EOG is the electrical potential response generated by eye movements. For EOG-based system, the response speed can be very high. It is a highly desirable feature especially for motion control, and there are a lot of studies about designing EOG human-machine interface [4] [5] [6]. However, the number of usable eye-movement patterns is usually quite limited. That makes the system can only support a small number of outputs, which does not satisfy a multi-task situation. In order to realize large numbers of outputs, people tried to use a menu-like interface, and let the user select items by continuous eye movements. However, it is hardly an optional choice since it is both time consuming and more importantly, fatiguing, to repeatedly perform eye movements.

ERPs are brain activities time-locked to experimental events of interest and contain typically N170, vertex positive potential (VPP), N200 and P300 etc.. P300-based BMI has relatively robust performance for target detection. One of its representative applications is P300 speller which can be used for inputting characters [7]. ERP interface can naturally support large numbers of outputs. In the scenario of "choosing a command from the list", using ERP the user is only required to focus on the desired target. This process needs no effort but only slight mental concentration, which is much more convenient than performing eye movements repeatedly. In the respect of processing speed, ERP interface may be even faster than the "cursor moving" approach if the number of commands is large and the accuracy is high enough. However, for frequently used commands, the effectiveness of ERP interface will be unsatisfactory. Even though the ITR of ERP interface is relatively high among EEG-based BMIs, it is still not comparable to EOG/EMG-based system.

The idea of hybrid BMI system is a relatively new topic. In 2011, Postelnicu *et al.* adopted eye saccade to control a robot arm to move towards 4 directions, and alpha band EEG (evoked by eye closing) to make the robot arm grip [8]. It is an interesting idea to add alpha band EEG component to EOG-based control. However in that experiment, the change of alpha band is also originated from eye movement, which shows no concrete advantage of combining EEG and EOG system.

In practice, a lot of control scenarios require multitask. These tasks have different characteristics that some are suitable to be realized by EOG and others by ERP, as mentioned

[1]J. Ma is with the Graduate School of Engineering, Kyoto University, Japan `ma.jiaxin.62c at st.kyoto-u.ac.jp`

[2]Y. Zhang is with the School of Information Science and Engineering, East China University of Science and Technology, China, and the RIKEN Brain Science Institute, Japan `zhangyu0112 at gmail.com`

[3]Y. Nam is with the School of Interdisciplinary Bioscience and Bioengineering, Pohang University of Science and Technology, Korea `druid at postech.ac.kr`

[4]A. Cichocki is with the RIKEN Brain Science Institute, Japan, and the Systems Research Institute, Polish Academy of Science, Poland `cia at brain.riken.jp`

[5]F. Matsuno is with the Graduate School of Engineering, Kyoto University, Japan `matsuno at me.kyoto-u.ac.jp`

978-1-4673-6357-0/13/$31.00 ©2013 IEEE

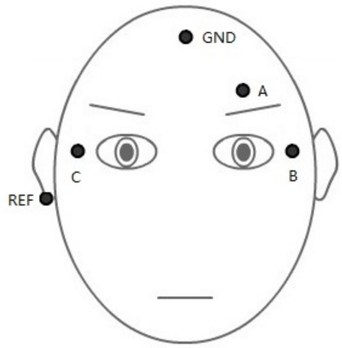

Fig. 1. The placement of EOG electrodes. A: vertical EOG electrode; B, C: horizontal EOG electrode; REF: reference; GND: ground.

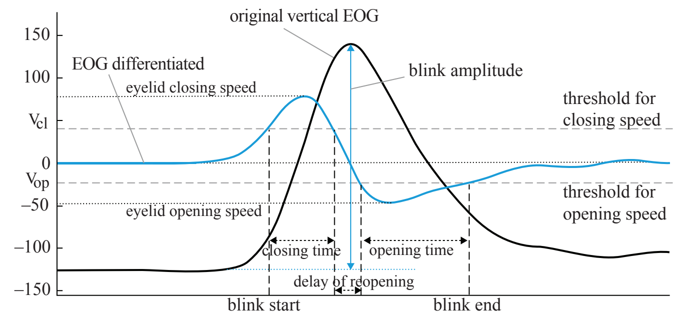

Fig. 2. The waveform of a blink in vertical EOG and its differentiated shape. The x-axis is time and y-axis is EOG (μV).

above. Combining EOG and ERP can potentially make full use of the advantages of both systems and avoid the disadvantages. In additional, user experience like convenience is also very important. For using EOG, frequent eye movements will easily accumulate fatigue on muscles. Although ERP casts no heavy burden on the user, continuously watching the flashing signs on the screen inevitably leads to impatience and weariness, and thus decrease the system performance. Using two control methods alternately can potentially relief user's burden both on body and mind.

In this study, we take robot control as the application of our proposed system. This is because our experiment robot (NAO) can perform different tasks, from simple movements to complex behaviors, which is good for showing the advantages of the proposed system. By EOG we adopt various eye movements to control the moving of robot with fast response. The eye movements include blinks (double and triple), winks (left and right), gazes (left and right), and frowns. Specified algorithms have been designed for detecting those different eye movements. By eye movement it also realizes the switching between EOG mode and ERP mode. ERP mode allows the user to select different behaviors from an item list. A novel multi-ERP paradigm based on stimuli of inverted faces is adopted, which has shown significantly higher ERP classification and ITR than the traditional P300-based BMI [9].

II. EOG ANALYSIS

A. EOG acquisition

In this study, we recognize various eye movements from EOG signals to implement the control scheme of the robot. The EOG electrodes are placed as shown in Fig. 1 where position A is for vertical EOG; position B and C are for horizontal EOG. Other two electrodes are ground and reference which are shared with EEG electrodes. The signals are recorded by g.USBamp with g.GAMMAbox (g.tec medical engineering GmbH, Austria).

The sampling rate of EOG is 32 Hz (down-sampled from 256 Hz input), since high sampling rate causes more vibration in the signal which is undesirable for our algorithm. The filter is chosen as 0.1-30 Hz (built-in). We have tested using 0.1-30 Hz filter, 0.01-30 Hz filter and no filter. Without high pass filter, the baseline drift effect is extremely obvious in the signal. The 0.1 Hz high pass filter can best suppress the effect of baseline drift so we choose that. The 30 Hz low pass filter can remove high frequency noise so it is also applied. In fact, our algorithm is not sensitive to baseline drift so it also works without filters. The reason of our choice is mainly due to that the drifting signal is inconvenient to observe in the scope.

B. Eye movement detection

Among various eye movements, blinks, winks, gazes, and frowns are chosen to be the control commands of the robot. We designed algorithms to detect these eye movements. Here the algorithm of blink detection is explained as an example.

The blink detection algorithm is an improved version of the double thresholds method which was also used in our previous study [10]. Its key idea is to set two thresholds which represent the eyelid's closing speed and opening speed, respectively. The two thresholds can locate blink waveforms in EOG. This method is applied on the vertical EOG signal after taking difference. Because it utilizes speed information but not amplitude information, it has resistance to baseline drift effect.

Figure 2 is a typical EOG waveform of blink, where the black curve is the vertical EOG, and the blue curve is the EOG after taking difference. When the black curve reaches top the blue curve is just 0, which means its speed is equal to 0 when the eyelid is completely closed. V_{cl} and V_{op} are the two thresholds, where the former is for eyelid closing, and the latter for eyelid opening.

The algorithm consists of the following steps:

Step 1: Calculate the first difference of the signal. The result is the approximate speed of the eyelid.

Step 2: Apply the thresholds V_{cl} and V_{op} to the signal. A blink event should consist of four successive points where the first two are equal to V_{cl} and the last two are equal to V_{op} (may be approximate to but not equal to since the data is discrete). All matched point series will be taken as blink candidates.

Step 3: For more accurately estimating the duration of blinks, the start point and the end point should be properly extended from $\{V_{cl}, V_{op}\}$ to $\{E_{cl}, E_{op}\}$.

Step 4: Combine the blink candidates which are very closed to each other. In some cases two adjacent blink events are so close that it is better to treat them as one blink. If the distance from the end point of one blink to the start point of another is not larger than a value $combTH$ (set to 1/32 s), we will combine them into a single blink.

Step 5: Verify the duration and amplitude of blink candidates. In most cases blinks should be shorter than 0.5 s, so those longer than 0.5 s are removed. The amplitude is calculated by $(A_{max} - A_{head} + A_{max} - A_{tail})/2$, where A_{max} is the maximal amplitude during a blink; A_{head} and A_{tail} are the amplitudes at the start point and the end point of the blink. If the calculated amplitude is smaller than a set value A_{min}, the blink is removed.

A calibration process is to determine the value of parameters (V_{cl}, V_{op}, and A_{min}) before on-line experiments. In the process, the subject is asked to do 10 times of triple blink (30 blinks in total). The system first use preset parameters to detect those blinks, then update each parameter according to the calibration result. Assume that the maximal speed during each eye closing is V^i_{clmax}, and the minimal speed during each eye opening is V^i_{opmin}, then

$$\begin{aligned} V_{cl} &= 0.5 \min_{i=1\ldots30} V^i_{clmax} \\ V_{op} &= 0.5 \max_{i=1\ldots30} V^i_{opmin} \\ A_{min} &= 0.8 \min_{i=1\ldots30} A^i_{max} \end{aligned} \tag{1}$$

For detecting winks, gazes, and frowns the algorithms are similar, but parameters need to be adjusted due to the different features of eye movements.

III. ERP ANALYSIS

A. EEG acquisition

In this study, 8 electrodes are used to record EEG, which are Fz, Cz, P7, P3, Pz, P4, P8, and Oz (Fig. 3). The device is the same as EOG acquisition. The electrodes of ground and reference are shared with EOG electrodes (placed in forehead and ear lobe). These positions cover the areas where N170, VPP, and P300 occur.

The sampling rate of EEG is 64 Hz (down-sampled from 256 Hz input). The filter is the same as that used in EOG. The data collection, stimuli presentation and online processing are controlled by Simulink and Matlab (Mathworks Inc., USA).

B. ERP paradigm

This study adopts a more advanced ERP paradigm based on configural processing of human face, which combines oddball presentation and inverted face perception [9]. This paradigm mainly exploits three ERP components, namely VPP, N170 and P300, instead of only P300, and hence significantly improves the target detection performance in contrast to the stimulus intensification pattern used in the conventional P300-based BCI.

Among the ERPs, N170 and VPP are evoked by the configural processing of facial image, and P300 is evoked by oddball event. The shapes of their waveforms are illustrated in Fig. 4.

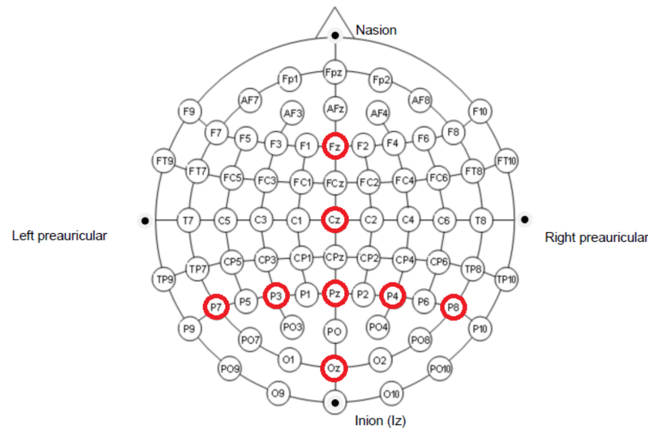

Fig. 3. Electrode positions for EEG recording. (Ground and reference electrodes are showed in Fig. 1.)

Our ERP interface has 8 items placed in 8 directions of screen (N, W, S, E, NW, NE, SW, SE). The training phase contains eight runs. Each run which consists of five trials is to train one of the 8 direction. The test phase also contains eight runs but each run consisting only two trials (see Fig. 5 for detailed timing of one run). Feedback is provided to the subject in each run of the test phase.

C. ERP classification

Linear Discriminant Analysis (LDA) is used to classify which target the subject is focusing on. Before the on-line experiment, the subject needs to train the LDA classifier first.

For the input of LDA, each data segment has a length of 700 ms which is calculated from the beginning of each stimulus. A total of 320 such data segments consisting of 40 targets and 280 non-targets were derived from each subject for classifier training. To reduce the length of input, the data segment is further down-sampled to 15 points (approximately 21 Hz). So the length of input is $8 \times 15 = 120$ (for 8-channel data), which means the feature vector has 120 dimensions.

To conduct the on-line experiment, the accuracy of test phase should achieve at leat 50%, or the subject need to redo the training.

IV. ROBOT CONTROL AND EXPERIMENT

A. Robot control strategy

In our proposed system, EOG mode and ERP mode can be switched over at any time, so the hybrid system can exert the advantages of both. In ERP mode, one output is generated after 16 stimuli is given (8 items, 2 trials) where each stimulus consists of 100 ms highlight and 100 ms dark, so the total time to generate an output is at least 3.2 s, whereas using eye movement to generate a command usually takes less than 1 s. Considering this fact, when we design the robot control policies we follow this principle: commands which require fast response and high accuracy should be assigned to EOG mode; commands which are infrequently

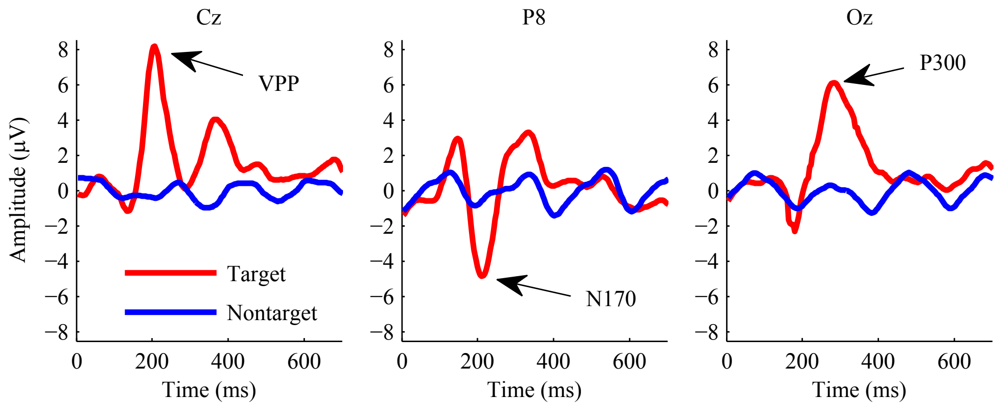

Fig. 4. Grand average ERP waveforms derived from the target and non-target stimuli of the inverted face image. VPP, N170, and P300 can be clearly observed from channel Cz, P8, and Oz.

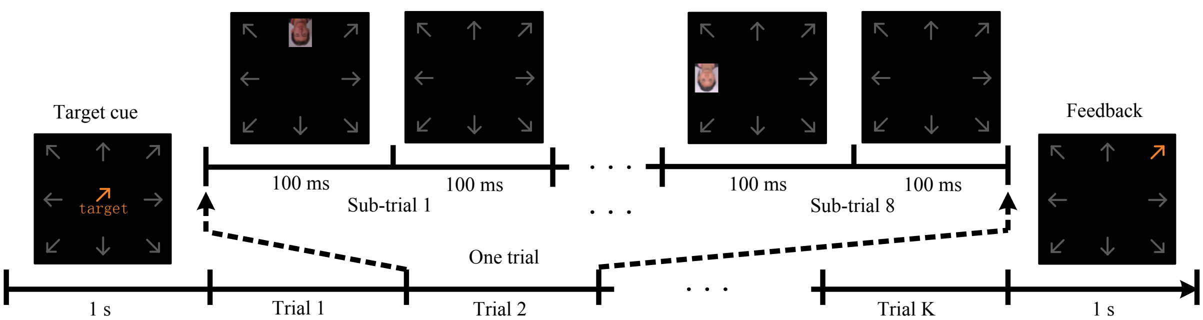

Fig. 5. The timing of one run. In the training phase, the number of trials K was 5 whereby each run consisted of 40 flash sub-trials (5 targets and 35 non-targets), and no feedback was provided to the subject. In the test phase, the number of trials K was 2 and feedback was provided. Each trial consisted of eight sub-trials in each of which one stimulus was randomly presented in one of the eight directions for 100 ms with an inter-stimulus interval of 100 ms.

used and does not require fast response are are assigned to ERP mode. In additional, the left/right eye movements should respectively correspond to left/right movement of robots for convenience.

Generally, The EOG mode controls the robot's moving as well as mode switching, and the ERP mode controls the robot's complex behavior. The control policies are described in detail as follows.

1) EOG mode: In EOG mode, the system is asynchronous, which means when the user performs an eye movement, the robot will respond at once. The timing of sending command is held by the user. If the user does not send any command, the robot should make no action.

- blink related:
  - double blink: robot stops the current behavior.
  - triple blink: robot goes ahead.

For recognizing, double blink should be performed within an 1 s interval, and triple blink should be performed in 1.5 s. Generally, during a triple blink the system will first detects a double blink, so the output will be like "stop" then "go ahead". This does not matter because in most cases we would like to make the robot stop first before the next move.

- wink related:
  - left wink: robot stops and turns 90° to the left.
  - right wink: robot stops and turns 90° to the right.

The robot will rotate by its feet. In order to turn a smaller angle, the user can stop the robot any time during the rotation by double blink. This is a good example showing the fast response of EOG.

- gaze related:
  - look left: robot turns its head to left if it is at center, to center if it is at right.
  - look right: robot turns its head to right if it is at center, to center if it is at left.

These commands are designed for the remote-control scenario where the user observes through the head-mounted camera of the robot. Because turning the whole body of NAO is very time consuming but turning head is quick, these commands are helpful for adjusting the robot's view.

- frown related:
  - frown: robot stops and enters ERP mode.

Every time the user switches to ERP mode, a new ERP trial starts after a short preparation time (1.5 s).

2) ERP mode: In ERP mode, the system is synchronous, which means the user have to follow the system's predefined pace. That is to say, in this mode the system cannot estimate whether the user is receiving ERP stimuli or not. Even the user does not look at the screen, the system will still

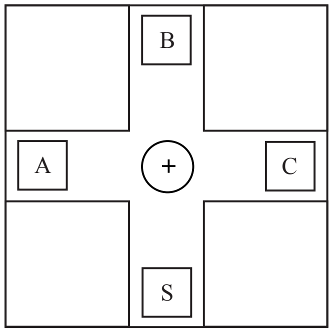

Fig. 6. Experimental room.

analyze the EEG and output the result arbitrarily. So when the user is idle, he should switch to EOG mode to stop the stimuli and outputs. Asynchronous ERP, as well as other asynchronous BMI systems, is still underdeveloped. But since our system can avoid the disadvantage of synchronous BMI by mode switch, we do not need to forcibly implement an asynchronous ERP system at the cost of performance decreasing.

Because the NAO is programmable thus can perform infinitely possible behaviors, it is rather free to assign any kind of behaviors to ERP mode. We can flexibly set it up to conform to different practical scenarios. As our ERP paradigm has 8 outputs, we appropriately assigned 8 behaviors as an example, which are: receiving object, handing over object, dancing, sitting down, standing up, taking a picture, ask for water, and ask for help.

The system does not further receive any ERP command if there is already one being executed. Also, in order not to influence the ERP mode, the system does not receive EOG commands except “frown” (mode switching). This command has another usage: when the ERP produces a wrong output (this happens because its accuracy can hardly achieve 100%), frown twice make the mode switch to EOG then switch back, which stops the current behavior and restart a new ERP trial (and the robot will also return to the standard pose).

Depending on different needs, we can also make more EOG commands available in ERP mode and assign new contents to them.

B. Real-time experiment

In the on-line experiment of this study, we design a scenario in which the user controls the robot to complete a complex task by EOG/ERP. The scenario is described as follows (see Fig. 6 for the setup of experiment room).

- The robot starts from point S, and moves to point A by provided route, then receives an object from the person at point A.
- Holding the object, the robot moves to point B, and gives the object to the person at point B.
- Then the robot moves to point C and performs a dance.
- Finally the robot moves back to point S and sits down.

This scenario can prove NAO, the multi-functional humanoid robot (Fig. 7) is competent for complex tasks, and can also show the usefulness of our hybrid HMI system.

Fig. 7. Humanoid robot NAO.

In the designed scenario, all the tasks are completed in a single session, but this makes the session last too long. Moreover, the execution time of each behavior of robot (receiving, giving, dancing, sitting) is different, which makes inconvenient to calculate the average time cost. In order to simplify the experiment, we divided the experiment into 4 sessions. Each session includes similar tasks and costs similar time.

Take the first session as an example, the detailed experiment steps are:

- The robot gets ready at point S (standing, in EOG mode).
- The robot moves ahead, until gets to the center point (+), then stops.
- The robot looks left, then looks back to center.
- The robot turns 90° left, then moves ahead.
- The robot stops at point A, then switches to ERP mode.
- The robot performs a provided behavior by ERP command.

If the behavior is correctly selected by ERP, the session ends up at once (no need to wait for the robot to complete the behavior). If not correct, the user need to cancel the command and redo the ERP trial.

For the second session, the robot starts from point A and moves to point B in the same manner (the length of route is also the same), and perform another ERP-mode behavior. The third and fourth session are done in the same way.

We have tested four subjects. Their performance is shown in Table I.

V. Discussion and Conclusion

From Table I we can see that the average completion time of subjects. The ideal completion time which operated by hands is 49.1 s (averaged by four runs). The 4th subject gives an outstanding result. His average completion time is 54.1 s, only 5 s longer than the ideal time. Except for his first session (in which he was unaccustomed to the experiment), his session 2, 3, and 4 are completed quite perfectly. This

TABLE I
THE TAKEN TIME (S) FOR THE ON-LINE EXPERIMENT

Subject	Session	No. of commands	Time cost	Time cost Avg.
1	1	16	62.9	70.2
	2	11	57.0	
	3	13	58.3	
	4	13	102.6	
2	1	12	57.2	72.6
	2	11	52.0	
	3	14	59.1	
	4	25	122.2	
3	1	18	78.3	67.5
	2	12	56.2	
	3	15	78.2	
	4	11	57.3	
4	1	17	62.7	54.1
	2	11	52.2	
	3	11	52.5	
	4	11	49.0	

result shows that our hybrid HMI can be potentially applied to able-bodied people due to its good effectiveness.

According to the feedback of the subjects, our hybrid HMI is easy to use. The first advantage is the fast response speed of EOG commands. Using eye movements subjects can send moving commands to the robot quickly and accurately. Moreover, because each eye movement corresponds to one certain movement of robot, it is a more direct way than to select commands from a menu, and thus can achieve better effectiveness. Secondly, the accuracy of ERP is another significant factor. If the ERP trials fail frequently, the user experience will be greatly harmed, and the HMI will not work as well. Because the ERP paradigm we used adopts inverted face stimuli which evoke N170 and VPP besides the conventional P300, the overall accuracy of system is increased.

Except for the multi-functional robot, this hybrid HMI can be also used in any controlling case which requires both moving and behaving. For example, traditional BMI based wheelchair is relatively weak in mobility. Using the proposed system, by EOG it can dexterously control the moving and by ERP it can realize the user's other demands (e.g. requiring water). The control strategy we designed is suitable for both direct control and remote control. In our experiment, the result shows a good overall performance which indicates that our hybrid HMI is not only applicable to the disabled people like traditional BMI is, but may also satisfy the healthy people.

REFERENCES

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

[2] F. A. Mussa-Ivaldi and L. E. Miller, "Brain–machine interfaces: computational demands and clinical needs meet basic neuroscience," *TRENDS in Neurosciences*, vol. 26, no. 6, pp. 329–334, 2003.

[3] J. R. Wolpaw, G. E. Loeb, B. Z. Allison, E. Donchin, O. F. do Nascimento, W. J. Heetderks, F. Nijboer, W. G. Shain, and J. N. Turner, "Bci meeting 2005-workshop on signals and recording methods," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 14, no. 2, pp. 138–141, 2006.

[4] Q. Ding, K. Tong, and G. Li, "Development of an eog (electro-oculography) based human-computer interface," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*. IEEE, 2006, pp. 6829–6831.

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

[6] Z. Lv, X. Wu, M. Li, and C. Zhang, "Implementation of the eog-based human computer interface system," in *Bioinformatics and Biomedical Engineering, 2008. ICBBE 2008. The 2nd International Conference on*. IEEE, 2008, pp. 2188–2191.

[7] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.

[8] C. Postelnicu, D. Talaba, and M. Toma, "Controlling a robotic arm by brainwaves and eye movement," *Technological Innovation for Sustainability*, pp. 157–164, 2011.

[9] Y. Zhang, Q. Zhao, J. Jin, X. Wang, and A. Cichocki, "A novel bci based on erp components sensitive to configural processing of human faces," *Journal of neural engineering*, vol. 9, no. 2, p. 026018, 2012.

[10] J.-X. Ma, L.-C. Shi, and B.-L. Lu, "Vigilance estimation by using electrooculographic features," in *Proce. of the 32nd International Conference of the IEEE Engineering in Medicine and Biology Society, Buenos Aires, Argentina*, 2010, pp. 6591–6594.