

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/268281484>

# A Novel EOG/EEG Hybrid Human–Machine Interface Adopting Eye Movements and ERPs: Application to Robot Control

Article in IEEE transactions on bio-medical engineering · November 2014

Impact Factor: 2.35 · DOI: 10.1109/TBME.2014.2369483 · Source: PubMed

---

CITATIONS

3

---

READS

115

4 authors, including:



**Yu Zhang**

East China University of Science and Technol...

42 PUBLICATIONS 386 CITATIONS

SEE PROFILE



**Andrzej Cichocki**

RIKEN

856 PUBLICATIONS 18,216 CITATIONS

SEE PROFILE



**Fumitoshi Matsuno**

Kyoto University

328 PUBLICATIONS 2,334 CITATIONS

SEE PROFILE

# A Novel EOG/EEG Hybrid Human-Machine Interface Adopting Eye Movements and ERPs: Application to Robot Control

Jiaxin Ma, Yu Zhang, Andrzej Cichocki, *Fellow, IEEE*, and Fumitoshi Matsuno, *Member, IEEE*

**Abstract**—This study presents a novel human-machine interface (HMI) based on both electrooculography (EOG) and electroencephalography (EEG). This hybrid interface works in two modes: an EOG mode recognizes eye movements such as blinks, and an EEG mode detects event related potentials (ERPs) like P300. While both eye movements and ERPs have been separately used for implementing assistive interfaces which help patients with motor disabilities in performing daily tasks, the proposed hybrid interface integrates them together. In this way, both the eye movements and ERPs complement each other. Therefore, it can provide a better efficiency and a wider scope of application. In this study, we design a threshold algorithm which can recognize four kinds of eye movements including blink, wink, gaze, and frown. In addition, an oddball paradigm with stimuli of inverted faces is used to evoke multiple ERP components including P300, N170, and VPP. To verify the effectiveness of the proposed system, two different online experiments are carried out. One is to control a multi-functional humanoid robot, and the other is to control four mobile robots. In both experiments, the subjects can complete tasks effectively by using the proposed interface whereas the best completion time is relatively short and very close to the one operated by hand.

**Index Terms**—Electrooculogram (EOG), Electroencephalogram (EEG), event-related potential (ERP), human-machine interface (HMI), robot control.

## I. INTRODUCTION

Brain-machine interface (BMI), also called brain-computer interface (BCI), is a communication system that allows direct connection between a human brain and a computer or other external device [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 become unavailable or occupied. There are a variety of noninvasive techniques measuring brain activities: functional magnetic resonance imaging (fMRI) [3], near-infrared spectroscopy (NIRS) [4], [5], magnetoencephalography (MEG) [6], electroencephalography (EEG) [7], and so on. Among them, EEG has high time resolution, less environmental limits, and requires relatively

inexpensive equipments [7]. It has been largely used in both clinical and research applications.

The EEG-based BCI modalities can be categorized into four different types: event-related desynchronization/synchronization (ERD/ERS) [8], [9], steady state visual evoked potentials (SSVEP) [10], event-related potentials (ERP) [11]–[13], and slow cortical potentials (SCP) [14]. Among these, ERP and SSVEP-based BCIs are more practical than others because they support large numbers of output commands, and need little training time. ERPs are brain responses to specific cognitive tasks. P300 is one of the most often used ERP components. It is a positive deflection in EEG over parietal and occipital cortex, occurring approximately 300 ms after a rare but task-relevant stimulus [15]. P300-based BCI has relatively robust performance for target detection. Although, its information transfer rate (ITR) is at a medium level, but unlike SSVEP-based BCI, it does not cause some subjects feeling annoyed or fatigued by the flickering stimuli [16], [17]. One of its representative applications is the P300 speller which is used for inputting characters [18].

Electrooculography (EOG) measures voltage fluctuations resulted from eye movements. EOG signals are generated by eye saccades or pursuit movements as well as blinks. EOG can be used to track the eye-gaze direction, doing similar work as an optical (video-based) eye tracker. It also contains highly recognizable information of eyelid movements such as blinks and winks. Although in EEG-based BCI, EOG signals are usually considered as a major noise source to be removed [19], [20], EOG alone can make up another kind of human machine interface. For an EOG-based system, the response speed can be considerably high, which is desirable especially for control applications. There have already been a lot of studies about designing EOG-based human-machine interfaces as described by [21]–[23].

However, most current techniques of EEG or EOG-based interfaces still face challenges that prevent these from being accepted by the majority in the clinical applications. For example, the main obstacle of the P300-based BCI (and almost all BCIs) is its relatively low ITR: the response time of the system is unsatisfactory for most daily tasks, and the accuracy is also not perfect. Moreover, P300 is a synchronous BCI system which receives inputs and generates outputs at specified time intervals. In other words, when users do not want to send commands, the system should be off to prevent unwanted outputs. Considering this point, additional means to make the system active/inactive are often needed. For example, SSVEP has been used in P300 BCI as a switch [24], but this approach still has some problems like causing vision tiredness.

For EOG interfaces, the main problem is that they do not adequately support large numbers of outputs. The number of

This study was supported in part by the Nation Nature Science Foundation of China under Grant 61305028, Fundamental Research Funds for the Central Universities under Grant WH1314023.

J. Ma and F. Matsuno are with the Department of Mechanical Engineering and Science, School of Engineering, Kyoto University, Kyoto 6158530, Japan (email: ma.jiixin.62c@st.kyoto-u.ac.jp; matsuno@me.kyoto-u.ac.jp).

Y. Zhang is with the Key Laboratory for Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China (email: yuzhang@ecust.edu.cn)

A. Cichocki is with the Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, and RIKEN-BSI Toyota Cooperation Center (BTCC), Wako-shi, Saitama 351-0198, Japan, and also the System Research Institute, Polish Academy of Sciences, Warsaw 00-901, Poland (e-mail: a.cichocki@riken.jp).

reliable EOG commands is usually limited to several eyelid movement patterns. Comparing with video-based eye trackers, an EOG-based eye tracking system is light-weight but poor in accuracy, thus making it unsuitable for fine control of a graphical user interface (e.g. as a replacement of the computer mouse). Moreover, unlike BCI, EOG interface is muscle-dependant which means that it inevitably causes fatigue over a longer use.

Another problem is arises due to the needs of users. The expected users of such interfaces may have various kinds of needs (e.g. controlling a wheelchair to move around, sending messages to contact other people, and so on). Usually, a single EEG or EOG-based system can only manage one certain kind of task. It is rather difficult to have a universally robust system applicable to different situations.

To overcome the limitations and disadvantages, an extensive amount of work in hybrid BCI has been invested in recent years [25]. A hybrid BCI is typically a combination of two different types of BCI systems, or BCI and non-BCI systems (EOG, EMG, etc.). These hybrid systems allow for different subsystems to be chosen for different tasks, making them more flexible. Table I (partially from [25]) shows the recent studies about hybrid brain or non-brain human-machine interfaces.

In this study, we propose a novel EOG/EEG hybrid human-machine interface adopting eye movements and ERPs, and then apply it to robot control tasks. The proposed method has following novelties and features.

- The proposed hybrid HMI combines EOG and ERP interfaces. According to Table I, there are no existing studies carried on EOG/ERP(P300) hybrid system.
- The EOG method used in this study can detect four kinds of eye movements: blink, wink, gaze, and frown, while other similar studies only focus on blink and gaze.
- Conventional ERP-based interfaces only utilize the P300 component. Our ERP paradigm uses inverted face images as stimuli which evoke not only the P300 component, but also the VPP and N170 components, and thus, is expected to give a better accuracy.
- Two different robot control experiments have been carried out to verify the proposed system. The experiments include single robot and multiple robots control.

The proposed hybrid interface works in two modes: an EOG mode and an EEG mode. The two subsystems are equally important and have separate functions: EOG for fast-response tasks and EEG (ERP) for menu-selection tasks. Therefore, the overall system becomes versatile and flexible. In many other studies on hybrid BCIs, one subsystem only works in assistance of the other. For example, among the SSVEP/P300 hybrid interfaces mentioned in Table I, SSVEP is used as an on/off switch of P300 [24], an idle/non-idle state detector of P300 [26], or is used to divide the character matrix of a P300 speller into several subareas to increase ITR [27]. All these implementations enhanced the original function of P300 interface but did not improve the versatility and flexibility.

The multi-ERP (P300, VPP and N170) paradigm based on stimuli of inverted faces belongs to one of our previous studies [28]. The basic idea of this hybrid HMI, as well as a detailed design for a humanoid robot control scheme, has

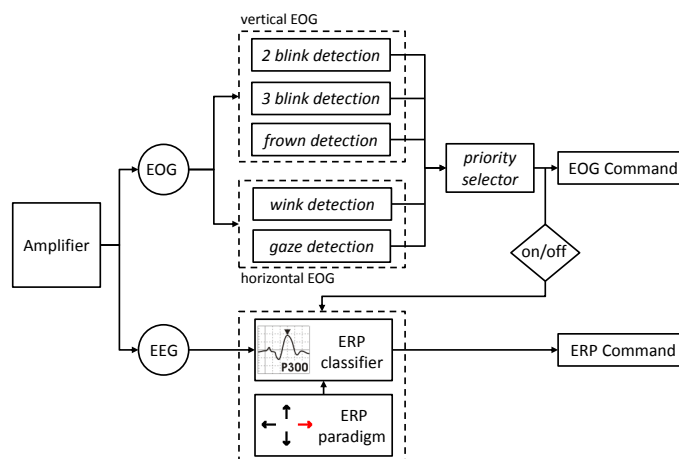


Fig. 1. The proposed model of the hybrid HMI

been introduced in another pilot study [29]. So these parts are predigested, and the main contents in this paper are about the hybrid HMI details, the eye movement detection algorithm, and discussions.

## II. MODEL OF HYBRID INTERFACE

In practice, a lot of control scenarios require multitasking. As mentioned in the introduction section, the significance of our hybrid interface is its versatility and flexibility. Combining eye movements and ERPs can potentially make full use of the advantages of both systems and help to overcome the disadvantages. By using eye movements, the system can achieve very high ITR, which compensates for the largest weakness of ERP interfaces. By using ERPs, a graphical user interface can be realized and large numbers of commands can be supported more easily. In addition, user experience like convenience is also an important feature. For using the EOG interface, repeatedly performing eye movements will easily accumulate fatigue on muscles. Although the ERP interface casts no physical burden on users, continuously watching the flashing cues on the screen can also lead to impatience and weariness, and thus decrease the system performance. With the hybrid interface, users do not need to constantly concentrate on the same operation mode, which potentially relieves users' burden, both on body and mind.

Figure 1 illustrates the whole model of the proposed hybrid interface. The upper part is EOG processing. This part works asynchronously, which means the system is always actively detecting the eye movements and the user can send EOG commands at any time. Input EOG signals are divided into vertical EOG and horizontal EOG to detect four kinds of eye movements: blink, frown, wink, and gaze. More specifically, the system detects double blinks, triple blinks, and frowns from the vertical EOG, and detects winks (left/right) and gazes (towards left/right) from the horizontal EOG. These are all common eye movements and can be performed easily (without great efforts) and non-voluntarily. The single blink is not involved in this model because people always spontaneously make single blinks. Generally, it is difficult to distinguish intentional blinks from spontaneous blinks. In our model,

TABLE I  
RECENT STUDIES ABOUT HYBRID BRAIN OR NON-BRAIN HUMAN-MACHINE INTERFACES

Literatures	Brain										
	Non-Invasive							Invasive	Non-Brain		
	EEG			Non-EEG					EOG	EMG	ECG
P300	SSVEP	ERD	SCP	fMRI	NIRS	MEG	ECoG	EOG	EMG	ECG	
[17] [30] [31] [32] [33]		✓	✓								
[34]			✓						✓		
[35] [36]			✓			✓					
[37]			✓							✓	
[24] [26] [27]	✓	✓									
[38] [39] [40]	✓		✓								

we avoid this problem by only concerning double and triple blinks but ignoring single blinks. The vertical gaze is also not involved because when compared with horizontal gazes, vertical ones are more unreliable and easy to be performed spontaneously.

The priority selector is used to determine the priority of different types of eye movements. In this model, each kind of eye movement is detected by an individual sliding window on the corresponding EOG channel. Because all kinds of eye movements are detected in parallel, there can be a time in which two or more eye movements exist. A most common example could be: a triple blink always contains double blinks, so at a time both “triple blink” and “double blink” will be detected. Since there can only be one output, eye movements are assigned priorities. In our study it is like this: wink > frown > gaze > triple blink > double blink.

The lower part is EEG processing. Similar to a conventional P300 system, it includes an ERP paradigm and an ERP classifier. The ERP paradigm contains a dynamic graphical interface. It displays icons (e.g. arrows) on the screen, where each icon represents one ERP command. Stimuli of inverted face images are continuously flashed upon these icons in a random sequence. When the user focuses on one target icon, the stimulus flashed on that icon will evoke ERPs including P300, VPP and N170. At the same time, the ERP classifier analyzes the EEG signals of each time interval following a stimulus. The length of the time interval should be proper to contain the evoked ERPs (in this study it is 700 ms from the beginning of each stimulus). The classifier identifies which time interval is most likely to contain the evoked ERPs, and thus determines the output ERP command.

Since the ERP interface works synchronously, as mentioned in the previous section, there must be an external command to enable/disable it. An eye movement is a suitable choice for this task. Our system has two modes: EOG mode and EEG mode. An eye movement (in this study, frown) is used to switch between the two modes. In EOG mode, the ERP interface (including ERP paradigm and ERP classifier) is inactive, which means no image stimulus will be shown, and no EEG signal will be analyzed. In EEG mode, the ERP interface is active. By this design, the user can easily start and end an ERP trial any time through switching between the two modes. Although ERP interface can also have a “sleep”

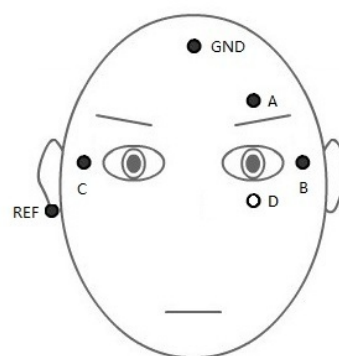


Fig. 2. The placement of EOG electrodes. A: vertical EOG electrode; B, C: horizontal EOG electrode; REF: reference; GND: ground.

command to shut down itself (turning on is impossible), but it is definitely more convenient to use an eye movement to do so.

On the other hand, even in EEG mode, the system still continuously detects eye movements from EOG signals. One reason is just mentioned above that the eye movement is responsible for mode switching. For other kinds of eye movements unused in EEG mode, the system can be set as unresponsive. However, eye movements can have more roles than this in EEG mode, one example is to report the error of ERP classification. The accuracy of ERP classification varies across individuals and is hardly perfect for most people. Incorrect results lead to unwanted command being sent and executed. To avoid this, it is better to have an error report mechanism as an assurance. In our current setting, a 1 sec delay is added before sending out an ERP command. During this time if the user finds the result shown on the screen is incorrect, he or she can perform a wink immediately to cancel it out and then no command will be sent.

The specific contents of EOG and ERP commands can be determined depending on various situations and tasks. But generally, EOG commands suit tasks that require fast response, while ERP commands suit a form of selection menu. In our online robot control experiments, EOG commands are mainly used to control the robot’s move, and ERP commands are to let the robot perform some preprogrammed behaviors, or to select the control target from multiple robots.

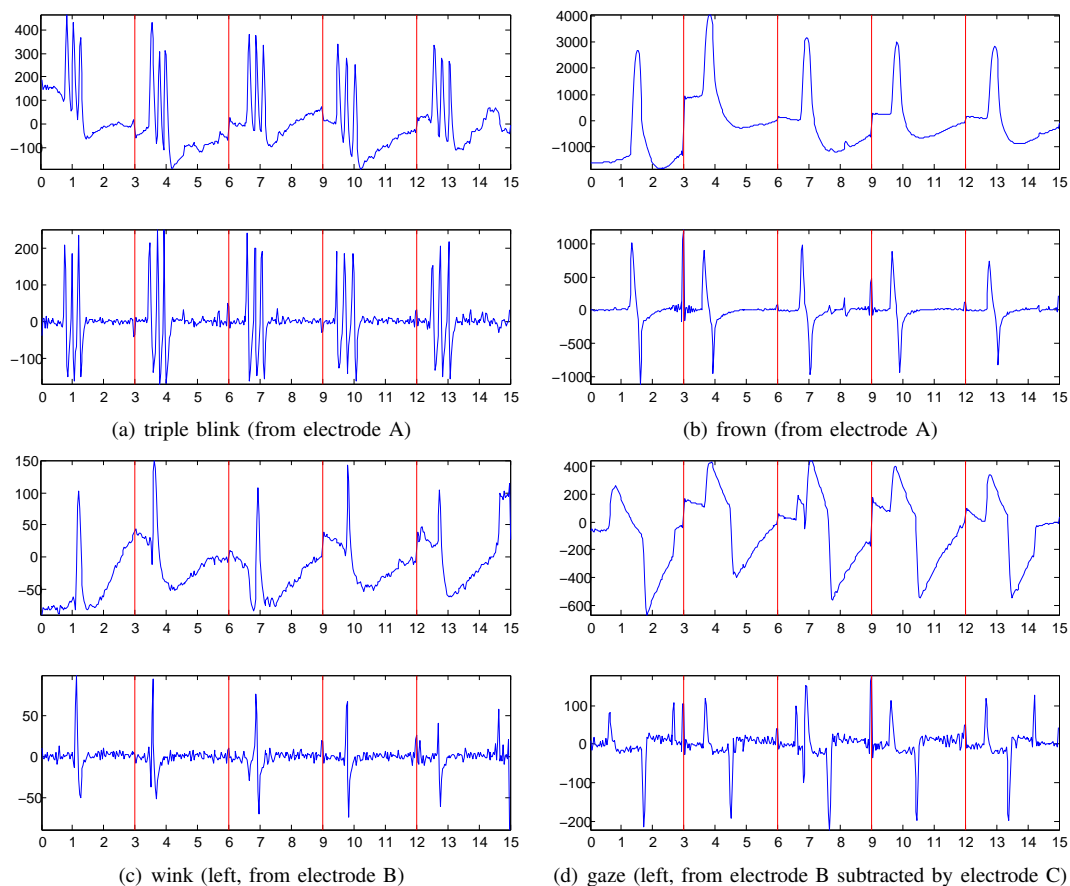


Fig. 3. Typical waveforms of (a) blinks, (b) frown, (c) wink, and (d) gaze. For each, the upper plot is the original waveform whose y-axis is voltage ( $\mu\text{V}$ ) and x-axis is time (s), and the lower plot is the differentiated result (first-order difference) of the upper one. Here each 3 second duration is considered as a single trial. Intervals between trials are omitted.

### III. EOG ANALYSIS

#### A. EOG acquisition

In this study, various eye movements from EOG signals are recognized and then applied as the commands of robot control. The EOG electrodes are placed as shown in Fig. 2 where electrode A is for recording vertical EOG; electrodes B and C are for recording horizontal EOG. All of the electrodes are monopolar. Other two electrodes: ground and reference, are shared with EEG recording. Unlike the commonly used dipolar placement of EOG electrodes (i.e., B-C dipolar and A-D dipolar), this placement shares the reference electrode with EEG, so that electrode D is saved.

The signals were recorded by g.USBamp with g.GAMMAbox (g.tec medical engineering, Austria). The original sampling rate of EOG was 256 Hz. It was down-sampled to 32 Hz because the proposed eye movement detection algorithm prefers smooth data. Higher sampling rate may cause undesirable fluctuations. The band-filter was chosen as 0.1-30 Hz (built-in). The 0.1 Hz lower-cutoff frequency is to eliminate the effect of baseline drift, and the 30 Hz upper-cutoff frequency is to remove high-frequency noise. In addition, the choice of lower-cutoff frequency influences the signal shape, which may also affect the optimal values of algorithm parameters. Our algorithm was based on 0.1 Hz lower-cutoff frequency.

#### B. Eye movement detection algorithm

In this study, four kinds of common eye movements: multi-blink (blink twice or three times quickly), frown, wink (one-sided blink), and gaze (horizontal, towards left or right) are detected. All these eye movements have recognizable waveform shapes. While threshold method is commonly used for blink detection [41], [42], we designed a simple and effective multi-threshold algorithm applicable to all these four kinds of eye movements.

Figure 3 illustrates the waveforms of the four kinds of eye movements, from which it can be observed that the waveforms of different eye movements have common characteristic. All of them consist of a positive peak in original EOGs, and a

TABLE II  
A SUMMARY OF DIFFERENT EYE MOVEMENTS

	blink	frown	wink	gaze
speed	M	L	S	S
amplitude	M	L	S	M
duration	S	L	M	M*
EOG channel	V	V	H	H**

L: large, M: medium, S: small, V: vertical, H: horizontal

\* only including the eyeballs moving forth

\*\* only considering horizontal gazes

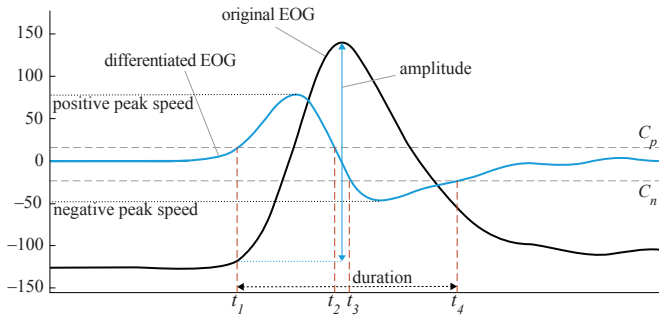


Fig. 4. The waveform of a blink in vertical EOG and its differentiated shape [41]. The x-axis is time and y-axis is voltage ( $\mu V$ ). The derivative is approximated by the first-order difference  $\frac{x(t) - x(t - \Delta t)}{\Delta t}$  in our program.

positive peak followed by a negative peak in differentiated signals. This feature can be utilized to locate eye movements. Next, it can also be observed from Fig. 3 that different types of eye movements have different speed, amplitude, and time duration features. A summary can be made as shown in Table II. According to Table II, it is possible to distinguish different types of eye movements by the same algorithm. The proposed multi-threshold algorithm takes the speed, amplitude, and duration as three kinds of features to determine an eye movement event.

To explain the rough idea of the algorithm, see Fig. 4 as an example. The figure shows a blink waveform and its differentiated shape. To locate the blink, we need to find a positive peak (the interval between  $t_1$  and  $t_2$ ) and a negative peak (the interval between  $t_3$  and  $t_4$ ) on the differentiated signal. Then the blink event is just the interval between  $t_1$  and  $t_4$ . To further verify if it is a blink, three kinds of thresholds (speed, amplitude, and duration) were applied to identify its eligibility. These threshold values are determined by a calibration process before each experiment.

Assume that  $f(t)$  is data of original EOG corresponding to a channel, and  $f'(t)$  is the data of time differential of the original EOG signal. For detecting different kinds of eye movements  $f(t)$  is also from different channels. As Fig. 2 shows the electrode positions, for blink and frown detection,  $f(t)$  is from channel A; for wink detection,  $f(t)$  is channel B (left) or C (right); for gaze detection,  $f(x)$  is channel B–C (left) or C–B (right). The detailed steps of the algorithm are as follows.

1) *Locate all the peaks*: Assume that we already have following thresholds:  $S_{min}$ ,  $S_{max}$ ,  $A_{min}$ ,  $A_{max}$ ,  $D_{min}$ , and  $D_{max}$ , which represent minimal and maximal thresholds of speed (S), amplitude (A), and duration (D), respectively.

The first step is to find all the points  $t_1$ ,  $t_2$ ,  $t_3$ , and  $t_4$  which satisfy the following inequalities:

$$\begin{aligned} f'(t_1 - \Delta t) < C_p, f'(t_1) \geq C_p, f'(t_1 + \Delta t) \geq C_p, \\ f'(t_2 - \Delta t) \geq C_p, f'(t_2) \geq C_p, f'(t_2 + \Delta t) < C_p, \\ f'(t_3 - \Delta t) > C_n, f'(t_3) \leq C_n, f'(t_3 + \Delta t) \leq C_n, \\ f'(t_4 - \Delta t) \leq C_n, f'(t_4) \leq C_n, f'(t_4 + \Delta t) > C_n, \end{aligned} \quad (1)$$

where in the algorithm, by default we have  $C_p = 10$ ,  $C_n = S_{min}$ . This step locates all the eligible peaks, where  $[t_1, t_2]$

has a positive peak whose initial and final value is equal to  $C_p$ , and  $[t_3, t_4]$  has a negative peak whose initial and final value is equal to  $C_n$ . Here  $C_p$  is given as a smaller number i.e. 10, because the more  $C_p$  is close to zero, the more accurate is  $t_1$ , which is also the start point of eye movements. In the same way, if  $C_n$  is close to zero, the end point of eye movements  $t_4$  will be accurate. However, in an online running environment, early detection is much more preferred than accurately locating the end point. That is why  $C_n$  is set equal to  $S_{min}$ , but not some value close to zero.

For clarity, the results are sorted to multiple pairs of  $\{t_1, t_2\}$  (for positive peaks) and  $\{t_3, t_4\}$  (for negative peaks).

2) *Apply speed thresholds ( $S_{min}$  and  $S_{max}$ )*: The second step is to pick out all the eligible peaks that satisfy the following inequalities:

$$\begin{aligned} \max_{t \in [t_1, t_2]} f'(t) \geq S_{max}, \\ \min_{t \in [t_3, t_4]} f'(t) \leq S_{min}. \end{aligned} \quad (2)$$

This step is to make sure that the maximal  $f'$  value of each positive peak exceeds the threshold  $S_{max}$ , and the minimal  $f'$  value of each negative peak is lower than the threshold  $S_{min}$ . Note that if in the previous step  $C_n = S_{min}$ , here the second equation will be automatically satisfied.

The result comprises of some pairs of  $\{t_1, t_2\}$  (positive peaks) and  $\{t_3, t_4\}$  (negative peaks). Then, adjacent  $\{t_1, t_2\}$  and  $\{t_3, t_4\}$  are grouped together to form a complete eye movement candidate  $\{t_1, t_2, t_3, t_4\}$ .

3) *Apply amplitude threshold ( $A_{min}$ ,  $A_{max}$ )*: The third step is to pick out all the eligible eye movement events that satisfy the following inequality:

$$A_{min} \leq \max_{t \in [t_1, t_4]} f'(t) - f(t_1) \leq A_{max}, \quad (3)$$

where  $\max_{t \in [t_1, t_4]} \{f'(t)\} - f(t_1)$  means the maximal voltage in an eye movement subtracted by its initial voltage, which leads to the amplitude.

4) *Apply duration threshold ( $D_{min}$ ,  $D_{max}$ )*: The last step is to pick out all the eligible eye movement events that satisfy the following inequality:

$$D_{min} \leq t_4 - t_1 \leq D_{max}. \quad (4)$$

5) *Special cases for wink and gaze*: As mentioned above, this algorithm is suitable for all four kinds of eye movements as long as the threshold values are properly set. But still, there are a few differences in processing wink and gaze.

First, wink and gaze have some similarities. Only using thresholds cannot separate them perfectly, so there is an additional step to distinguish them from each other. For gaze (left/right), the signals of channel B and C are almost inverse, while in the case of wink (left/right), the signals from these two channels are more likely to have the same trend. For example, a left wink leads to a large positive peak in channel B, and a small positive peak or no peak in channel C. Therefore, we can compare the linear correlation between channel B and C, to determine if the eye movement event is a wink or a gaze. The criterion is:

$$\begin{aligned} \rho(f_B(T_1), f_C(T_1)) > -0.8 \rightarrow \text{wink}, T_1 = [t_1, t_4] \\ \rho(f_B(T_2), f_C(T_2)) < -0.8 \rightarrow \text{gaze}, T_2 = [t_1, t_2], \end{aligned} \quad (5)$$

where  $\rho(X, Y)$  is Pearson's linear correlation coefficient ranging from -1 to +1. Here +1 means total positive correlation, 0 means no correlation, and -1 means total negative correlation.

The above criterion uses  $\{t_1, t_2\}$  to determine a gaze instead of  $\{t_1, t_4\}$ . The reason is explained as follows. For a gaze, times of the positive peak and the negative peak in a differentiated EOG are highly separated (see Fig. 3(d)), as the process of the eyes moving to one side and then moving back is much slower compared to other eye movements. As waiting for the occurrence of the negative peak will cause too much delay, for gaze detection the negative peak is completely ignored. So to detect gaze, for step (1) and (2), remove all the procedures related to  $t_3$  and  $t_4$ , and for other steps, replace  $t_4$  by  $t_2$ .

### C. EOG calibration and thresholds calculation

A calibration process is needed to determine the specific threshold values. The time duration of eye movements is stable and can be fixed, but speed and amplitude values change in every experiment because the subject conditions and the electrode positions are inconsistent. Before a subject begins the real-time experiment, he or she is asked to go through a calibration process first. During the calibration, on the screen it shows the texts of "ready (1 s)→movement name (3 s)→relax (1 s)" repeatedly, where "movement name" includes "frown", "triple blink", "left wink", "right wink", "gaze left", and "gaze right", and is also displayed in this sequence. For each movement the cue is repeated for ten times.

In the calibration process, all the eye movements are time-locked so they can be easily detected even without thresholds. For frown, wink, and gaze, we simply find the peak amplitude in the middle 3 seconds of each trial, and define a short interval  $T$  (2 s for frown and wink, 1 s for gaze) centered on the peak value to represent the eye movements. For triple blink, since there are three individual peaks, we use predefined threshold values ( $S_{max} = 10$ ,  $S_{min} = -10$ ,  $A_{min} = 150$ ) to locate all the three blink events. With these eye movement samples, positive peak speed  $S_p$ , negative peak speed  $S_n$ , and amplitude  $A$  for each eye movement can be calculated as

$$\begin{aligned} S_p &= \max_{t \in T} f'(t), \\ S_n &= \min_{t \in T} f'(t), \\ A &= \max_{t \in T} f(t) - f(t_0), \end{aligned} \quad (6)$$

where  $t_0$  is the first time point of  $T$ . Then the thresholds are calculated as in Table III. Thresholds, which are not mentioned in this table, are not necessarily needed for our experiments.

## IV. EEG ANALYSIS

### A. EEG acquisition

In this study, 8 electrodes were used to record EEG: Fz, Cz, P7, P3, Pz, P4, P8, and Oz (Fig. 5). 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.

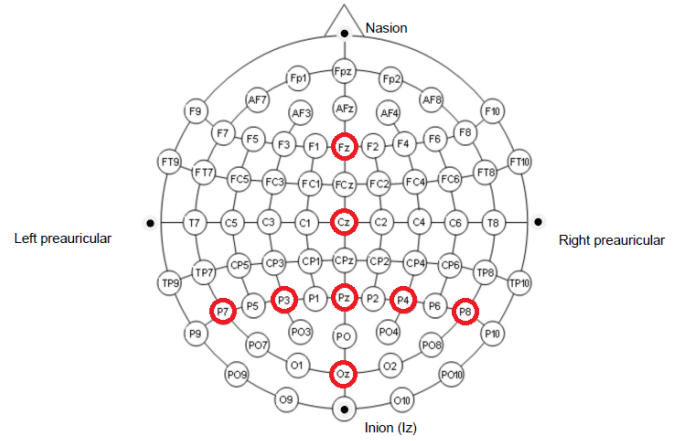


Fig. 5. Electrode positions for EEG recording. (Ground and reference electrodes are shown in Fig. 2.)

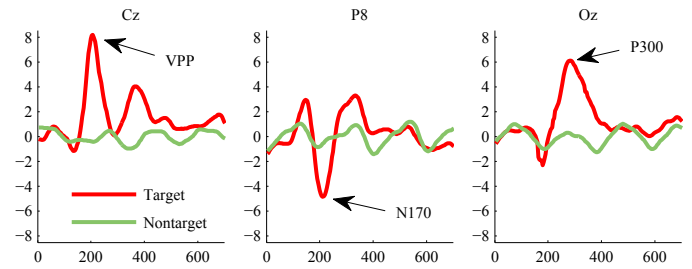


Fig. 6. Grand average ERP waveforms derived from the target and nontarget stimuli of the inverted face image. VPP, N170, and P300 can be clearly observed from channel Cz, P8, and Oz. The y-axes are voltage ( $\mu\text{V}$ ) and x-axes are time (ms).

The original sampling rate of EEG was 256 Hz and was down-sampled to 64 Hz. This was not specifically designed, but just for reducing redundant information that might result in over-fitting. The filter was the same which was used in EOG.

### B. ERP paradigm

This study adopts a more advanced ERP paradigm which combines oddball presentation and inverted face perception. This paradigm mainly exploits three ERP components, namely VPP, N170 and P300, instead of only P300. According to our previous work [28], it can significantly improve the target

TABLE III  
THRESHOLDS FOR EYE MOVEMENT DETECTION

Th.	Speed	Amplitude	Duration (s)
Blink	$S_{max} = 0.5 \min S_p^*$ $S_{min} = 0.5 \max S_n$	$A_{min} = 0.8 \min A$	$D_{max} = 0.5$
Frown	$S_{max} = 0.4 \min S_p$ $S_{min} = 0.5 \max S_n$	$A_{min} = 0.8 \min A$	$D_{min} = 0.4$ $D_{max} = 2.0$
Wink	$S_{max} = 0.5 \min S_p$ $S_{min} = 0.5 \max S_n$	$A_{min} = 0.8 \min A$	$D_{min} = 0.1$ $D_{max} = 0.5$
Gaze	$S_{max} = 0.8 \min S_p$	$A_{min} = 0.8 \min A$ $A_{max} = 1.4 \min A$	$D_{max} = 0.5$

\* Here  $\min S_p$  means the minimal positive peak speed of all the blinks in the calibration process.

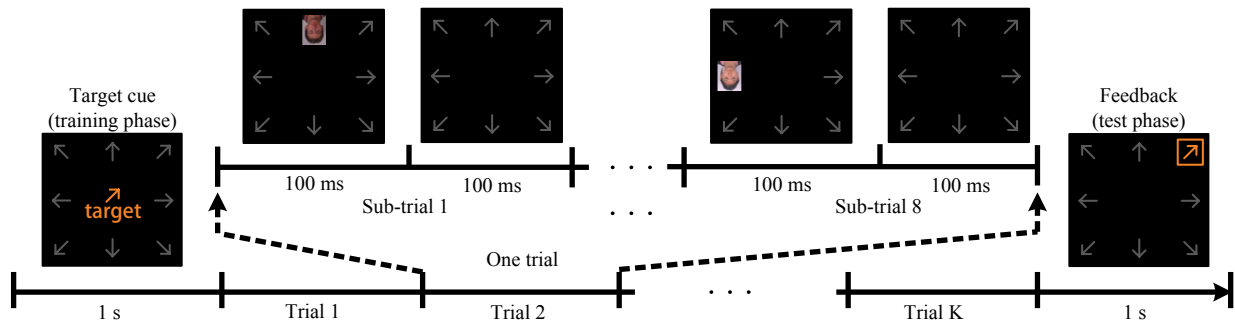


Fig. 7. The timeline of a single run. In the training phase, a target cue was provided, and the number of trials  $K$  was 5. In the test phase, the number of trials  $K$  was 2 and feedback was provided. Each trial consisted of eight sub-trials in each of which one stimulus was randomly presented in one of the eight directions for 100 ms with an inter-stimulus interval of 100 ms.

detection performance in contrast to the stimulus intensification pattern used in the conventional P300-based system.

Among the ERP components, N170 and VPP are evoked by the configural processing of facial image, and P300 is evoked by oddball event. Typical ERP waveforms are illustrated in Fig. 6.

The ERP interface used in this study has 8 arrow icons placed in 8 directions of screen (N, W, S, E, NW, NE, SW, SE). For one trial, stimuli (invert facial images) were displayed upon each arrow icon once, in random order. One stimulus was presented for 100 ms. After another 100 ms interval, the next stimulus was displayed. So one trial has  $200 \text{ ms} \times 8 = 1.6 \text{ s}$ .

It is possible to classify which icon the subject is focusing on by only analyzing a single trial, but to improve the classification accuracy, usually more than one trial are used to determine one output. We call such group of trials a run. Our training phase contains eight runs. Each run corresponds to one of the eight arrow icons as the target, and is of 5 successive trials. At the beginning of each run, there will be a 1 s cue instructing the subject which target he or she should focus on.

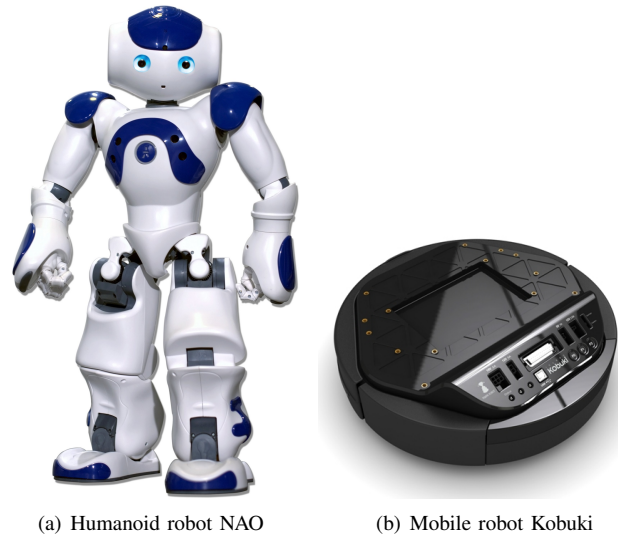
After training, subjects will go through a test phase which also contains eight runs. In the test phase each run only contains 2 trials so as to speed up the process and increase the difficulty. The test phase does not have cues at the beginning of each run; instead, the classification result (one of the 8 arrows) will be highlighted as a feedback to subjects after each run. With the feedbacks, subjects can be aware of their performance.

Figure 7 illustrates a detailed timeline of a single run.

### C. ERP classification

Linear discriminant analysis (LDA) was used to classify which target the subject is focusing on. Before the online experiment, EEG data collected from the training phase were used to train the LDA classifier first.

A 700 ms data segment was extracted from the beginning of each flash stimulus and baseline corrected by a 100 ms pre-stimulus interval. A total of 320 such data segments consisting of 40 targets and 280 non-targets were derived from each subject. To reduce the feature dimensionality, each data segment was further downsampled to 15 temporal points (approximately 21 Hz). The training samples (i.e., feature



(a) Humanoid robot NAO

(b) Mobile robot Kobuki

Fig. 8. Experimental robots.

vectors) were then formed by the concatenation of 15 temporal points at 8 channels from the data segments. That is, the dimensionality of feature vector is  $8 \times 15 = 120$ . The extracted feature vectors were then fed to train the LDA classifier for the subsequent online application.

## V. EXPERIMENTS

### A. Experiment introduction

In this study, two kinds of online experiments were carried out to verify the proposed system. One was humanoid robot control, in which the experimental robot was NAO (Aldebaran robotics, Inc., Fig. 8(a)). The other was multiple mobile robots control, where the experimental robots were Kobuki (Yujin Robot Co. Ltd., Fig. 8(b)). For simplicity, hereinafter the two experiments are called the “NAO experiment” and the “Kobuki experiment”. The experimental platform was MATLAB Simulink (The MathWorks, Inc.). The details of the two experiments are introduced in this section.

Totally 13 subjects (numbered S1~S13) were involved in the two experiments. Among them, two (S7 and S10) are female and others are male. All the subjects are able-bodied, aged from 22 to 30. Four subjects: S1, S2, S3, and S4 participated in the NAO experiment, and ten subjects: S1,



S5~S13 participated in the Kobuki experiments. In addition, S1~S4 have some knowledge about BCI while others do not.

In both experiments, EOG and EEG were acquired as the descriptions of Section III and IV. For each subject, EOG was calibrated and EEG was trained before doing the online experiments. The target of the online experiments was to control robots to complete a series of tasks. Subjects needed to send different commands to robots in EOG mode and EEG mode. In EOG mode, the system was asynchronous, which means whenever the user performs an eye movement, the robot responds. The timing of sending commands was held by the user. If the user did not send any command, the robot would make no action. The eye movement detection algorithm must be robust enough to prevent any false detection which leads to undesired output. On the other hand, EEG mode was synchronous, which means the user had to follow the system's predefined pace. In this mode, even the user did not look at the screen, the system would still analyze the EEG and output most probable results. So, when the user was idle, he needs to switch to EOG mode in order to inactivate the ERP paradigm.

In EEG mode, one output was generated after 16 stimuli were given (8 icons by 2 trials) where each stimulus consisted of 100 ms highlight and 100 ms dark, which was 3.2 s. Including the time of showing the result and any other interval of delay, the total time to generate an ERP command was about 6 s, whereas using eye movement to generate a command only took around 1 s. Considering this fact, commands requiring fast response such as moving and turning are assigned to EOG mode; other functions which do not require fast response are assigned to EEG mode. In addition, the left/right eye movements respectively correspond to leftward/rightward movements of robots. The detailed commands for both experiments are listed in Table IV.

Note that in the NAO experiment, when a wrong ERP command was generated, the subject needed to "frown twice" to cancel that command (frown once to stop the current output and switch to EOG mode, then frown again to switch back). Since this operation was too ineffective, we implemented a new command "right wink" in EEG mode to report classification error and cancel the corresponding execution. However, this function was unused in the Kobuki experiment where the ERP commands were only related to robot selection, not behavior execution.

### B. Humanoid robot control

NAO is a multi-functional humanoid robot, programable to perform various complex behaviors. The NAO experiment stimulates a scenario that a person controls a humanoid robot to make simple communication with other people. The detailed experiment scenario is described as follows.

- The robot starts from the point S (see Fig. 9), and moves to the point A by a provided route, then receives an object from a person at the point A.
- Holding the object, the robot moves to the point B, and gives the object to a person at the point B.
- Then the robot moves to the point C and performs a dance.

TABLE IV  
COMMANDS FOR THE TWO ROBOT-CONTROL EXPERIMENTS

Mode	Command	NAO Exp.	Kobuki Exp.
EOG mode	double blink	stop	stop
	triple blink	go head	go head
	left wink	turn 90° left	keep turning left
	right wink	turn 90° right	keep turning right
	left gaze	head 90° left	-
	right gaze	head 90° right	-
	frown	mode switch	mode switch
EEG mode	icon 1	receive an object	select robot No.1
	icon 2	hand over an object	select robot No.2
	icon 3	dance	select robot No.3
	icon 4	sit down	select robot No.4
	icon 5~8	-	-
	frown	mode switch	mode switch
	right wink	-*	error report**

\* unimplemented  
\*\* implemented, but unused

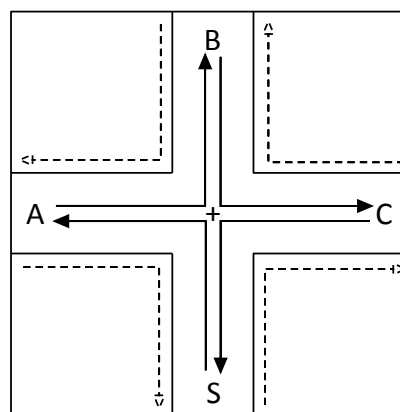


Fig. 9. Experimental layout. The solid arrows are routes of NAO, and the dashed arrows are routes of Kobukis.

- Finally the robot moves back to the point S and sits down.

The whole experiment was divided into 4 sessions. These sessions consist of similar tasks. Take the first session as an example, the detailed experiment steps are:

- The robot gets ready at the 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 goes ahead.
- The robot stops at the point A, then switches to EEG mode.
- The robot performs a provided behavior by ERP command.

For other sessions, only the robot positions and the provided behaviors are different. The four sessions (S→A, A→B, B→C, C→S) were scheduled separately.

### C. Mobile robots control

The purpose of the Kobuki experiment is to simulate a scenario that a person controls multiple robots to move around

and gather information. In this experiment, four Kobuki robots were used and moved to destinations one by one. These robots only have basic function of mobility. For simplicity, in this experiment they were placed near each other under the direct observation of the subject. But if necessary, they can also be mounted with a laptop and sensors like a camera and a GPS, which can enable them to remotely perform complicated tasks.

The experimental layout (Fig. 9) was designed similar to the previous experiment. Four Kobuki robots numbered 1~4 were placed at the right side of the marks (so they were unlikely to run into each other) S, C, B, and A, respectively. The detailed experiment steps are:

- The subject switches to EEG mode, and selects the No.1 near the point S (see Fig. 9).
- The No.1 moves ahead, and stops when near the center point.
- The No.1 turns right, and stops when facing its destination.
- The No.1 moves ahead, and stops when reaching its destination (near C).
- Repeated until all four robots are selected and moved to their destinations.

Each subject completed the experiment without any break, and repeated the experiment again after the first completion. The time for each completion was recorded.

## VI. RESULTS

### A. Offline evaluation

For all the 13 subjects involved in online experiments, their offline data (EOG calibration and ERP training) were recorded and evaluated. Since one subject participated in two different experiments, there are totally 14 groups of data, and the results are listed as follows.

1) *EOG*: 14 groups of calibration data are used for evaluation. Each group includes 6 kinds of eye movements: triple blink, frown, left wink, right wink, left gaze, and right gaze, with 10 repetitions of each. A comparison has been made between the proposed threshold method and a standard pattern matching method. The proposed method calculated the thresholds according to Table III. For the pattern matching method, the implementation is as follows: 10 repetitions of eye movements are averaged, and the result is considered as a standard eye movement pattern. Then a sliding window on original signals is compared with the standard pattern, by calculating the Pearson's linear correlation coefficient  $\rho$ . If  $\rho$  is larger than a predefined threshold, it means an eye movement is detected. We have carefully chosen a proper length of the time window, as well as a proper threshold of  $\rho$  for each kind of eye movement.

Figure 10 shows the recall of each kind of eye movement detected by the two methods. Lower recall means higher miss rate.

2) *ERP*: Figure 11 depicts the classification accuracy averaged on all subjects using one to five trials. The classification accuracy is calculated by cross-validation. Since each subject took 8 ERP runs (one run includes five trials) in training process, the cross-validation took 7 runs as training data, and

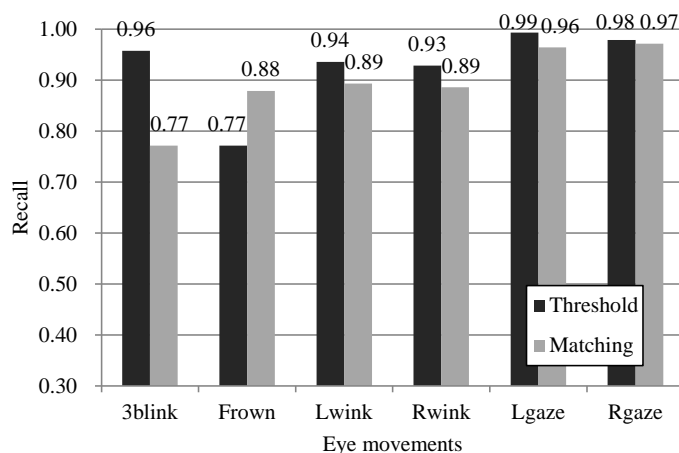


Fig. 10. The recall of six kinds of eye movements, detected by the proposed method and a pattern matching method. One kind of eye movement has 140 repetitions.

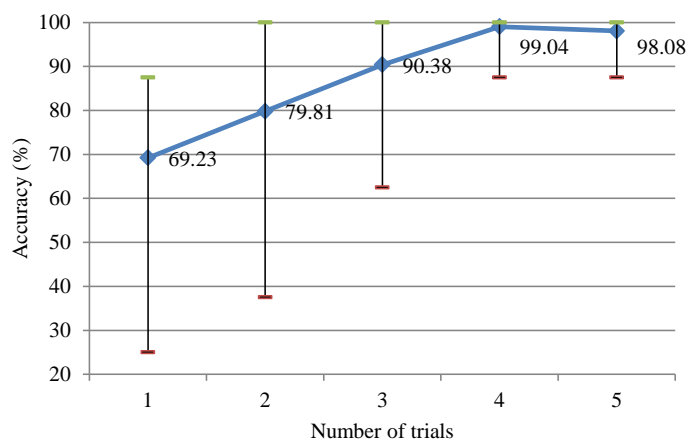


Fig. 11. The ERP offline classification accuracy averaged on subjects using various trials. Vertical lines show maximal and minimal accuracies among all the subjects. The x-axis means how many trials were used.

the rest one as test data. This process was repeated 8 times, and each time the test data was a different run.

### B. NAO experiment

The performance of subjects and experiment details are shown in Table V. The time listed in the table does not include any behavior performing (e.g. dance) so that each session has the same expected time cost. For comparison and evaluation, the experiment was also completed by using Choregraphe, a GUI software platform for NAO, with a mouse. Operated by hand, the completion time of one session was about 49.1 s, which was averaged by four runs.

One session without any mistake should include nine commands (eight EOG, one ERP) which are sequentially: go ahead, stop, look left, look center, turn left, go ahead, stop, EEG mode, and behavior selection. So the minimal number of commands for one experiment is 32 EOG commands and 4 ERP commands. In Table V, the results are evaluated by: i) the total number of detected commands, ii) the time cost, iii) the ERP accuracy, and iv) the EOG accuracy. The

TABLE V  
RESULTS OF THE NAO EXPERIMENT

Subject	Session	No. of commands	Time(s)	Avg. time (s)	ERP accuracy	EOG accuracy
S1	1	12	62.9	70.2	4/4	35/39
	2	9	57.0			
	3	11	58.3			
	4	11	102.6			
S2	1	10	57.2	72.6	4/9	41/45
	2	9	52.0			
	3	11	59.1			
	4	24	122.2			
S3	1	14	78.3	67.5	4/6	35/40
	2	9	57.3			
	3	10	56.2			
	4	13	78.2			
S4	1	15	62.7	54.1	4/4	33/38
	2	9	52.2			
	3	9	52.5			
	4	9	49.0			

Note: The minimum number of commands for one session is 9. If operated by hand, the average time cost is 49.1 s.

ERP and EOG accuracy is expressed as “number of effective commands/number of detected commands”. For ERP, the number of effective commands is always equal to 4, because a complete NAO experiment contains 4 ERP commands. For EOG, the effective commands include the minimum necessary commands to complete the task (which is 32), plus the commands to correct earlier mistakes. For example, if the robot received a wrong command and did “look right” instead of “look left”, there should be an additional command ordering the robot to “look center” first. Although this command does not belong to the minimum necessary commands, it is still counted as an effective command. Additionally, in Table V for each subject, the denominator of “ERP accuracy” plus that of “EOG accuracy” means the total number of commands, so it should always be equal to the sum of “No. of commands”.

The EOG accuracy of online experiment shows the number of false detections, while the miss rate (false negative) can only be referred from the offline evaluation. This is because in the online experiment the EOG detection process was asynchronous: the subject could perform any eye movement at any time. So it is difficult to tell which eye movements were missed. However, the miss rate can also be estimated from the time cost. If the time cost was very long in spite of high EOG accuracy, it is very likely that there is relatively a larger number of misses.

### C. Kobuki experiment

Table VI shows the performance and details about the Kobuki experiment. Four robots were used in the experiment, and each was given seven commands: EEG mode, robot selection, EOG mode, go ahead, turn right, go ahead, and stop. Therefore the minimal number of commands for one experiment should be 28 (4 ERP and 24 EOG). The evaluation

TABLE VI  
RESULTS OF THE KOBUKI EXPERIMENT

Subject	Exp.	No. of commands	Time(s)	Avg. time (s)	ERP accuracy	EOG accuracy
S1	1st	29	123.7	124.9	8/9	52/56
	2nd	36	126.1			
S5	1st	45	190.7	172.5	8/10	52/76
	2nd	41	154.2			
S6	1st	35	160.9	157.5	8/9	52/62
	2nd	36	154.1			
S7	1st	51	212.0	221.5	8/9	56/103
	2nd	62	230.9			
S8	1st	37	147.5	136.5	8/9	52/59
	2nd	31	125.5			
S9	1st	34	137.3	140.8	8/10	50/61
	2nd	37	144.3			
S10	1st	37	147.0	141.8	8/9	51/65
	2nd	37	136.6			
S11	1st	35	136.4	140.3	8/8	51/67
	2nd	40	144.1			
S12	1st	39	141.1	134.2	8/10	55/63
	2nd	34	127.2			
S13	1st	34	138.0	141.6	8/9	50/60
	2nd	35	145.1			

Note: The minimum number of commands for one session is 28. If operated by hand, the average time cost is 122.7 s.

is similar to the NAO experiment, except that this experiment was not timed in separated sessions, and each subject completed this experiment twice.

For this experiment, we also calculated the hand-operated time for comparison. Because Kobuki has no corresponding GUI software platform, we created an operation menu in Simulink as its GUI which enables speed adjustment, direction adjustment, and robot selection. By using this operation menu, the whole experiment could be completed in about 122.7 s

## VII. DISCUSSIONS

### A. Performance

1) *EOG*: According to the results shown in Fig. 10, when compared to the pattern matching method, the proposed threshold method has a distinct advantage in detecting triple blinks, but remains below expectations in frown detection. Regarding wink and gaze detection, both methods have very high recall, but the proposed method performs better.

The main reason of the performance gaps between the proposed method and the pattern matching method is the duration of blink and frown. First, blinks are usually very quick, so a long blink and a short blink will have relatively large deviation. Because the pattern matching method checks the similarity between a blink candidate and the standard pattern, large deviation of blink candidates will lead to a high miss rate. For a triple blink, if the method misses at least one of the three blinks, the detection will fail. On the contrary, the threshold method does not have this problem as long as the blink candidates do not exceed the threshold.

The case of frown detection is just opposite. Generally a frown is more difficult to detect relative to other eye movements because it is usually close to a facial expression and is more affected by individual habits. But as its duration is usually very long, the deviation of frown candidates becomes small. In that case the pattern matching method will have some advantages. For the threshold method, we set the duration threshold to 2 s, so frowns longer than 2 s will have no chance to be detected. We have told subjects not to make very long frowns because that is not effective in doing a control experiment. However, some frowns were made longer than 2 s in the calibration process.

Although we have summarized the thresholds as in Table III, they need to be determined for individual subjects. A safe strategy is to manually adjust the thresholds for every subject, and in that way, the gap in frown detection can also be filled. However, only using pattern matching method to detect frowns, and using the proposed method to detect other eye movements could be a simpler solution.

2) *ERP*: The feature extraction based on P300 activity with high classification accuracies have already been discussed in many studies [43]–[47]. However, a P300-based BCI usually still requires several (more than five) trials average to obtain good performance. The inverted face paradigm adopted in our study used stimuli of facial images with loss of configural face information. The configural face perception has been demonstrated to evoke not only significantly stronger P300 but also VPP and N170 [13], [28], [48]. The outstanding effects of the inverted face paradigm can help the BCI to accurately decide the desired command with only two trials or even single trial [28]. This is why we chose the inverted face paradigm instead of the conventional P300 paradigm for our system development.

In fact, many studies have suggested other components, in addition to P300, may also contribute significantly to the target detection in an ERP-based BCI [13], [28], [49], [50]. The adopted inverted face paradigm exploits a joint effect of multiple ERP components (including N170, VPP and P300) for target detection, which has been confirmed by giving very good classification accuracy while using the ordinary LDA even in a small sample size scenario [28]. Whereas other algorithms such as SVM, regularization FDA, etc. could be a better choice for classification when only a low number of samples are available, they require a regularization parameter that is typically determined by the time-consuming cross-validation. Considering a trade-off between efficiency and accuracy, we adopted the simple LDA for classification in the developed BCI system.

3) *Online experiments*: In the NAO experiment, all the four subjects got a satisfactory result (see Table V), except for the 4th session of S1 and S2. Although S1 has a high EOG accuracy, but as mentioned above, the online accuracy does not show the miss rate (which is only shown in offline evaluation). According to the data, when S1 tried to make frowns in his 4th session, they were not detected. The data showed that his frowns in the 4th session became weaker than those he did in the calibration process, where the most possible reason might be tiredness. S2 failed the ERP command for 5 times

successively in his 4th session, but before that his ERP was 100% accurate. It is possible that the subject was distracted or tired at that time, nevertheless, such kind of successive failures never happened on other subjects. Apart from these two sessions, the experiment was successful. Especially, S4 gave an outstanding result, whose average completion time was 54.1 s, which means only 5 s longer than operating by hand. While he was unaccustomed to the experiment in his first session, all of his session 2, 3, and 4 were completed by minimal necessary commands and without any error.

Regarding the results of the Kobuki experiment (Table VI), most subjects had a stable performance except S5 and S7. Among all the 10 subjects, only S1 had participated in the NAO experiment and was familiar with BCI. This may be the reason that he had a clear advantage in terms of time. His average completion time was only 2.2 s longer than the hand-operated time. S5 had poor performance in his first attempt, he failed some winks, and confused double blinks and triple blinks many times. However, in his 2nd attempt he made a clear improvement where his completion time was reduced from 190 s to 154 s. In the case of S7, she spent obviously longer time than the other subjects. The reason for this was that she had some troubles in making winks. More specifically, her intentional winks were as weak as normal blinks, so the algorithm confused these two eye movements. As a result, she could not effectively make use of our hybrid HMI.

The proposed hybrid HMI received positive feedback from most subjects. The results show that the proposed hybrid HMI can be potentially applied to able-bodied people due to its good effectiveness. However, considering the fact that a few people might not be able to perform certain kinds of eye movements, more attention is needed for irregular eye movement detection.

## B. Practicality

In this study two online experiments were carried out: one using a multi-functional humanoid robot, the other using multiple mobile robots. On the basis of these two experiments, the proposed hybrid interface can also be expanded to applications requiring multiple controllable agents with multifunction. In that case, the ERP paradigm should include two ERP menus: one for target selection, and the other for behavior selection. The two ERP menus can be switched by ERP commands or eye movements.

The proposed control strategy is suitable for both direct control and remote control. Mounted with cameras, the robots are able to perform various tasks remotely, which can be applied for the case when the user has difficulty with mobility. On the other hand, the proposed method can also be used to assist the user's self-move. For example, nowadays there is a new concept called BCI smart home which is designed for assisting people with disabilities in their daily house life [51]. A smart BCI house usually includes a BCI wheelchair for moving between rooms, and some other BCI operations for using household electrical appliances like TV, lights, and so on. Compared to traditional EEG-based wheelchair which is relatively poor in mobility, using the proposed hybrid HMI, the user can move dexterously and freely, unrestricted by

predefined routes. Moreover, the EOG mode and ERP mode are supposed to work well together in this application. For example, we can imagine a scene such as a user moved into living room by EOG controlled wheelchair, then he chose to operate the TV from an ERP menu which listed all the operable targets in the living room. The TV turned on, and all the available channels appeared on the ERP menu. He chose a channel, made a wink to gradually turn up the volume, and winked again when he felt the volume was just good.

Compared with pure BCI or brain-based hybrid BCI such as SSVEP/P300, the proposed hybrid HMI has advantages in versatility and flexibility. Of course, since it takes eye movements as commands, those disabled persons who cannot control eye movements normally (e.g. serious facial paralysis) are unable to use this hybrid interface. Because most disabled people have preserved eye movement capabilities, the proposed hybrid HMI shows promise.

### C. Extensions

As mentioned above, the hybrid EOG/EEG interface requires the user to, at least, preserve normal function of making eye movements. However, even able-bodied people may have troubles in making certain kinds of eye movements. Our eye movement detection is implemented by using predefined thresholds which are only optimized for regular eye movements. Unfortunately, the proposed method still has some flaws when the eye movements are not regular, which can be seen from the result of the S7 in the Kobuki experiment. For this reason, powerful algorithms are needed to correctly recognize irregular eye movements. We have tried extracting standard eye-movement patterns from the subject's EOGs and doing online matching, but it was only advantageous in frown detection. Using classifiers can be a potential solution, except for two problems. First, EOG mode works asynchronously, so the training set of classifier will be largely biased where most data belongs to null class. Second, the training time is another important issue. Short training time affects the accuracy, while longer training time harms user experience. These difficulties need to be overcome to find a proper solution.

In aspect of ERP paradigm, the most important issue has always been to increase the information transfer rate. When the classification accuracy has been improved through using N170 and VPP together with P300, error report is another way to achieve the goal. If the classification error is reported in time, a second trial can be presented at once, or even the second probable result can be directly selected. As a result, the overall time consuming will be reduced. In this study, it has been proposed that eye movements can be used to report errors. However, there is another interesting method using error-related potential, which enables the system to automatically identify whether the subject considers the result correct or not [52], [53]. If this automatic error detection can achieve a high accuracy, the ERP paradigm will be more convenient to use.

## VIII. CONCLUSION

In this study, a novel EOG/EEG hybrid human-machine interface is proposed. The hybrid interface works in two

modes, where EOG mode detects eye movements including blink, frown, wink and gaze, and EEG mode adopts multi-component ERPs to judge the user's visual focus. The usefulness of the hybrid interface is verified by two online experiments about controlling different robots. In each experiment, most subjects have satisfactory performance, and a few subjects were even able to complete the tasks with high effectiveness comparable with hand operation. The experiment results suggest that there is a complimentary relationship between eye movements and ERP paradigm, and also suggest that the introduced hybrid interface is promising for BCI related applications. Future work will focus on further optimization of the eye movement detection method, and to improve its ability to detect irregular eye movements so that the proposed hybrid interface can better be applied to disabled persons.

## REFERENCES

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "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.-H. Lee, J. Ryu, F. A. Jolesz, Z.-H. Cho, and S.-S. Yoo, "Brain-machine interface via real-time fMRI: preliminary study on thought-controlled robotic arm," *Neuroscience Letters*, vol. 450, no. 1, pp. 1-6, 2009.
- [4] N. Naseer and K.-S. Hong, "Classification of functional near-infrared spectroscopy signals corresponding to the right-and left-wrist motor imagery for development of a brain-computer interface," *Neuroscience Letters*, vol. 553, pp. 84-89, 2013.
- [5] N. Naseer, M. J. Hong, and K.-S. Hong, "Online binary decision decoding using functional near-infrared spectroscopy for the development of brain-computer interface," *Experimental Brain Research*, vol. 232, no. 2, pp. 555-564, 2014.
- [6] J. Mellinger, G. Schalk, C. Braun, H. Preissl, W. Rosenstiel, N. Birbaumer, and A. Kübler, "An MEG-based brain-computer interface (BCI)," *Neuroimage*, vol. 36, no. 3, pp. 581-593, 2007.
- [7] 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," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 138-141, 2006.
- [8] G. Pfurtscheller and F. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842-1857, 1999.
- [9] Y. Li, J. Long, T. Yu, Z. Yu, C. Wang, H. Zhang, and C. Guan, "An EEG-based BCI system for 2-D cursor control by combining Mu/Beta rhythm and P300 potential," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 10, pp. 2495-2505, 2010.
- [10] 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*, vol. 6, no. 4, p. 046002, 2009.
- [11] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, "Single-trial analysis and classification of ERP components: a tutorial," *NeuroImage*, vol. 56, no. 2, pp. 814-825, 2011.
- [12] Y. Zhang, G. Zhou, Q. Zhao, J. Jin, X. Wang, and A. Cichocki, "Spatial-temporal discriminant analysis for ERP-based brain-computer interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, no. 2, pp. 233-243, 2013.
- [13] T. Kaufmann, S. Schulz, C. Grünzinger, and A. Kübler, "Flashing characters with famous faces improves ERP-based brain-computer interface performance," *Journal of Neural Engineering*, vol. 8, no. 5, p. 056016, 2011.
- [14] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor, "A spelling device for the paralysed," *Nature*, vol. 398, no. 6725, pp. 297-298, 1999.

- [15] S. Sutton, M. Braren, J. Zubin, and E. John, "Evoked-potential correlates of stimulus uncertainty," *Science*, vol. 150, no. 3700, pp. 1187–1188, 1965.
- [16] B. Allison, T. Luth, D. Valbuena, A. Teymourian, I. Volosyak, and A. Graser, "BCI demographics: How many (and what kinds of) people can use an SSVEP BCI?" *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 2, pp. 107–116, 2010.
- [17] C. Brunner, B. Allison, C. Altstätter, and C. Neuper, "A comparison of three brain-computer interfaces based on event-related desynchronization, steady state visual evoked potentials, or a hybrid approach using both signals," *Journal of Neural Engineering*, vol. 8, no. 2, p. 025010, 2011.
- [18] 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.
- [19] H. V. Semlitsch, P. Anderer, P. Schuster, and O. Presslich, "A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP," *Psychophysiology*, vol. 23, no. 6, pp. 695–703, 1986.
- [20] G. Gratton, M. G. Coles, and E. Donchin, "A new method for off-line removal of ocular artifact," *Electroencephalography and Clinical Neurophysiology*, vol. 55, no. 4, pp. 468–484, 1983.
- [21] R. Barea, L. Boquete, M. Mazo, and E. López, "System for assisted mobility using eye movements based on electrooculography," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 10, no. 4, pp. 209–218, 2002.
- [22] 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.
- [23] R. Barea, L. Boquete, M. Mazo, and E. López, "Wheelchair guidance strategies using EOG," *Journal of Intelligent and Robotic Systems*, vol. 34, no. 3, pp. 279–299, 2002.
- [24] R. C. Panicker, S. Puthusserypady, and Y. Sun, "An asynchronous P300 BCI with SSVEP-based control state detection," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1781–1788, 2011.
- [25] S. Amiri, A. Rabbi, L. Azinfar, and R. Fazel-Rezai, "A review of P300, SSVEP, and hybrid P300/SSVEP brain-computer interface systems," in *Brain-Computer Interface Systems—Recent Progress and Future Prospects*. InTech, 2013.
- [26] Y. Li, J. Pan, F. Wang, and Z. Yu, "A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 11, pp. 3156–3166, 2013.
- [27] E. Yin, Z. Zhou, J. Jiang, F. Chen, Y. Liu, and D. Hu, "A speedy hybrid BCI spelling approach combining P300 and SSVEP," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 2, pp. 473–483, 2014.
- [28] 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.
- [29] J. Ma, Y. Zhang, Y. Nam, A. Cichocki, and F. Matsuno, "EOG/ERP hybrid human-machine interface for robot control," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 859–864.
- [30] B. Allison, C. Brunner, V. Kaiser, G. Müller-Putz, C. Neuper, and G. Pfurtscheller, "Toward a hybrid brain-computer interface based on imagined movement and visual attention," *Journal of Neural Engineering*, vol. 7, no. 2, p. 026007, 2010.
- [31] G. Pfurtscheller, T. Solis-Escalante, R. Ortner, P. Linortner, and G. R. Müller-Putz, "Self-paced operation of an SSVEP-based orthosis with and without an imagery-based brain switch: a feasibility study towards a hybrid BCI," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 4, pp. 409–414, 2010.
- [32] A. Savić, U. Kisić, and M. Popović, "Toward a hybrid BCI for grasp rehabilitation," in *Proceedings of the 5th European Conference of the International Federation for Medical and Biological Engineering*, 2012, pp. 806–809.
- [33] J. Li, H. Ji, L. Cao, D. Zang, R. Gu, B. Xia, and Q. Wu, "Evaluation and application of a hybrid brain computer interface for real wheelchair parallel control with multi-degree of freedom," *International Journal of Neural Systems*, vol. 24, no. 04, 2014.
- [34] Y. Punsawad, Y. Wongsawat, and M. Parnichkun, "Hybrid EEG-EOG brain-computer interface system for practical machine control," in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2010, pp. 1360–1363.
- [35] S. Fazli, J. Mehnert, J. Steinbrink, G. Curio, A. Villringer, K.-R. Müller, and B. Blankertz, "Enhanced performance by a hybrid NIRS-EEG brain computer interface," *Neuroimage*, vol. 59, no. 1, pp. 519–529, 2012.
- [36] M. J. Khan, M. J. Hong, and K.-S. Hong, "Decoding of four movement directions using hybrid NIRS-EEG brain-computer interface," *Frontiers in Human Neuroscience*, vol. 8, 2014.
- [37] R. Leeb, H. Sagha, R. Chavarriaga, and J. del R Millán, "A hybrid brain-computer interface based on the fusion of electroencephalographic and electromyographic activities," *Journal of Neural Engineering*, vol. 8, no. 2, p. 025011, 2011.
- [38] B. Rebsamen, E. Burdet, Q. Zeng, H. Zhang, M. Ang, C. Teo, C. Guan, and C. Laugier, "Hybrid P300 and Mu-Beta brain computer interface to operate a brain controlled wheelchair," in *Proceedings of the 2nd International Convention on Rehabilitation Engineering & Assistive Technology*, 2008, pp. 51–55.
- [39] Y. Su, Y. Qi, J.-x. Luo, B. Wu, F. Yang, Y. Li, Y.-t. Zhuang, X.-x. Zheng, and W.-d. Chen, "A hybrid brain-computer interface control strategy in a virtual environment," *Journal of Zhejiang University SCIENCE C*, vol. 12, no. 5, pp. 351–361, 2011.
- [40] H. Riechmann, N. Hachmeister, H. Ritter, and A. Finke, "Asynchronous, parallel on-line classification of P300 and ERD for an efficient hybrid BCI," in *Proceedings of the 5th International IEEE/EMBS Conference on Neural Engineering*, 2011, pp. 412–415.
- [41] B. Jammes, H. Sharabty, and D. Esteve, "Automatic EOG analysis: A first step toward automatic drowsiness scoring during wake-sleep transitions," *Somnologie-Schlafforschung und Schlafmedizin*, vol. 12, no. 3, pp. 227–232, 2008.
- [42] A. Bulling, J. A. Ward, H. Gellersen, and G. Troster, "Eye movement analysis for activity recognition using electrooculography," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 741–753, 2011.
- [43] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayouhd, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "A comparison of classification techniques for the P300 speller," *Journal of Neural Engineering*, vol. 3, no. 4, p. 299, 2006.
- [44] V. Abootalebi, M. H. Moradi, and M. A. Khalilzadeh, "A new approach for EEG feature extraction in P300-based lie detection," *Computer Methods and Programs in Biomedicine*, vol. 94, no. 1, pp. 48–57, 2009.
- [45] A. Turnip, K.-S. Hong, and M.-Y. Jeong, "Real-time feature extraction of P300 component using adaptive nonlinear principal component analysis," *Biomedical Engineering Online*, vol. 10, no. 1, p. 83, 2011.
- [46] J. N. Mak, D. J. McFarland, T. M. Vaughan, L. M. McCane, P. Z. Tsui, D. J. Zeililin, E. W. Sellers, and J. R. Wolpaw, "EEG correlates of P300-based brain-computer interface (BCI) performance in people with amyotrophic lateral sclerosis," *Journal of Neural Engineering*, vol. 9, no. 2, p. 026014, 2012.
- [47] A. Turnip and K.-S. Hong, "Classifying mental activities from EEG-P300 signals using adaptive neural network," *International Journal of Innovative Computing Information and Control*, vol. 8, no. 9, pp. 6429–6443, 2012.
- [48] R. J. Itier, M. Latinus, and M. J. Taylor, "Face, eye and object early processing: what is the face specificity?" *NeuroImage*, vol. 29, no. 2, pp. 667–676, 2006.
- [49] B. Z. Allison and J. A. Pineda, "ERPs evoked by different matrix sizes: implications for a brain computer interface (bci) system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 110–113, 2003.
- [50] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient p300-based brain-computer interface for disabled subjects," *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 115–125, 2008.
- [51] G. Edlinger, C. Holzner, and C. Guger, "A hybrid brain-computer interface for smart home control," in *Human-Computer Interaction. Interaction Techniques and Environments*. Springer, 2011, pp. 417–426.
- [52] A. Combaz, N. Chumerin, N. V. Manyakov, A. Robben, J. A. Suykens, and M. M. Van Hulle, "Towards the detection of error-related potentials and its integration in the context of a P300 speller brain-computer interface," *Neurocomputing*, vol. 80, pp. 73–82, 2012.
- [53] M. Spüler, M. Bensch, S. Kleih, W. Rosenstiel, M. Bogdan, and A. Kübler, "Online use of error-related potentials in healthy users and people with severe motor impairment increases performance of a P300-BCI," *Clinical Neurophysiology*, vol. 123, no. 7, pp. 1328–1337, 2012.



**Jiaxin Ma** received B.S. and M.S. degrees in computer science from Shanghai Jiao Tong University, China, in 2008 and 2011, respectively, and also a M.S. degree in engineering from Waseda University, Japan, in 2009. He is currently a Ph.D. candidate in the Department of Mechanical Engineering and Science at Kyoto University, Japan. He was a visiting scholar in the Department of Biomedical Engineering at the Johns Hopkins University, US, in 2012. His current research interests

include brain-computer interface, electroencephalogram, electrooculogram, and electromyogram signal processing and their applications.



**Yu Zhang** received the Ph.D. degree in control science and engineering from the School of Information Science and Engineering, East China University of Science and Technology, Shanghai, China, in 2013. He worked as an International Program Associate (IPA) for two years (from Dec. 2010 to Nov. 2012) in the Laboratory for Advanced Brain Signal Processing (LABSP) at RIKEN Brain Science Institute, Wako-shi, Japan. He is currently working as an Assistant Professor at the School of Information Science and Engineering, East China University

of Science and Technology, Shanghai, China. His research interests include brain-computer interface, signal processing, tensor analysis, machine learning and pattern recognition.



**Andrzej Cichocki** received the M.Sc. (with Hons.), Ph.D., and Dr.Sc. (Habilitation) degrees, all in electrical engineering, from Warsaw University of Technology, Warsaw, Poland. Since 1972, he has been with the Institute of Theory of Electrical Engineering, Measurement and Information Systems, Faculty of Electrical Engineering at the Warsaw University of Technology, where he received the title of a Full Professor in 1995. He spent several years at the University Erlangen-Nuerenberg, Germany, at the Chair of Applied and Theoretical Electrical

Engineering directed by Prof. R. Unbehauen, as an Alexander-von-Humboldt Research Fellow and Guest Professor. From 1995 to 1997, he was a team leader of the laboratory for Artificial Brain Systems, at the Frontier Research Program RIKEN, Japan, in the Brain Information Processing Group. He is currently the head of the laboratory for Advanced Brain Signal Processing, at RIKEN Brain Science Institute, Wako-shi, Japan. He is author of more than 250 technical papers and four monographs (two of which have been translated to Chinese). His research interests include signal processing, inverse problems, neural network and learning algorithms, tensor analysis, and brain-computer interface.



**Fumitoshi Matsuno** received the Ph.D. (Dr. Eng.) degree from Osaka University in 1986. In 1986 he joined the Department of Control Engineering, Osaka University. He became a Lecturer in 1991 and an Associate Professor in 1992, in the Department of Systems Engineering, Kobe University. In 1996 he joined the Department of Computational Intelligence and Systems Science, Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology as an Associate Professor.

In 2003 he became a Professor in the Department of Mechanical Engineering and Intelligent Systems, University of Electro-Communications. Since 2009, he has been a Professor in the Department of Mechanical Engineering and Science, Kyoto University. He holds also a post of the Vice-President of NPO International Rescue System Institute (IRS). His current research interests lie in robotics, swarm intelligence, control of distributed parameter system and nonlinear system, and rescue support system in disaster. Dr. Matsuno received many awards including the Outstanding Paper Award in 2001 and 2006, Takeda Memorial Prize in 2001 from the Society of Instrument and Control Engineers (SICE), and the Best Paper Award in 2013 from Information Processing Society of Japan. He is a Fellow member of the SICE, the JSME, and a member of the IEEE, the RSJ, the ISICE, among other organizations. He served as a co-chair of IEEE RAS Technical Committee on Safety, Security, and Rescue Robotics, a chair of Steering Committee of SICE Annual Conference, a General Chair of IEEE SSR2011 and IEEE/SICE SII2011 etc. He is an Editor-in-Chief of Journal of RSJ, an Editor of Journal of Intelligent and Robotic Systems, an Associate Editor of Advanced Robotics, International Journal of Control, Automation, and Systems, etc. and on the Conf. Editorial Board of IEEE CSS.