IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 27, 2018, accepted January 21, 2019, date of publication February 8, 2019, date of current version March 1, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2898324

An Electrooculogram-Based Interaction Method and Its Music-on-Demand Application in a Virtual Reality Environment

JING XIAO[1,2], JUN QU[1,2], AND YUANQING LI[1,2], (Fellow, IEEE)

[1]School of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China
[2]Guangzhou Key Laboratory of Brain Computer Interface and Applications, South China University of Technology, Guangzhou 510640, China

Corresponding author: Yuanqing Li (auyqli@scut.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFB1002505, in part by the National Natural Science Foundation of China under Grant 61633010, and in part by the Guangdong Natural Science Foundation under Grant 2014A030312005.

ABSTRACT Recently, the devices commonly used for interaction with a virtual reality (VR) environment include game controllers, data gloves, and motion tracking systems. However, these devices may limit the immersive experience or cause inconvenience, especially for patients with limb paralysis. This paper proposes a new nonmanual human–computer interface (HCI) based on a single-channel electrooculogram (EOG) signal and enables real-time interactions with the VR environment. The graphical user interface of the EOG-based HCI in VR includes several buttons flashing in a random order. The user needs to blink in synchrony with the corresponding button's flashes to issue a command, while the algorithm detects the eye blinks from the EOG signal and determines the users' target button. Furthermore, with the EOG-based HCI, we developed a music-on-demand system in the VR environment. By blinking, users can search and select target songs in the music library, play the selected songs, switch to the next song, and delete songs from the list. Ten healthy subjects and five patients with spinal cord injuries (SCIs) participated in online experiments. The experimental results demonstrated the effectiveness of the EOG-based HCI as a new method for interacting with the VR environment. Furthermore, the music-on-demand system can be used for entertainment not only by healthy people but also by SCI patients with limb paralysis.

INDEX TERMS Electrooculogram (EOG), virtual reality (VR), human–computer interface (HCI), eye blink, music-on-demand.

I. INTRODUCTION

Virtual reality (VR) is an interactive technology in the virtual world [1], which provides users with an immersive 3D simulation environment [2]. VR has been widely used in various areas, such as education, medicine, the military and video games [3]. In a normal VR application, users generally generate control commands by using external physical devices, including game handle, data gloves, and motion tracking systems [4], [5]. As users are engaged in a VR environment, those external physical devices often limit the immersive experience or cause inconvenience [6], [7]. Barrier-free and nonmanual human-computer interfaces (HCIs) for remote control of a VR environment have received a great deal of attention [8], [9].

The associate editor coordinating the review of this manuscript and approving it for publication was Bora Onat.

Electroencephalogram (EEG)-based brain-computer interfaces (BCIs) form a class of nonmanual HCIs that detect brain patterns in EEG signals and translate them into commands for external devices, such as wheelchairs, robots and spellers. These brain patterns include steady-state visually evoked potentials (SSVEPs) [10], P300 potentials [11], and motor imagery (MI)-related mu/beta rhythms [12]. In recent years, several BCIs have been developed for VR control. For example, a BCI combined with VR was used to assess and train hemiparetic upper limbs in patients with chronic stroke [13]. A rehabilitation game system based on VR was proposed for patients with stroke; in this system, a BCI was used for interaction with the VR environment [14]. In [15], a VR system with a BCI was presented for the rehabilitation of the upper limbs of patients with cervical spinal cord injuries (SCIs), and its effectiveness was demonstrated by comparison with traditional methods. However, BCIs designed for interaction with VR environments have several limitations, including the inconvenience of wet electrodes, the poor signal quality of dry electrodes, and the relatively low information transfer rate (ITR) for an asynchronous BCI [16].

EOG-based HCIs can allow users to issue various commands by monitoring their eye movements, such as blinking, fixating, and looking up/down/left/right [17]. Compared with EEG-based BCIs, the advantages of EOG-based HCIs include the following: (i) eye movements are natural activities that generally do not cause discomfort to users [18]; (ii) it is easier to detect eye movements in EOG signals than to detect brain patterns in EEG signals [19]; and (iii) fewer channels are used in EOG-based HCIs. Several EOG-based HCIs have been developed to operate various types of applications [20]. For instance, Deng *et al.* [21] implemented an EOG-based HCI that detected horizontal and vertical directions of eye movements from EOG signals and designed several applications, such as TV remote controls, games and vision tests. Ma *et al.* [22] proposed an EOG- and EEG-based HCI to control a robot, where P300 was used to generate commands and the blink-based EOG was used to verify these commands. Huang *et al.* [23] presented an EOG-based HCI for wheelchair control that could quickly and accurately generate control commands for left and right turns, forward and backward movements, acceleration, deceleration, and stopping. He *et al.* [24] developed a single-channel EOG-based speller that allowed users to input characters by blinking their eyes. These EOG-based systems achieved satisfactory performance but had not been used for interaction with VR. Only in [25], Kumar and Sharma tried to design an HCI for a VR game using EOG signals. The online VR game had 3 target trains for selection, and the average accuracy of 10 participants (8 males and 2 females) was 80%. Thus far, there has been no effective EOG-based method proposed for the interaction with VR.

In summary, the existing VR interaction methods have some shortcomings. First of all, the traditional VR interaction method requires some external equipment, which imposes an extra burden on the user, especially for patients with hand and foot inconvenience or the elderly. For example, the handle must be carried by the user and the tracker needs to be tracked; likewise, there is a need to track the locator for gesture recognition and the gesture types are limited, resulting in an insufficient number of control commands. Second, the existing non-manual HCIs often fails to achieve satisfactory performance. Therefore, This study first presents a nonmanual EOG-based HCI for interaction with a VR system. We design a graphical user interface (GUI) in VR that includes several buttons flashing in a random order. Each button corresponds to a specific command. The user can issue a command by blinking in synchrony with the corresponding buttons flashes. The HCI algorithm detects the eye blinks from the recorded EOG signal to determine the users target button. Furthermore, a music-on-demand system is developed on the basis of the EOG-based HCI. By blinking to use this system, users can search and select target songs in the music library, play the selected songs, switch to the next song, delete the current songs in the list, etc. Online experiments involving 8 healthy subjects and 5 SCI patients were conducted and showed that the healthy subjects and patients could quickly and accurately complete all operations, thus demonstrated the effectiveness of our EOG-based interaction method and music-on-demand system. This study not only provides a reliable VR interaction method for patients with limb paralysis, but also assists the interaction-related research in VR field.

The remainder of this paper is organized as follows: Section II describes the EOG-based HCI, including the system structure, GUI, control method, data processing algorithm, and music-on-demand system. The experimental procedure and results are presented in Section III. Further discussion is provided in Section IV, and Section V concludes this study.

II. METHODS

In this section, we first describe the EOG-based HCI for interaction with a VR environment and then present a music-on-demand system based on the HCI.

A. EOG-BASED HCI

As shown in Fig. 1(a), the EOG-based HCI is composed of an EOG signal acquisition device, a central control unit (CCU) and a VR system. The CCU is realized on a laptop, whereas the stimulus and execution interfaces are presented in the VR system and were developed using Unity software. As the most popular game engine, Unity software is often utilized to create video game content, including scenes, game objects, and components in VR. The detailed structures of the system are as follows:

1) COMMUNICATION

The data transmission between the data acquisition device and CCU is based on the Bluetooth communication protocol, while the data transmission between CCU and VR system is based on the WIFI protocol. Furthermore, we have developed an interactive software development kit (SDK) to connect to VR system and the CCU, which provides us a user datagram protocol (UDP) network connection to read and write data streams such that allows the external client can remotely control the objects in VR. Specifically, the SDK provides a communication interface for CCU and Unity, which allows the CCU to control the objects or components in the Unity. The CCU processes the incoming EOG signals to obtain control commands, generates instructions for the GUI and send them to the VR system.

2) EOG SIGNAL ACQUISITION DEVICE

As shown in Fig. 1(b), the EOG signal acquisition device is a self-designed hardware circuit that records and amplifies the EOG signal. The main modules include a preamplifier, a right leg drive circuit (RLD), a bandpass filter (1-30 Hz), a microcontroller unit (MCU), and a Bluetooth communication module. The RLD is used to reduce common-mode interference.

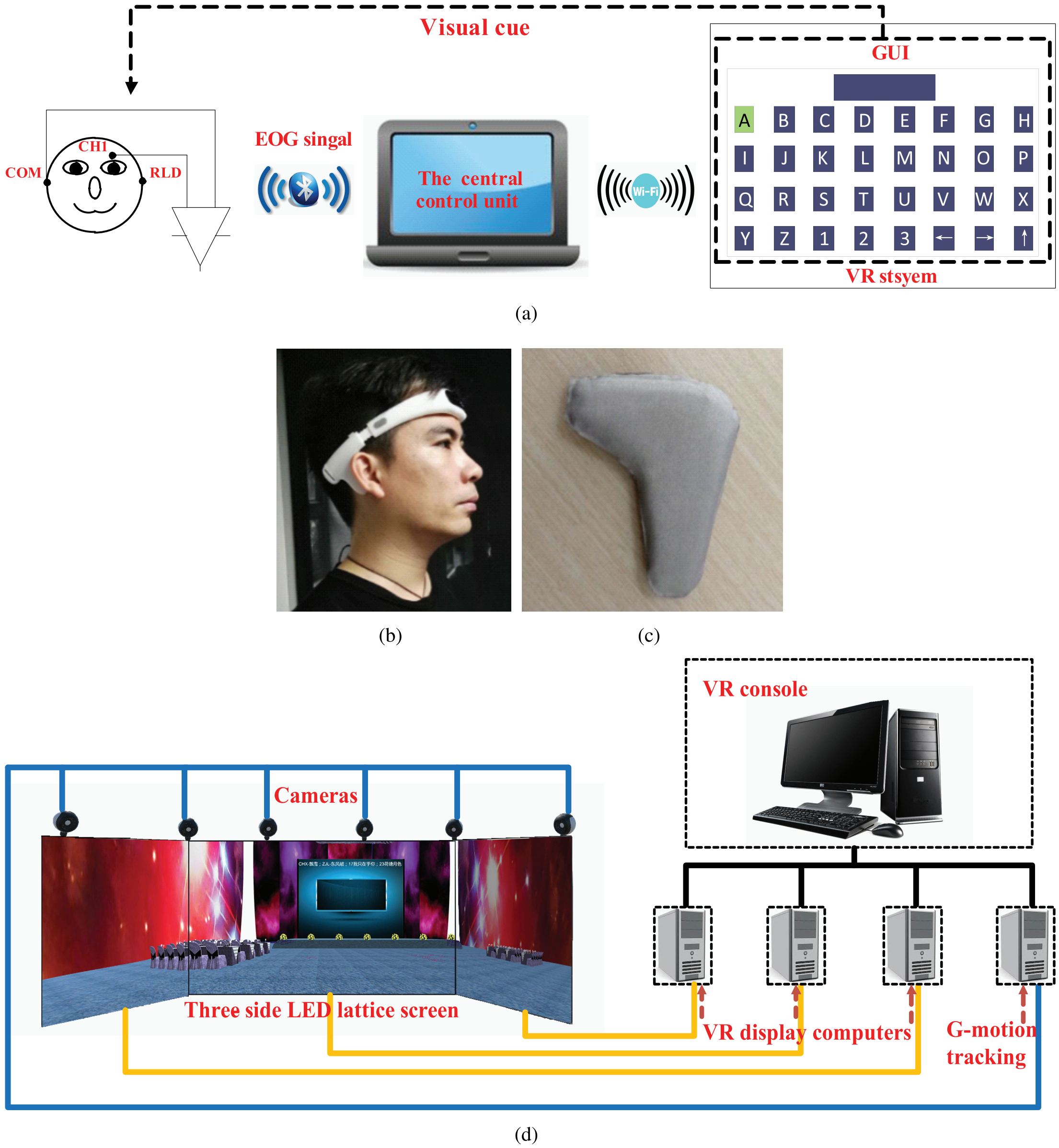


FIGURE 1. The EOG-based system: (a) The EOG-based HCI is composed of an EOG data acquisition device, a central control unit (CCU) in laptop, and a VR system. The signal terminal of the EOG data acquisition device consists of three electrodes: "COM", which is located under the left ear, is the reference electrode; "CH1", which is located above the right eyebrow, is used to record the EOG; and "RLD", which is located under the right ear, provides the common-mode signal feedback end of the human body. (b) Wearing diagram of the wearable EOG signal acquisition device. (c) Flexible electrode developed by the Shulin team of South China University of Technology. (d) The VR system is composed of a console, a visual tracking device and display devices.

The overall gain of the amplifier is approximately 1500, and the sampling rate is 250 Hz. The circuit board for the hardware circuit is small (2×2.7 cm). The signal terminal consists of three flexible dry electrodes ("COM", "CH1", and "RLD"), which are made of metal wire fabric, as shown in Fig. 1(c). Specifically, "COM", which is located under the left ear, is the reference electrode; "CH1", which is located above the right eyebrow, is used for EOG recording; and "RLD", which is located under the right ear, provides the common-mode signal feedback end of the human body.

3) VR SYSTEM

The VR system is provided by the Man Heng company in Shanghai, China, and is composed of a console, a visual tracking device and display devices, as shown in Fig. 1(d). The console controls the selection/operation/stopping of the VR system and the preview/adjustment/replacement of the VR scene. The visual tracking device is called "G-motion tracking", which tracks the body posture of the user in real time to ensure VR immersion. The display device consists of three computers and a three-side LED lattice screen including a main screen and two auxiliary screens with an angle of 120 degrees between each two adjacent screens; the main and auxiliary screens are 5 × 2.8 and 2.2 × 2.8 meters, respectively.

4) GUI AND EXECUTION

As shown in Fig. 2, the GUI includes a 4×8 simulation matrix containing 32 buttons and a text box for the user to provide entries. When the GUI is running, all buttons are activated with alternately flashing blue/green in a random order, which is similar to the P300-based scintillation method described in [26]. The user can select the target button by synchronous blinking. In this case, the control state indicates that the user is currently performing blinking operations in accordance with a flashing target button, while the idle state indicates that the user does not want to perform any action. The duration of each button flicker is 100 ms, with a stimulus onset asynchrony (SOA) of 60 ms between the onsets of two adjacent flashes (there is an overlap of 40 ms between two adjacent flashes). A round of flashing is defined as a cycle in which each button flashes once. The proposed system works in an asynchronous mode. Specifically, in a synchronous mode, the program provides a time cue for the user to perform each operation, and the buttons simultaneously begin to flash [24]. In an asynchronous mode, each operation is initiated by the user (self-paced), and the flashes of the buttons are independent of the users state (control or idle state).

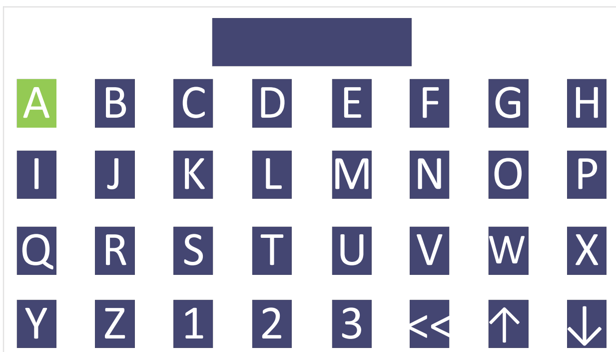


FIGURE 2. GUI of the EOG-based HCI, which includes a 4 × 8 simulation matrix containing 32 buttons and a text box.

B. DETECTION ALGORITHM

When the buttons of the GUI flash, the user needs to blink in synchrony with the flashes of the target button to issue a command. We perform blink detection and decision making to select a button in each round of button flashes (in a round, each button flashes once in a random order). When the same button is selected in two consecutive rounds of button flashes, then this button is determined to be the target, and a corresponding command is issued. More details are presented in the following.

1) BLINK DETECTION

For each round of button flashes, we first perform blink detection. Specifically, a feature vector corresponding to each

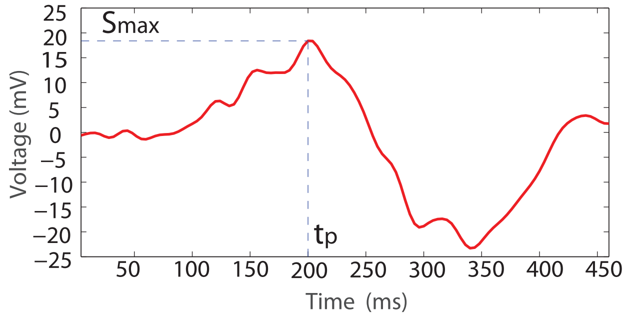


FIGURE 3. A differential EOG waveform, where t_p is the delay between the flash start and the peak of the waveform and S_{max} is the maximum amplitude of the differential waveform.

button flash is extracted as follows. A segment of the EOG signal (150-550 ms after the onset of a flash) is extracted and bandpass filtered in the 0.1-5 Hz range to eliminate the effects of baseline wander and high frequency noise. Next, the differential waveform is obtained through a first-order difference calculation, as shown in Fig. 3. Third, the following features are extracted: the peak, denoted S_{max}, of the differential waveform signal; and the delay, denoted t_p, between the starting time of this flash and the time point corresponding to the peak of the waveform. The blink detection based on the differential waveform of the $i-th$ buttons flash is based on the following criterion:

$$\begin{cases} 1, & S_{max} \geq S_{on}, t_{min} \leq t_p \leq t_{max} \\ 0, & otherwise \end{cases} \tag{1}$$

where S_{on} is an amplitude parameter, and the parameters t_{min} and t_{max} represent the range of the delay. These parameters are determined in a calibration session.

A blink is detected corresponding to the $i-th$ button if $b_i = 1$; otherwise, no blink is detected.

2) DECISION

At the end of each round of flashes, a decision is performed. Specifically, if $b_i = 0$ for all i, no candidate target button is recognized. Otherwise, the button i with a nonzero bi is identified as a candidate target button. When than one candidate targets are identified in a round, we calculate the following evaluation value e:

$$e = |t_p - T_p| \tag{2}$$

where T_p denotes the average delay determined through a calibration process. The button with the minimal evaluation value is selected for this round.

Furthermore, if a button is selected twice in this and previous rounds of button flashes (this selection implies that the detected blinks are synchronized with the buttons flashes), this button is determined to be a target, the corresponding command is issued, and the flashes of all buttons of the GUI are paused. After a predefined pause time (for example, 1000 ms), a new round of flashes starts again to generate a new command. If no target is detected in this round, the buttons in the GUI continue to flash.

3) CALIBRATION

For a new subject, a calibration is performed before the online experiment. Specifically, before the test, one button ("TRAIN") appears at the center of the three-side LED screen and flashes once every 1200 ms. The user is required to blink ten times according to the flashes of the "TRAIN" button. Using the recorded EOG signal, we obtain the peak value S_{max} and the delay t_p based on the differential EOG waveform of each flash (refer to the description of feature extraction in the detection algorithm). The average values of S_{max} and t_p for all flashes are taken as the threshold S_{on} and the average delay T_p, respectively. Furthermore, the duration thresholds t_{min} and t_{max} are obtained by multiplying the average value of t_p by empirical factors in this study (e.g., $t_{min} = 0.8 \times T_p$ and $t_{max} = 1.2 \times T_p$).

C. THE MUSIC-ON-DEMAND SYSTEM

As shown in Fig. 4, we first constructed a virtual concert hall environment of approximately 2000 square meters using 3Ds Max and Unity. The environment includes a stage, a large projection screen and audience seats, in which users can experience Karaoke Television (KTV) or other music programs in a highly realistic simulation. Next, we design a music-on-demand system similar to KTV, by which the user can search and select the target songs in the music library by blinking and then play the songs. As illustrated in Fig. 5, the system includes a GUI tree structure that is composed of seven sub-GUIs labeled sub-GUI0 - sub-GUI6. The user performs the operations in the music-on-demand system using the EOG-based interaction method previously described. Specifically, each sub-GUI contains several blue buttons highlighted in green (i.e., a flash). The buttons and

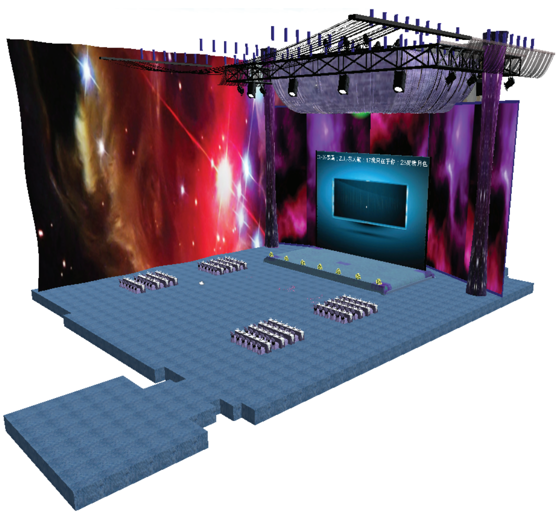


FIGURE 4. Virtual concert hall of the music-on-demand system created by 3Ds Max.

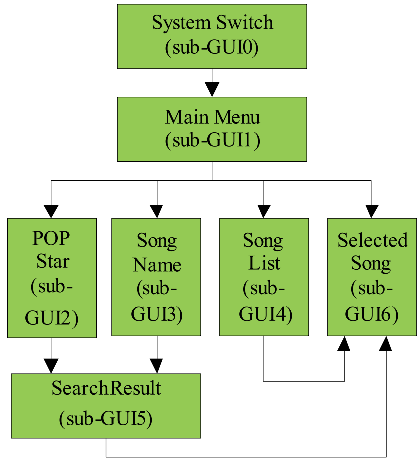


FIGURE 5. The GUI tree structure of the music-on-demand system, with seven sub-GUIs: system switch (sub-GUI0), main menu (sub-GUI1), song search based on POP star (sub-GUI2), search based on song name (sub-GUI3), song list (sub-GUI4), search result (sub-GUI5) from sub-GUI2 or sub-GUI3, and the list of songs that have been selected by the user (sub-GUI6).

the corresponding functions for each sub-GUI are shown in Fig. 6. The user can issue a command by blinking in synchrony with the corresponding buttons flashes, as shown in Fig. 6(d). In the following, we describe these sub-GUIs and the corresponding operations.

(i) Fig. 6(a) (sub-GUI0) is the switch button, which is the initial sub-GUI of the system. There is a system locking mechanism as below: (1) if the user want enter the main menu, he/her needs to open "ON1" and "ON2" switches in turn. (2) When the "ON1" switch is turned on, if the "ON2" switch is not turned on in 10 seconds, the sub-GUI will automatically return the "ON1" switch. (3) When the user is not focused on the flicker button (i.e., in the idle state) for more than 120s, the system returns the initial sub-GUI and is locked to avoid accidental operation.

(ii) Fig. 6(b) (sub-GUI1) is the main menu, consisting of two columns with 12 buttons. From the upper left corner, "POP Star"/"Song Name" is a way to search and select target songs, allowing the user first spells the singer name/song name (short for singer name, such as singer "Celine Dion", which is abbreviated as "CD"; short for song name, such as song "Big Big World", which is abbreviated as "BBW") in the search criteria input box, then all singers/ songs in the music library that meet the conditions are retrieved and displayed by search engine, and finally the user selects the target song. The "Song List" is the third way to do the song, which shows all the songs for the music library in alphabetical order, and the user can turn the page and select the target Song. "Selected Song" shows the list of songs waiting for play that the user has selected. "Original Sing" is a toggle button for the original or accompaniment when playing a song. "Play" finishes playing the current song in "Selected Song". "Pause" performs the suspension of the current song. "Delete" is the delete button where users select

FIGURE 6. **The four types of sub-GUIs in the music-on-demand system: (a) system switch (sub-GUI0). (b) main menu (sub-GUI1). (c) song list menu (sub-GUI4, sub-GUI5 and sub-GUI6). (d) song search menu (sub-GUI2 and sub-GUI3).**

a song from the "Selected Song" list, then deleted target song. "Switch Song" is a song switching button, when this button is selected, the next song will be played regardless of whether the current song is playing. "Priority" is a priority button; which is selected a target song from the "Selected Song" list and performed a priority operation. The playback order of target song in the "Selected Song" list has the highest priority, while other songs are delayed. "Hide Menu" hides the main menu, then the system return to the initial sub-GUI ("ON1").

(iii) Fig. 6(c) (sub-GUI4, sub-GUI5 and sub-GUI6) is a sub-GUI for displaying song information, which is used to display the search results or the selected songs. This sub-GUI has 16 buttons, wherein the left 8 buttons perform a blinking operation, and the right 8 buttons display corresponding information. Specifically, "≪" / "≫" performs to the previous/next page of the song list; "↑" is selected to return back the main menu. The song information display the song number, song name, and singer name. The specified song (including the song name and artist name) can be selected by selecting the song number.

(iv) Fig. 6(d) (sub-GUI2 and sub-GUI3) is a sub-GUI for search consisting with 32 buttons which contains 26 English letters and 1-3 numbers, as well as 3 function keys("≪", "↓", "↑"). The "≪" button is selected to delete the output error character; "↓" button is selected to return back the main menu; "↑" button is selected to activate music library search engine. At the top of the screen, the currently entered characters are displayed in the search text box.

The sub-GUIs are displayed on a three-side LED screen with a resolution of 1920×1080 and a refresh rate of 120 Hz. As the SOA is set to a small duration, a particular flicker may be associated with several adjacent buttons. To reduce the interference of the temporally neighboring buttons, we design the following rules: the sub-GUIs such as sub-GUI1 ÍC sub-GUI6, which have large buttons, are alternately flashing in random order; and sub-GUI0 flash in a fixed sequence. If the music-on-demand system outputs a command, a red box appears around the target button, as shown in Fig. 6(d). As described above, the users can select the corresponding buttons of seven sub-GUIs by blinking to issue a sequence of commands to perform the operations of music-on-demand application.

III. EXPERIMENTS AND RESULTS

In this study, two online experiments were conducted in the VR environment to test the performance of our EOG-based interaction method and music-on-demand system. Eight healthy subjects with normal or corrected-to-normal vision first participated in Experiment 1. Next, five patients with SCIs from the Guangdong Provincial Work Injury Rehabilitation Hospital participated in Experiment 2. In these experiments, the subjects operated the music-on-demand system using the EOG-based interaction method. This study was approved by the Ethics Committee of the Guangdong Provincial Rehabilitation Hospital, Guangzhou, China, and conformed to the Code of Ethics of the World Medical Association. Written informed consent for the experiments and

TABLE 1. Results of experiment 1 for eight healthy subjects.

Healthy	Time (s)	Operations	Command ratio (%)	Accuracy (%)	RT (s)	FPR (events/min)	FOR (events/min)
H1	429.1	52	104	98.1	5.3	0	0
H2	442.9	59	118	89.8	4.5	0.05	0
H3	430.4	57	114	94.7	4.6	0.1	0
H4	387.0	54	108	94.4	4.0	0	0
H5	336.1	50	100	100.0	3.8	0.05	0
H6	359.8	53	106	94.3	3.9	0.1	0
H7	361.5	55	110	92.7	3.5	0	0
H8	343.5	51	102	98.0	3.8	0.15	0
Avg. ± SD	386.29 ± 39.76	53.88 ± 2.85	107.70 ± 5.70	95.25 ± 3.09	4.18 ± 0.55	0.06 ± 0.05	0 ± 0

the publication of individual information were obtained from each healthy subject and each SCI patients legal surrogate.

A. PERFORMANCE INDICES

The following performance indices were used to evaluate the proposed method and system:

(i) Time required to implement all tasks of an experiment;

(ii) Number of operations generated in an experiment;

(iii) Operation ratio, the ratio of the number of commands actually issued in an experiment to the optimal number of commands required for all tasks in the experiment;

(iv) Accuracy, the ratio of the number of correctly selected buttons to the total number of selected buttons;

(v) Response time (RT), the time required to generate a command;

(vi) False positive rate (FPR): The number of commands incorrectly issued per minute when the subject was in the idle state;

(vii) False operation rate (FOR): the number of erroneous operations that occurred per minute when the subject was in the idle state.

B. ONLINE EXPERIMENTS

1) EXPERIMENT I: FOR HEALTHY SUBJECTS

Eight healthy subjects (2 females and 6 males aged between 23 and 39) participated in this experiment to test the effectiveness of our system. For each subject, the experiment tasks included the following sequence of operations:

(i) Calibration to determine the system parameters.

(ii) Opening the main menu (2 operations).

(iii) Search for four songs through the "POP Star"/"Song Name" sub-GUI, which includes the following operation sequence: entering the search sub-GUI (sub-GUI2/ sub-GUI3) and typing the search criteria (short names of singer/song name); selecting the target song in the search results (sub-GUI5) such that this song is included in the list of "Selected Song" (28 operations).

(iv) Select two target songs in the "Song List", which includes the following operations: entering the music database, where the preselected songs are displayed in alphabetical order (sub-GUI4); selecting the target song in the list and putting them into the "Selected Song" list (6 operations).

(v) Test the function buttons of "Main Menu" (sub-GUI1), which includes the following operations: "Original Sing"/"Accompaniment Sing" operation; "Delete" operation; "Prioritize" operation; "Play" operation; "Pause" operation; "Switching Song" operation; and "Hidden Menu" operation (14 operations).

After the above experimental tasks were performed, the subject was asked to remain in an idle state for 30 minutes without controlling the system (the subject could relax, listen to songs, sing, or chat with others). At the beginning of the idle state, the system was in the initial state (off state). If the switch "ON1" was selected, then this selection was counted as a false positive (FP). If two switch keys ("ON1" and "ON2") were continuously selected within a time interval of 10 s, the main menu (sub-GUI1) was opened, and a false operation (FO) was counted. If the main menu was wrongly activated, a selection of any button was counted as an FP, and an operation to the system was counted as an FO (Note: an operation might involve more than one button selection sometimes).

Table 1 summarizes the results of Experiment 1. Specifically, for all subjects, 386.29±39.76 s was taken to complete all experimental tasks. The average RT to issue a command was 4.18 ± 0.55 s, and the average accuracy was 95.25 ± 3.09%. In addition, each subject sent 53.88±2.85 commands to perform all the experimental tasks on average, which theoretically required 50 commands. Therefore, the average ratio between the actual number of commands to complete all tasks and the optimal number of commands was 107.75 ± 5.70%. These additional EOG commands were generated because if an error command was issued, the subject needed to correct the error. Considering the idle state, the average FPR for all subjects was 0.06 ± 0.05/min. However, these FPs did not lead to any false operations; thus, the FOR of each healthy subject was 0/min. The zero FOR during the idle state was due to our design: (i) a switch mechanism that would switch the system to the off state if there were no operations in a time interval of 120 s; (ii) a verification mechanism that turned on the system only when the two keys "ON1" and "ON2" were continuously selected within 10 s.

2) EXPERIMENT II: FOR PATIENTS

Five patients with SCIs from the Guangdong Provincial Work Injury and Rehabilitation Hospital, Guangzhou, China, participated in this experiment to illustrate the potential

TABLE 2. Clinic information of five sci patients.

Patient	Gender	Age	Time of onset (month)	Pathogenesis	ASIA
P1	male	27	60	traffic accident	C3-ASIA-A
P2	male	33	5	fall accident	C4-ASIA-A
P3	male	26	24	traffic accident	C4-ASIA-A
P4	Female	28	4	Myelitis	C1-ASIA-C
P5	male	38	8	traffic accident	C3-ASIA-A

ASIA: American Spinal Injury Association. Details of ASIA impairment scales can be found in [32].

TABLE 3. Results of experiment 2 for five sci patients.

Patients	Time (s)	Operations	Command ratio (%)	Accuracy (%)	RT (s)	FPR (events/min)	FOR (events/min)
P1	275.0	35	106	97.1	4.9	0	0
P2	289.2	35	106	94.4	5.4	0.1	0
P3	239.9	38	115	94.7	3.4	0.15	0
P4	281.8	39	118	92.3	4.3	0.1	0
P5	294.7	40	121	90.0	4.4	0	0
Avg. ± SD	276.12 ± 19.30	37.40 ± 2.06	113.20 ± 6.18	93.70 ± 2.40	4.48 ± 0.67	0.07 ± 0.06	0 ± 0

FIGURE 7. A SCI patient choose one song in the "POP Star" sub-GUI by synchrony blinking.

application of our system. The patients' clinical data are shown in Table 2. All patients were able to perform normal eye movements but had no experience using EOG-based HCIs.

The experimental procedure was similar to that in Experiment 1 except that there were different numbers of operations. Specifically, prior to the experiment, each patient received a brief introduction to the system and completed the calibration process described in Section II. Next, the subjects were asked to switch the system on and choose 4 songs by the "Song List" and "Song Name" sub-GUIs, as shown in Fig. 7. Specifically, the experimental tasks for each SCI patient included the following sequential operations:

(i) Calibration for determining the system parameters.

(ii) Open the main menu (2 operations).

(iii) In the "POP Star" sub-GUI, select two target songs via the search engine (14 operations).

(iv) In the "Song List" sub-GUI, select two target songs from the song list (8 operations).

(v) Test the function buttons of the "Main Menu" (sub-GUI1), which include the following operations: "Prioritize" operation; "Play" operation; "Pause" operation; "Switching Song" operation; and "Hidden Menu" operation (9 operations).

After the above experimental tasks (33 operations in total) were performed, the patient was asked to remain in an idle state for 30 minutes without controlling the system (the subject could relax, listen to songs, sing, or chat with others). At the beginning of the idle state, the system was in the initial state, i.e., off state. Then, we counted the FPs and FOs during the idle state as in Experiment 1.

As shown in Table 3, all patients successfully performed the experimental tasks, with an average of time of 276.12 ± 19.30 s, an average accuracy of 93.70±2.40%, an average RT of 4.48 ± 0.67 s, and an average command ratio of 113.20 ± 6.18%. Furthermore, the average FPR and FOR in the idle state were 0.07 ± 0.06/min and 0/min for all patients.

C. WORKLOAD ASSESSMENT

In this study, we used the NASA Task Load Index (NASA-TLX) score to assess the subjective workload and satisfaction of each subject after using the proposed EOG-based music-on-demand system in Experiment 1 or 2. The NASA-TLX questionnaire contains six factors: mental needs, physical needs, temporary needs, performance, effort and setbacks [27]. Each factor has 20 intervals of scores ranging from 0 to 100. High NASA-TLX scores represent high subjective workloads. The NASA-TLX scores obtained from the healthy subjects and patients with SCIs are shown in 4 and 5, respectively. Except for the "performance" in Experiment 1, the average scores for most of the metrics were less than 20.

TABLE 4. Workload ratings for eight healthy subjects.

Healthy	Mental	Physical	Temporal	Performance	Effort	Frustration level
H1	10	15	15	10	10	10
H2	20	20	20	25	15	20
H3	15	15	20	10	10	10
H4	15	15	20	15	20	10
H5	10	5	10	5	10	0
H6	20	15	20	20	20	10
H7	15	5	20	20	10	10
H8	5	15	10	5	10	5
Avg. ± SD	13.75 ± 4.85	13.13 ± 4.96	16.88 ± 4.28	13.75 ± 6.96	13.13 ± 4.28	9.38 ± 5.27

TABLE 5. Workload ratings for five sci patients.

Patients	Mental	Physical	Temporal	Performance	Effort	Frustration level
P1	15	15	20	10	20	5
P2	15	15	20	15	15	10
P3	15	10	20	10	15	5
P4	20	10	15	20	15	15
P5	20	15	20	20	20	15
Avg. ± SD	17.00 ± 2.45	13.00 ± 2.45	19.00 ± 2.00	15.00 ± 4.48	17.00 ± 2.45	10.00 ± 4.48

This finding indicated that the subjects were satisfied with the proposed system. However, there were some differences in scores in different subjects because the subjects had different proficiency levels when using the proposed system, and those who lacked experience had high scores.

IV. DISCUSSION

This study presented an EOG-based asynchronous method for users to interact with a VR environment. The GUI included several flashing buttons, each of which corresponded to a command for VR control. The user could issue a command by blinking in synchrony with the corresponding buttons flashes. The algorithm detected the eye blinks from the recorded EOG signal to determine the users target button. As an illustrative application of this method, a music-on-demand system with a multilayer GUI structure was developed in a VR environment. Through this system, the users could search and select target songs in the music library, play, pause, switch to the next song, give priority and delete the current songs in the list. Online experiments involving 8 healthy subjects and 5 SCI patients were conducted, and the experimental results demonstrated the effectiveness of our EOG-based interaction method and music-on-demand system.

The advantages of the EOG-based HCI for interaction with a VR environment include the following: First, the hardware of EOG-based HCI is simple and wearable. Second, the EOG-based HCI achieved satisfactory performance. According to our experiments, the average number of commands generated to complete the tasks was 53.88 ± 2.85 and 37.40 ± 2.06 for healthy subjects and patients, respectively, while the optimal number was 50 and 34. The additional commands were due to the users erroneous selections. The results demonstrated that the number of erroneous selections during the tasks was kept under an acceptable level. Moreover, the average accuracy was $95.25 \pm 3.09\%$ and $93.70 \pm 2.40\%$ and the RT was 4.18 ± 0.55 s and 4.48 ± 0.67 s for healthy subjects and SCI patients to generate a command, respectively. This result was due to our system paradigm design. Specifically, EOG signals involving blinks, which are relatively higher in intensity than EEG signals, are easily detected. Furthermore, the users needed to blink according to the target buttons flashes to issue a command, and the algorithm detected the blinks that were in synchrony with these flashes. In this way, spontaneous or unintended blinks could be excluded. Third, through the EOG-based HCI, we might implement various VR applications, although only a music-on-demand application was presented in this study.

In several studies, the synchronous control mode was used; this mode requires the user to perform control operations according to the prompts of the computer system at a certain time interval [24]. However, the asynchronous mode of HCI in VR environments is more suitable for practical applications; this mode can infer whether the user intends to control the system and provide a more natural mode of operation to distinguish between the control state and the idle state [33]. To improve the asynchronous performance of our detection algorithm, we need to improve the accuracy of control state detection and reduce the FPR when the user is in an idle state. In this study, the satisfactory asynchronous performance of our system was ensured using the following strategies: First, we designed the system switch with two buttons, "ON1" and "ON2". If the subjects successfully selected the button "ON1" and then the button "ON2" within 10 s, the system entered the main menu; otherwise, the system stayed in the initial GUI. Only one "ON1/ON2" button in the initial GUI

TABLE 6. Comparison with state-of-the-art non-manual control systems.

Publication	Pattern (Signal)	Commands	Subject type	Accuracy(%) (s: synchronous; a: Asynchronous)	ITR (bits/min)	RT (s)	FRR (events/min)	FOR (events/min)
Kumar et al.[25]	Left/right/up/down/wink/ double-blinks /gaze (EOG)	7	healthy subjects	80% (s)	-	-	-	-
Coogan et al.[31]	SMR-based BCI (EEG)	4	healthy subjects	60% (s)	-	-	-	-
Tidoni et al.[16]	P300-based BCI (EEG)	9	healthy subjects	MovI+, 89.06% (s)	10.09	-	-	-
				MovI-, 83.30% (s)	9.16	-	-	-
			SCI patients	77.80% (s)	6.73	-	-	-
Our system	Single blink (EOG)	32	healthy subjects	95.25% (a)	-	4.18	0.06	0
			SCI patients	93.70% (a)	-	4.48	0.07	0

Note: "MovI+": proprioceptive stimulation with illusion; "MovI-": proprioceptive stimulation without illusion.Among all the GUIs for the music-on-demand system in this study, the song search menu (sub-GUI2 and sub-GUI3) contains the most buttons (i.e., 32 buttons, see Fig. 6(d)), and each corresponds to a control command. Therefore, the maximum number of available control commands for our system is 32.

was placed in the upper left corner of the screen, which would not hinder the immersion in the system. Second, when the current GUI did not generate any operation for 120 s, the system automatically returned to the initial GUI with only one button, "ON1". Our experimental results showed that the normal subjects and the patients with SCIs achieved a FPR of 0.06/min and 0.07/min, respectively and an FOR of 0/min. Therefore, the false operations in the idle state could be effectively reduced through the above strategies.

Previously, several interaction methods have been for VR. Three VR head display manufacturers, Oculus, SONY, and HTC, adopt a VR handle as the standard interactive mode; this handle is held by two hands and contains buttons and vibration feedback [28]. The disadvantage is that an external handle must be carried, and the tracking locator must be installed around the human body, which limits its application scenarios. Another common mode for VR interaction is gesture tracking [29], which mainly includes two categories: the first is based on optical tracking, such as in the Leap Motion and Nimble VR depth sensors; the second is based on sensors placed on the hands, such as data gloves. The drawback is that the types of gestures are limited, resulting in a small number of control commands that can be generated. In addition, the corresponding relationship between gestures and control targets can easily be confused in a complex scene. Some developers have tried to use voice control for VR interaction [30]. Due to interference from surrounding noise, this control mode has not been widely used.

In comparison with the existing nonmanual control systems for VR environments, our system might exhibit several advantages, as shown in 6. EOG signals are rarely used in the design of VR systems because of the limited number of commands (usually equal to the number of eye movements). In [25], a hands-free natural interaction system was presented that was based on EOG, which enhanced the immersion in VR games. In this system, the EOG signal acquisition device included six electrodes, and seven commands were produced by eye movements (gaze, left, right, up, down, wink, double-blink). The average accuracy for ten healthy subjects to control three targets was 80%. In our system, the generation of multiple control commands is based on the synchronization between blinking and button flashes. In this way, the proposed system requires only a single channel and an EOG mode (blinking) to provide sufficient control commands for a complex task. EEG-based BCIs have been used in the interaction with VR environments. For instance, by the help of Unity and BCI2000, an EEG-based BCI was developed to control a VR environment and several commercial internet of thing (IOT) devices [31]. In the two experiments involving 31 healthy subjects, the average accuracies were 60%. In [16], a P300 BCI-based VR interactive application was presented using an Oculus head-mounted VR display. Eleven subjects, that is, eight healthy participants and three patients with SCIs, attended the experiment to play a digital selection game in a VR environment. For the healthy subjects, an average accuracy of 89.06 $\pm$ 5.04% and an average ITR of 10.09 $\pm$ 1.55 bits per minute were achieved with the hybrid BCI combining P300 and myoelectric signals (proprioceptive stimulation with illusion, MovI+), and an average accuracy of 83.33 $\pm$ 8.63% and an average ITR of 9.16 $\pm$ 1.53 bits per minute were obtained with the P300 BCI (proprioceptive stimulation without illusion, MovI-). The three patients with SCIs achieved an average accuracy of 77.8% and an average ITR of 6.73 bits per minute with the P300 BCI. However, compared with our system, these EEG-based HCIs seem to have several disadvantages. First, all of these EEG-based HCIs were synchronous systems that could not effectively control/reduce FP when the user was in an idle state. The EOG-based HCI in the present study is an asynchronous system that achieved a low FPR (0.06 events/min for healthy subjects, 0.07 events/min for SCIs) and a zero FOR when the user was in an idle state. Second, when the user operates a system over long periods, the performance of the EEG-based HCIs might decrease due to visual fatigue and the low SNR of EEG signals. However, the subjects were satisfied with the proposed system even after a long time of use (the NASA-TLX score 9.38±5.27 to 16.88±4.28 for healthy subjects and 10.00 ± 4.48 to 19.00 ± 2.00 for SCIs). This satisfaction was because of the asynchronous nature of the system and the capability for the user to rest with little or no FPR. Third, the EEG acquisition device is generally complex.

With the EOG-based HCI, we developed a music-on-demand system in the VR environment, imitating a commercial karaoke system. A GUI tree structure including seven sub-GUIs was designed, and the user needed to blink in synchrony with the flashes of a button in a sub-GUI to issue the corresponding command. In this way, the users could search and select target songs in the music library, play the selected songs, switch to the next song, delete the current songs in the list, etc. The effectiveness and convenience of the system were demonstrated in our experiments involving both healthy subjects and SCI patients. This music-on-demand system provided entertainment for the users, especially the SCI patients, as indicated by conversations with the subjects after the experiments.

This study has several limitations. First, the RT of the proposed HCI could be further reduced to improve the efficiency of interaction with the VR environment. Second, the small number of songs in the music library (60 songs) were not sufficient for the users entertainment. Finally, only five patients with SCIs used/tested our system. In subsequent studies, we will improve our system considering these limitations.

V. CONCLUSIONS

This study proposed an EOG-based method for interaction with a VR environment. Furthermore, as an application of this method, a music on-demand system in a VR environment was established. Using this system, the subjects could search and select target songs in the music library, play, pause, switch to the next song, give priority and delete the current songs in the list, and so on. The experimental results from the healthy subjects and the patients with SCIs demonstrated the effectiveness of our EOG-based interaction method and music-on-demand system. More importantly, using this system was convenient and enjoyable for the subjects, especially the patients with SCIs. In the future, we will improve the system by enhancing the performance of the EOG-based HCI, expanding the music library and recruiting more subjects, especially patients with SCIs or stroke, to use our system.

REFERENCES

[1] G. N. Lewis, C. Woods, J. A. Rosie, and K. M. Mcpherson, ‘‘Virtual reality games for rehabilitation of people with stroke: Perspectives from the users,’’ *Disab. Rehabil. Assistive Technol.*, vol. 6, no. 5, pp. 453–463, 2011.

[2] A. Lecuyer, F. Lotte, R. B. Reilly, R. Leeb, M. Hirose, and M. Slater, ‘‘Brain-computer interfaces, virtual reality, and videogames,’’ *Computer*, vol. 41, pp. 66–72, Oct. 2008.

[3] R. Leeb *et al.*, ‘‘Walking by thinking: The brainwaves are crucial, not the muscles!’’ *Presence*, vol. 15, no. 5, pp. 500–514, 2006.

[4] D. Waller, A. C. Beall, and J. M. Loomis, ‘‘Using virtual environments to assess directional knowledge,’’ *J. Environ. Psychol.*, vol. 24, no. 1, pp. 105–116, 2004.

[5] A. Y. Dvorkin, R. V. Kenyon, and E. A. Keshner, ‘‘Reaching within a dynamic virtual environment,’’ *J. Neuroeng. Rehabil.*, vol. 4, no. 1, p. 23, 2007.

[6] D. Waller, J. M. Loomis, and D. B. Haun, ‘‘Body-based senses enhance knowledge of directions in large-scale environments,’’ *Psychonomic Bull. Rev.*, vol. 11, no. 1, pp. 157–163, 2004.

[7] C. F. Doeller, C. Barry, and N. Burgess, ‘‘Evidence for grid cells in a human memory network,’’ *Nature*, vol. 463, no. 7281, pp. 657–661, 2010.

[8] E. A. Suma, S. Clark, D. Krum, S. Finkelstein, M. Bolas, and Z. Warte, ‘‘Leveraging change blindness for redirection in virtual environments,’’ *Proc. IEEE Virtual Reality Conf.*, Mar. 2011, pp. 159–166.

[9] G. Bruder, F. Steinicke, P. Wieland, and M. Lappe, ‘‘Tuning self-motion perception in virtual reality with visual illusions,’’ *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 7, pp. 1068–1078, Jul. 2012.

[10] M. Middendorf, G. McMillan, G. Calhoun, and K. Jones, ‘‘Brain-computer interfaces based on the steady-state visual-evoked response,’’ *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 8, no. 2, pp. 211–214, Feb. 2000.

[11] E. W. Sellers, D. J. Krusienski, D. J. McFarland, T. M. Vaughan, and J. M. Wolpaw, ‘‘A P300 event-related potential brain-computer interface (BCI): The effects of matrix size and inter stimulus interval on performance,’’ *Biol. Psychol.*, vol. 73, no. 3, pp. 242–252, 2006.

[12] T. Zhang *et al.*, ‘‘Structural and functional correlates of motor imagery BCI performance: Insights from the patterns of fronto-parietal attention network,’’ *Neuroimage*, vol. 134, pp. 475–485, Jul. 2016.

[13] J. Broeren, M. Rydmark, A. Björkdahl, and K. S. Sunnerhagen, ‘‘Assessment and training in a 3-dimensional virtual environment with haptics: A report on 5 cases of motor rehabilitation in the chronic stage after stroke,’’ *Neurorehabilitation Neural Repair*, vol. 21, pp. 180–189, Mar./Apr. 2007.

[14] M. Cameirao *et al.*, ‘‘Neurorehabilitation using the virtual reality based rehabilitation gaming system: Methodology, design, psychometrics, usability and validation,’’ *J. Neuroeng. Rehabil.*, vol. 7, pp. 48-1–48-14, Sep. 2010.

[15] I. Dimbwadyo-Terrer *et al.*, ‘‘Virtual reality system Toyra: A new tool to assess and treatment for upper limb motor impairment in patients with spinal cord injury,’’ in *Converging Clinical and Engineering Research on Neurorehabilitation*. Berlin, Germany: Springer, 2013, pp. 853–858.

[16] E. Tidoni *et al.*, ‘‘Local and remote cooperation with virtual and robotic agents: A P300 BCI study in healthy and people living with spinal cord injury,’’ *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 9, pp. 1622–1632, Sep. 2017.

[17] R. J. Krauzlis, ‘‘The control of voluntary eye movements: New perspectives,’’ *Neuroscientist*, vol. 11, no. 2, pp. 124–137, 2005.

[18] A. B. Usakli and S. Gurkan, ‘‘Design of a novel efficient human–computer interface: An electrooculagram based virtual keyboard,’’ *IEEE Trans. Instrum. Meas.*, vol. 59, no. 8, pp. 2099–2108, Aug. 2010.

[19] J. F. Wu *et al.*, ‘‘Efficient implementation and design of a new single-channel electrooculography-based human–machine interface system,’’ *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 62, no. 2, pp. 179–183, Feb. 2015.

[20] T. Strandvall, ‘‘Eye tracking in human-computer interaction and usability research,’’ *Human-Computer Interaction—INTERACT* (Lecture Notes in Computer Science), vol. 5727, T. Gross *et al.*, Eds. Berlin, Germany: Springer, 2009.

[21] L. Y. Deng, C.-L. Hsu, T.-C. Lin, J.-S. Tuan, and S.-M. Chang, ‘‘EOG-based human–computer interface system development,’’ *Expert Syst. Appl.*, vol. 37, no. 4, pp. 3337–3343, 2010.

[22] J. Ma, Y. Zhang, A. Cichocki, and F. Matsuno, ‘‘A novel EOG/EEG hybrid human–machine interface adopting eye movements and ERPs: Application to robot control,’’ *IEEE Trans. Biomed. Eng.*, vol. 62, no. 3, pp. 876–889, Mar. 2015.

[23] Q. Huang *et al.*, ‘‘An EOG-based human–machine interface for wheelchair control,’’ *IEEE Trans. Biomed. Eng.*, vol. 65, no. 9, pp. 2023–2032, Sep. 2018.

[24] S. He and Y. Li, ‘‘A single-channel EOG-based speller,’’ *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 11, pp. 1978–1987, Nov. 2017.

[25] D. Kumar and A. Sharma, ‘‘Electrooculogram-based virtual reality game control using blink detection and gaze calibration,’’ in *Proc. IEEE Int. Conf. Adv. Comput., Commun. Inform.*, Sep. 2016, pp. 2358–2362.

[26] H. Zhang, C. Guan, and C. Wang, ‘‘Asynchronous P300-based brain–computer interfaces: A computational approach with statistical models,’’ *IEEE Trans. Biomed. Eng.*, vol. 55, no. 6, pp. 1754–1763, Jun. 2008.

[27] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," *Adv. Psychol.*, vol. 52, pp. 139–183, Apr. 1988.

[28] J. Kang, "Effect of interaction based on augmented context in immersive virtual reality environment," *Wireless Pers. Commun.*, vol. 98, no. 2, pp. 1931–1940, 2018.

[29] H. Cai and Y. Lin, "An integrated head pose and eye gaze tracking approach to non-intrusive visual attention measurement for wide FOV simulators," *Virtual Reality*, vol. 16, no. 1, pp. 25–32, 2012.

[30] D. Hong *et al.*, "Advances in tangible interaction and ubiquitous virtual reality," *IEEE Pervas. Comput.*, vol. 7, no. 2, pp. 90–96, Apr. 2008.

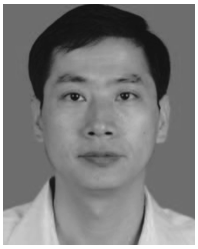
[31] C. G. Coogan and B. He, "Brain-computer interface control in a virtual reality environment and applications for the Internet of Things," *IEEE Access*, vol. 16, pp. 10840–10849, 2018.

[32] J. F. Ditunno, Jr., W. Young, W. H. Donovan, and G. Creasey, "The international standards booklet for neurological and functional classification of spinal cord injury," *Paraplegia*, vol. 32, no. 2, pp. 266–274, 1994.

[33] J. Pan, Y. Li, R. Zhang, Z. Gu, and F. Li, "Discrimination between control and idle states in asynchronous SSVEP-based brain switches: A pseudo-key-based approach," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 3, pp. 435–443, May 2013.

JING XIAO received the B.S. degree in electronic and information engineering from Xiangtan University, Xiangtan, China, in 2003, and the M.S. degree in electrical circuit and system from Hunan University, Changsha, China, in 2013. He is currently pursuing the Ph.D. degree in pattern recognition and intelligent systems with the South China University of Technology. His research interests include virtual reality and brain–computer interfaces.

JUN QU received the B.S. degree in measurement-control technology and instrumentation and the M.S. degree in microelectronics and solid electronics from Xiangtan University, Xiangtan, China, in 2005 and 2008, respectively. He is currently pursuing the Ph.D. degree in pattern recognition and intelligent systems with the South China University of Technology. His research interests include virtual reality and brain–computer interfaces.



YUANQING LI (F'16) received the B.S. degree in applied mathematics from Wuhan University, Wuhan, China, in 1988, the M.S. degree in applied mathematics from South China Normal University, Guangzhou, China, in 1994, and the Ph.D. degree in control theory and applications from the South China University of Technology, Guangzhou, in 1997. From 2002 to 2004, he was a Researcher with the Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Saitama, Japan. From 2004 to 2008, he was a Research Scientist with the Laboratory for Neural Signal Processing, Institute for Infocomm Research, Singapore. Since 1997, he has been with the South China University of Technology, where he became a Full Professor, in 2004. He has authored or co-authored more than 80 scientific papers in journals and conference proceedings. His research interests include blind signal processing, sparse representation, machine learning, brain–computer interface, electroencephalogram, and fMRI data analysis.