Behavior-based SSVEP Hierarchical Architecture for Telepresence Control of Humanoid Robot to Achieve Full Body Movement

Article *in* IEEE Transactions on Autonomous Mental Development · March 2016

DOI: 10.1109/TCDS.2016.2541162

CITATION
1

READS
101

6 authors, including:

Wei Li
California State University, Bakersfield
149 PUBLICATIONS **1,673** CITATIONS
SEE PROFILE

Hong hu
Tianjin University
2 PUBLICATIONS **1** CITATION
SEE PROFILE

Linwei Niu
West Virginia State University
28 PUBLICATIONS **259** CITATIONS
SEE PROFILE

Genshe Chen
Intelligent Fusion Technology, Inc.
258 PUBLICATIONS **1,848** CITATIONS
SEE PROFILE

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2016.2541162, IEEE Transactions on Cognitive and Developmental Systems

Behavior-based SSVEP Hierarchical Architecture for Telepresence Control of Humanoid Robot to Achieve Full Body Movement

Jing Zhao, Wei Li, *Member, IEEE*, Xiaoqian Mao, Hong Hu, Linwei Niu, and Genshe Chen

***Abstract*—The challenge to telepresence control a humanoid robot through a steady-state visual evoked potential (SSVEP) based model is to rapidly and accurately control full-body movement of the robot because a subject has to synchronously recognize the complex natural environments based on live video feedback and activate the proper mental states by targeting the visual stimuli. To mitigate this problem, this paper presents a behavior-based hierarchical architecture, which coordinates a large number of robot behaviors using only the most effective five stimuli. We defined and implemented fourteen robot behaviors for motion control and object manipulation, which were encoded through the visual stimuli of SSVEPs, and classified them into four behavioral sets. We proposed switch mechanisms in the hierarchical architecture to coordinate these behaviors and control the full-body movement of a NAO humanoid robot. To improve operation performance, we investigated the individual sensitivities of visual stimuli and allocated the stimuli targets according to frequency-responsive properties of individual subjects. We compared different types of walking strategies. The experimental results showed that the behavior-based SSVEP hierarchical architecture enabled the humanoid robot to complete an operation task, including navigating to an object and picking the object up with a fast operation time and a low chance of collision in an environment cluttered with obstacles.**

***Index Terms*—Brain-robot interaction (BRI), full-body movement, humanoid robot, steady-state visual evoked potential (SSVEP), electroencephalogram (EEG), telepresence control**

This wok was supported in part by the State Key Laboratory of Robotics at Shenyang Institute of Automation under Grant No. 2014-Z03; and National Nature Science Foundation of China under Grant No. 61473207.

Jing Zhao is with the School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China. (Email: zhaoj55@tju.edu.cn).

Wei Li is with the State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China; School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China; and Department of Computer & Electrical Engineering and Computer Science, California State University, Bakersfield, California 93311, USA. (phone: +661-654-6747; fax: +661-654-6960; Email: wli@csub.edu). Wei Li is the author to whom correspondence should be addressed.

Xiaoqian Mao is with the School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China. (Email: maoxiaoqian@tju.edu.cn).

Hong Hu is with the School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China. (Email: honghu@tju.edu.cn).

Linwei Niu is with the Department of Math and Computer Science, West Virginia State University, Institute, WV 25112, USA. (Email: lniu@wvstateu.edu)

Genshe Chen is with the Intelligent Fusion Technology, Inc, Germantown, MD 20876, USA. (Email: gchen@intfusiontech.com).

I. Introduction

BRAIN-robot interaction (BRI) providing an innovative communication between a human and a robotic device via brain signals is prospective for applying robot systems in many areas [1-3]. A commonly used technology for designing brain-computer interfaces (BCIs) for BRI is acquiring electroencephalogram (EEG) signals from electrodes placed on the scalp due to its non-invasive and inexpensive assay, ease of use, and acceptable temporal resolution [4, 5]. Among a variety of robotic devices, humanoid robots are advanced: they are created to imitate some of the same physical and mental tasks that humans undertake daily, to perform complicated tasks such as personal assistance, where they should be able to assist the sick and elderly, and to perform dirty or dangerous jobs [6, 7].

Three types of brainwave-based models, event-related potentials (ERPs), motor-imagery (MI) potentials, and steady-state visual evoked potentials (SSVEPs), are commonly used for control of a humanoid robot. Bell *et al.* (2008) used a P300-based strategy to select targets that a humanoid robot reached using visual feedback in an environment with prior knowledge, e.g., to discern an object the robot should pick up and a location the robot should bring the object to [8]. Li *et al.* (2011) used motor imagery potentials to control humanoid robot walking behaviors formulated by the robot kinematics [9]. Chung *et al.* (2011) presented a SSVEP-based model for training and controlling a humanoid robot in a simulated home environment by learning high-level commands [10]. Chae *et al.* (2102) proposed a posture-dependent control paradigm to translate three motor imagery states into five robot motions to navigate a humanoid robot in a maze using "artificial landmarks," such as a circle on the wall presenting a waypoint and an arrow sign indicating the path direction in [11]. Choi *et al.* (2013) combined the P300 potentials, SSVEP signals, and MI potentials to solve a multi-task problem using an experimental scenario by remotely controlling a humanoid robot in a properly sized maze with "artificial landmarks" [12]. Zhao *et al.* (2013) proposed an OpenViBE-based system to evaluate the P300 potentials, SSVEPs, and MI potentials for controlling a humanoid robot [13]. Li *et al.* (2013) developed a P300-based adaptive model for controlling a humanoid robot, including an off-line training and an on-line execution [14]. Bouyarmane *et al.* (2014) used motor imagery potentials to control the walking gait of a simulated humanoid robot through

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2016.2541162, IEEE Transactions on Cognitive and Developmental Systems

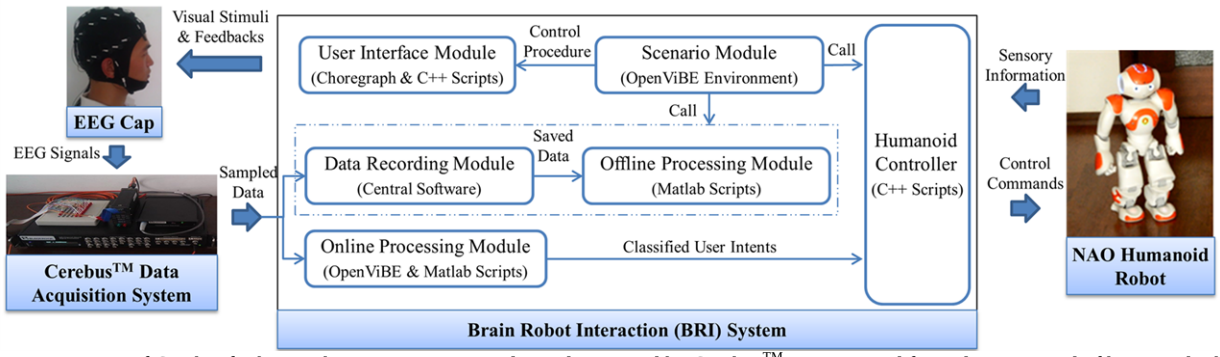

Fig. 1. System structure of Cerebot for brain robot interaction. Neural signals acquired by Cerebus™ are processed for real time control of humanoid robot, or are saved for off-line analysis by modules in dashed box. Major software tools utilized in building the modules are listed between parentheses.

waypoints in [15]. Li *et al.* (2014) navigated a humanoid robot in an office environment with an obstacle and picked an object up via motion-onset visual evoked potentials [16].

Compared with semi-autonomous control of a humanoid robot via brainwaves in an environment with prior knowledge [8, 10] or with "artificial landmarks" [11, 12], telepresence control of a humanoid robot to accomplish complex operation tasks in uncertain environments, especially under dim illumination conditions where a visual guide system is ineffective, is much more challenging for the following reasons: First, controlling the full-body movement of a humanoid robot with high degrees of freedom (DOF) requires a large number of visual stimuli encoding human mental intentions corresponding to a variety of robot behaviors, which need to be effectively coordinated under live video feedback. Second, the current BCI approaches have to optimize the balance between on-line success rate and information transfer rate (ITR). In this paper, we comprehensively extend our initial report in [17] to a novel behavior-based SSVEP hierarchical architecture to mitigate the first issue, i.e., to coordinate a large number of humanoid robot behaviors to achieve full-body movement. A behavioral approach to solving a complicated multi-task problem in robotics is to decompose the task into a set of behaviors with simple features [18-20] and activate them using a coordination mechanism. In this study, the operation task is defined: a subject telepresence controls a humanoid robot based on live video feedback to move from a start point to an object in an uncertain environment and to pick up the object by activating his/her mental activities. To complete this task, we define fourteen behaviors for motion control level and object manipulation level, which are classified into four different sets: walking with gross movement (WGM), walking with fine tuning (WFT), environment inspection (EI), and manipulation operation (MO). A NAO humanoid robot, made by Aldebaran in France [21], is used in this study. We formulate each behavior using the NAO robot programming language in Choregraphe and C++, and activate each of the fourteen behaviors under the proposed hierarchical architecture.

The hierarchical architecture proposed herein is built upon an adaptive content-coding SSVEP interface, i.e., a visual stimulus with a single frequency adapts its flashing robot images representing an encoded human mental activity, according to the status of the robot and its surroundings based on live video feedback. Under the hierarchical architecture implemented by only the five SSVEP stimuli, including the four visual stimuli encoding robot behaviors and the one selecting the behavioral sets, we are able to telepresence control the humanoid robot to achieve full-body movement for accomplishing the operation task, which requires the fourteen behaviors. The adaptive content-coding SSVEP interface enables a subject to understand intuitively the corresponding robot behavior encoded by each stimulus and helps reduce training time and habitual mistakes. The reason for choosing the SSVEP stimuli is that this model yields a relatively high ITR and requires less training. The on-line open-loop control experiments across five subjects, including non-experienced subjects, demonstrate that the SSVEP model reaches an average success rate of over 88%, an average response time of 3.48 s, and an average ITR of 27.3 bits/min. The experiments of the operation task show that operation performance depends not only on BCI system performance but also on the definition of the robot behaviors as well as their coordination mechanisms. The proposed hierarchical architecture exhibited a navigation time of 236.1 s and the a total collision number of 0.6 times/trial, including both sideswipes and obstacle hits, thereby achieving a better trade-off than was achieved by only using the behaviors defined in the WFT or WGM sets, which exhibited navigation times of 271.2 s and 167.2 s and total collision number of 0.2 times/trial and 1.8 times/trial. The four metrics for the evaluations of the control performance are discussed in detail in Section V-B.

This paper is organized as follows. Section II states the challenge of the operation task to telepresence control a humanoid robot and proposes the behavior-based SSVEP hierarchical architecture to solve this problem. Section III describes the implementation of the SSVEP model for evoking and processing brain signals. Section IV details the implementation of the proposed hierarchical architecture of the BRI system. Section V describes the evaluation procedure of the BRI system for both on-line open-loop control and on-line closed-loop control of the humanoid robot. Section VI discusses the experimental results of 5 subjects including non-experienced subjects. Section VII concludes this study with discussion and future works.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2016.2541162, IEEE Transactions on Cognitive and Developmental Systems

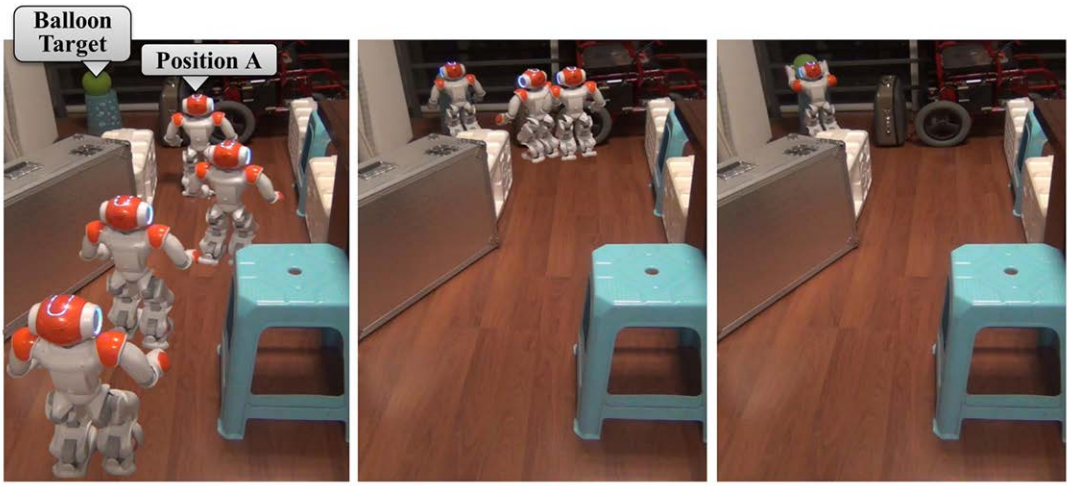

Fig. 2. Operation task of controlling NAO robot to walk in cluttered environment, to reach and pick up balloon target. (a) The robot was controlled to walk to Position A using Walking with Gross Movement (WGM) behaviors; (b) The robot reached the green balloon target using Walking with Fine Tuning (WFT) behaviors; (c) The robot picked up the balloon target.

II. Problem Description and Behavior Based Control

A. *Cerebot Platform*

We use Cerebot – the brainwave-controlled humanoid robot platform [9, 13], consisting of a Cerebus™ Data Acquisition System and a NAO humanoid robot, to implement and evaluate the behavior-based SSVEP hierarchical architecture. The Cerebus™ Data Acquisition System provides a number of analog I/Os and digital I/Os and is able to record up to 256 signal channels simultaneously up to a sampling rate of 30 kHz with 16-bits resolutions [22]. Its Software Development Kits provides libraries in C++ and MATLAB to design an experimental procedure, e.g., to pre-process and display bio-potential signals acquired by various types of electrodes. The NAO robot is equipped with multiple sensors, including 2 cameras, 4 microphones, 2 sonar rangefinders, 2 IR emitters and receivers, 1 inertial board, 9 tactile sensors, and 8 pressure sensors. It provides a graphical programming environment – Choregraphe and a Software Development Kit in C++ for creating and editing robot behavior.

Fig. 1 shows the system structure of Cerebot developed under the OpenViBE environment [9, 13]. OpenViBE is an open-source and general-purpose software package with powerful protocols for off-line analysis and on-line process of brain signals by exchanging data between modules and algorithms developed on a variety of programming platforms [23]. Cerebot integrates a number of software packages, such as OpenGL, OpenCV, WEBOTS, Choregraphe, Central software, and user-developed programs in C++ and MATLAB to present visual stimuli, to feedback live video about an environment, to acquire brain signals through Cerebus™, to extract their features and classify them according to their patterns, and to control humanoid robot behavior via brainwaves through wireless connection.

B. *Operation Task and Robot Behavior Definition*

Fig. 2 shows the defined on-line operation task: based on live video feedback, a subject navigates the humanoid robot from a start point to reach and pick up a balloon via brain signals. This is very popular in robotics research because it consists of walking through an unstructured environment with obstacle avoidance and finely controlling the robot to approach and pick up a desired object. The two factors that challenge autonomous accomplishment of the operation task by the humanoid robot include the following: first, detecting all irregular obstacles and recognizing landmarks in a low-contrast environment with uncertainty; second, effectively controlling movements of the humanoid robot with high degrees of freedom. Telepresence control, based on live video feedback, is a good way to manage with the first issue. A behavior-based strategy is effective to solve the second problem because the complicated operation task can be completed by decomposing the operation into a set

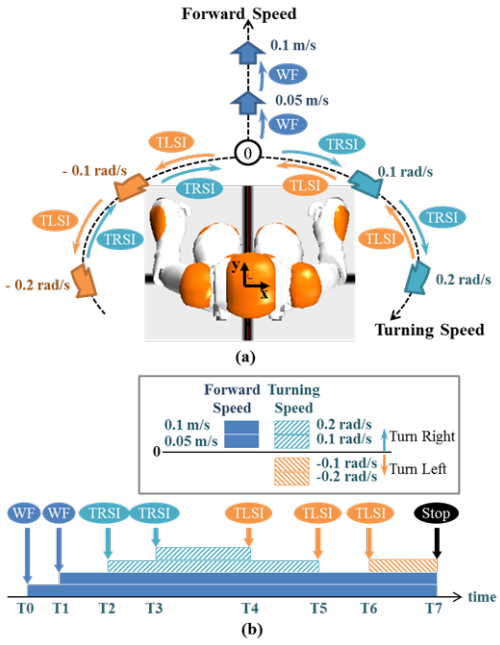

Fig. 3. (a) Coordinate system on the humanoid robot during walking. (b) Rules for increasing speeds of the WF, TRSI and TLSI behaviors. The legend of Figure 3b lists the forward and turning speeds represented by different number of colored bars, e.g., one navy blue bar represents the forward speed of 0.05 m/s and two bars represent the forward speed of 0.1 m/s.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2016.2541162, IEEE Transactions on Cognitive and Developmental Systems

of behaviors with simple features [18-20] and activating them using a coordination mechanism. To accomplish the operation task proposed herein, we define the following fourteen behaviors and implement them using the C++ SDK for the NAO humanoid robot:

Scan from left to right (SLR): This behavior controls the robot head to turn -0.75 rad (left), to scan surroundings from -0.75 rad to +0.75 rad (from left to right), and to return its head to face forward.

Scan from up to down (SUD): This behavior controls the robot head to raise +0.5 rad, to scan from +0.5 rad to -0.5 rad (from up to downward), and to return its head to face forward.

Activate bottom camera (ABC): This behavior activates the bottom camera embedded in the robot head to deliver live video of the surroundings in front of the robot feet, which assists a subject to telepresence control the robot to walk through a cluttered environment.

Activate top camera (ATC): This behavior activates the top monocular camera embedded in the robot head, which assists the subject to target an object to be reached.

Pick up object (PUO): This behavior controls the NAO robot to pick up the object if the robot hands are not holding an object.

Put down object (PDO): This behavior controls the robot to put down the object if the robot hands are holding an object.

Walk forward (WF): This behavior increases a forward speed by 0.05 m/s for walking forward. The maximum forward speed is set as 0.1 m/s. Fig. 3a defines the coordinate system embedded in the humanoid robot body. Turning right is positive whereas turning left is negative.

Turn right with speed increment (TRSI): This behavior controls the robot to make a right turn by an increasing speed interval of 0.1 rad/s. The maximum right turn speed is set as 0.2 rad/s, as shown in Fig. 3b.

Turn left with speed increment (TLSI): This behavior controls the robot to make a left turn by an increasing speed interval of -0.1 rad/s. The maximum left turn speed is set as -0.2 rad/s shown in Fig. 3b.

Stop: This behavior terminates all the behaviors and resets the forward and turning speeds to zero.

Step forward (SF): This behavior controls the robot to slowly step forward by a fixed distance of 0.1 m at 0.05 m/s. This behavior is important for the subject to control the robot to pass through a narrow corridor or to adjust robot locations to approach an object to be picked up.

Step backward (SB): This behavior controls the robot to slowly step backward by a fixed distance of 0.1 m at 0.05 m/s. This behavior is important for the subject to adjust robot locations when the robot faces an obstacle too closely.

Turn right with fixed radian (TRFR): This behavior controls the robot to make a right turn with a fixed radian of 0.4 rad.

Turn left with fixed radian (TLFR): This behavior controls the robot to make a left turn with a fixed radian of -0.4 rad.

An SSVEP model yields relatively high ITR and requires less training, but controlling these fourteen robot behaviors to accomplish the on-line operation task via SSVEP signals is very challenging because it is difficult to encode human mental intentions by flashing 14 visual stimuli simultaneously on a single computer LCD monitor with a refresh frequency of 60 Hz. In addition, simultaneously flashing too many visual stimuli may distract a subject's attention, increase the necessary training time to become familiar with the coding relations between the visual stimuli and the robot behaviors, and even activate unexpected robot behaviors, especially under live video feedback.

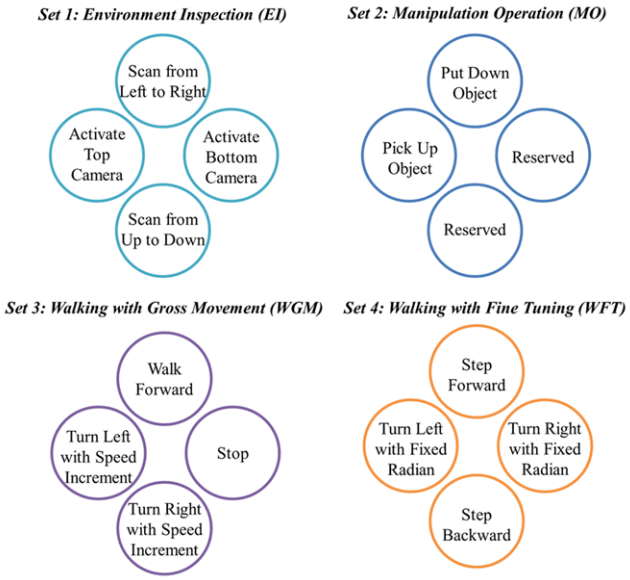

Fig. 4. Fourteen types of humanoid robot behaviors implemented in the humanoid controller. Our hierarchical architecture classifies them into four sets for menu-based selection.

We use the hierarchical architecture to solve these problems by classifying these behaviors into four sets: environment inspection (EI), manipulation operation (MO), walking with gross movement (WGM), and walking with fine tuning (WFT), as shown in Fig. 4. The EI set consists of four types of humanoid robot behaviors: SLR, SUD, ABC, and ATC, which are used to activate cameras embedded in the robot head and feedback surrounding information using live video. The MO set consists of two behaviors for manipulating an object: PUO and PDO. The WGM set consists of four behaviors: WF, TLSI, TRSI, and Stop, which are able to control robot compound movements by combining WF and TLSI or TRSI with large paces. Fig. 3b shows an example of how to modulate speeds by combining the WF, TRSI, TLSI, and Stop behaviors. At time T0, the subject activates WF to navigate the robot forward at a speed of 0.05 m/s, and then increases its forward speed to 0.1 m/s at T1. When the robot confronts a corner at T2, the subject activates TRSI twice to control the robot to make a right turn at the maximum speed of 0.2 rad/s, while the forward speed maintains 0.1 m/s. After the robot leaves the corner at T4, the subject uses TLSI to decrease the turning speed to 0.1 rad/s for adjusting robot directions, and then activates TLSI a second time to make the robot walk straight forward. Activating TLSI again at T6 allows the robot to make a left turn at 0.1 rad/s. Activating Stop terminates all walking behaviors at T7.

The behaviors in WGM navigate the robot to walk fast in an

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2016.2541162, IEEE Transactions on Cognitive and Developmental Systems

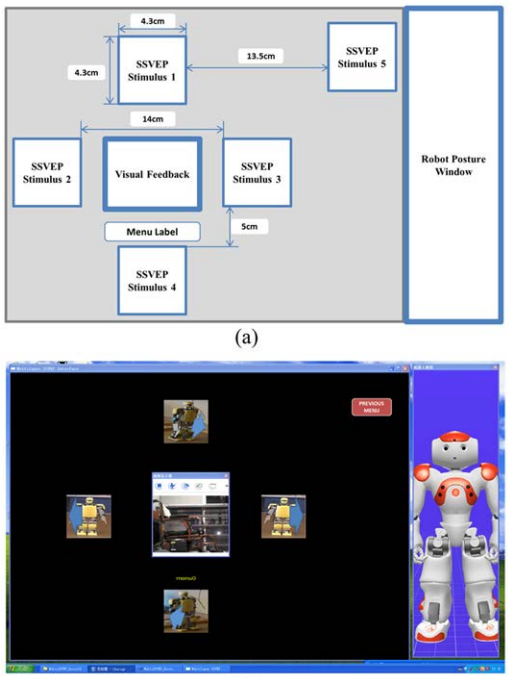

Fig. 5. (a) Layout of user interface. (b) User interface for behavior-based SSVEP hierarchical architecture.

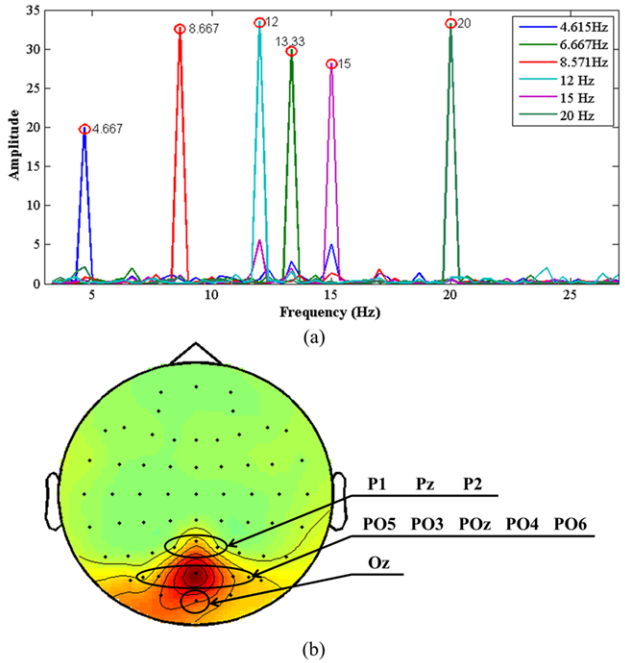

Fig. 6. (a) SSVEP power spectrum of subj1 after CCA-based signal processing. (b) Topographical distribution of the power spectral density of subj1 averaged when staring at the flickering target. Nine electrodes are selected to extract the SSVEP features.

environment with large free space as well as stop movement. The WFT set consists of four behaviors: SF, TRFR, TLFR, and SB. The behaviors in WFT are used to finely tune robot locations when the robot moves in a cluttered environment or approaches a selected object when the cameras deliver live video with poor quality. The hierarchical architecture arranges the four behavior sets using different stimulus interfaces and activates each of the fourteen humanoid robot behaviors by switching to its corresponding stimulus interface.

III. SSVEP Model for Evoking and Processing Brain Signals

In 1966, Regan [24] discovered harmonics of electrical potentials evoked by flickering sinusoidally modulated light using an analogue Fourier series analyzer. These types of brainwaves respond to visual stimuli at a given frequency and are known as SSVEPs. SSVEPs prominently appear over the visual cortex in the occipital region of the scalp [25]. The brain signal, $y_i(t)$, evoked by the i^{th} SSVEP stimulus at the time t is described by [26].

$$y_i(t) = \sum_{k=1}^{N_h} a_{i,k} \sin\left(2\pi k f_i t + \Phi_{i,k}\right) + B_{i,t} \quad i = 1, 2, ..., N \quad (1)$$

where f_i is the flickering frequency of the i^{th} visual stimulus, N is the total number of stimuli, N_h is the number of considered harmonics, $a_{i,k}$ and $\Phi_{i,k}$ are the amplitude and the phase of each sinusoid, and $B_{i,t}$ includes noise, artifacts and any components irrelevant to the SSVEP response.

Our model uses a stimulus interface, as shown in Fig. 5, to elicit SSVEP responses. The stimulus interface is implemented based on Windows DirectX API and Choregraphe on a 22-inch LCD monitor with a resolution of 1440 × 900 pixels and a refresh rate of 60 frames per second. The visual stimuli on the interface are modulated by square stimulation signals, and thus have to flicker at a period of N frames on the LCD screen (N = 1,2,3,…,30) [27]. Considering that each stimulus elicits SSVEP responses at the same frequency and its harmonics, it is unreliable to flicker stimuli at frequencies that are each other harmonics [28]. As a result, we scanned all possible stimulus frequencies and selected the most responsive six frequencies: 4.615 Hz, 6.67 Hz, 8.57 Hz, 12 Hz, 15 Hz and 20 Hz, which correspond to thirteen, nine, seven, five, four and three frames within a flickering period. Fig. 6a shows the power spectrums of the SSVEP signals induced by the six stimuli from the subject subj1. The SSVEP signal associated with each selected frequency is labeled near its peak, as depicted by curves in Fig. 6a. For each subject, we exclude the frequency evoking SSVEPs with the lowest signal-to-noise ratio (SNR) and reserve the remaining five frequencies to flicker the five robot images on the interface as the visual stimuli, e.g., for the subject subj1, the frequency of 4.615 Hz is excluded due to its SSVEP responses with the lowest feature amplitude as shown in Fig. 6a. For the on-line closed-loop control evaluation, the middle window on the interface displays live video feedback and the right panel shows the robot's current posture in 3D. Fig. 5b shows a screenshot of the user interface during an on-line closed-loop control experiment.

We acquire brain signals at a sampling rate of 1 kHz using a standard EEG cap with 64 channels placed with respect to the international 10-20 system. The ground electrode is placed at FPz on the forehead and an electrode at vertex of the head is used as a reference. Fig. 6b shows the power scalp distribution on the channels averaged over the six SSVEP frequencies

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2016.2541162, IEEE Transactions on Cognitive and Developmental Systems

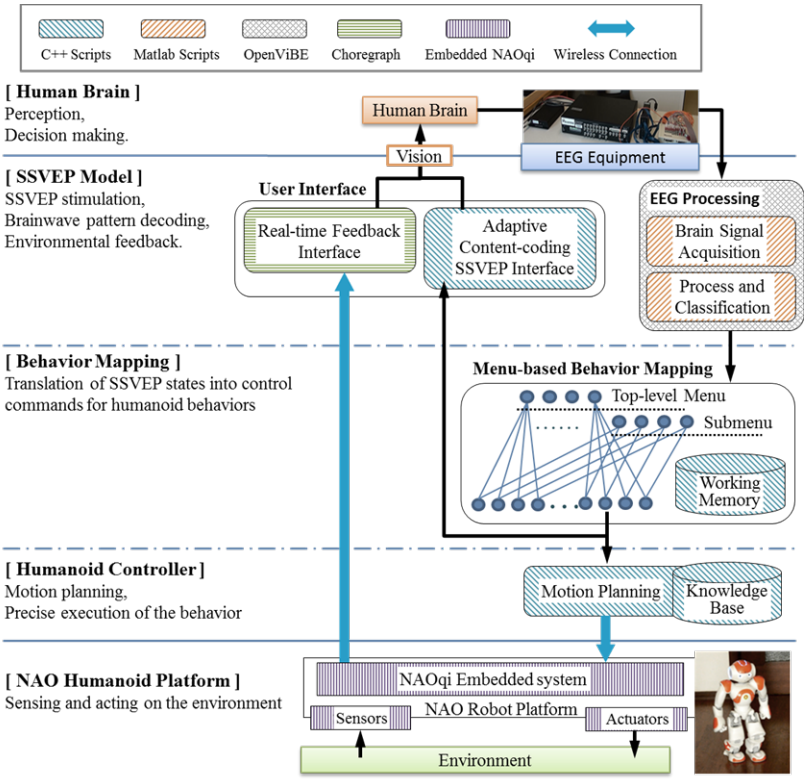

Fig. 7. Behavior-based SSVEP hierarchical architecture establishes connection between human brain and humanoid robot.

acquired from subj1. According to the power distribution, nine channels, Oz, POz, PO3, PO4, PO5, PO6, Pz, P1 and P2, are selected to extract the SSVEP features for classifying the brainwaves to mitigate the individual issue. In every 0.5s, we process the most recent 3s segment of EEG data. We use a low-pass filter of 30Hz to filter the data segment from each channel and apply the canonical correlation analysis (CCA) method to spatially filter the multi-channel brain signals. CCA is a multivariable statistical method used to find the underlying correlation of two sets of data. Considering multi-channel EEG data X and reference data Y with the same length of 3s, CCA finds the weight vectors W_x and W_y to maximize the correlation coefficient ρ between $x = X^T W_x$ and $y = Y^T W_y$ by solving the following equations [29]:

$$\max_{W_x, W_y} \rho(x, y) = \frac{\mathrm{E}\left[x^T y\right]}{\sqrt{\mathrm{E}\left[x^T x\right]\mathrm{E}\left[y^T y\right]}} = \frac{\mathrm{E}\left[W_x{}^T X Y^T W_y\right]}{\sqrt{\mathrm{E}\left[W_x{}^T X X^T W_x\right]\mathrm{E}\left[W_y{}^T Y Y^T W_y\right]}} \tag{2}$$

The reference data Y are set to be:

$$Y = \begin{pmatrix} \sin(2\pi N_1 f_1 t) \\ \cos(2\pi N_1 f_1 t) \\ \sin(2\pi N_2 f_2 t) \\ \cos(2\pi N_2 f_2 t) \\ \vdots \\ \cos(2\pi N_5 f_5 t) \\ \cos(2\pi N_5 f_5 t) \end{pmatrix} \tag{3}$$

where $f_1, f_2, \cdots, f_5$ are the flickering frequencies of the five visual stimuli, and $N_1, N_2, \cdots, N_5$ indicate the most responsive harmonic for each stimulus.

The canonical variant vector W_x projects the multi-channel data X to one-dimension data x, which ensure the most prominent correlation with the linear combination of five stimuli's reference data. Fast Fourier Transform (FFT) calculates the power spectral density (PSD) of the data x.

$$P(\omega) = \left|\mathrm{FFT}(x, N_{FFT})\right|^2 / N_{FFT} \tag{4}$$

where N_{FFT} is the sample size and $\mathrm{FFT}(x, N_{FFT})$ returns the N_{FFT}-point discrete Fourier transform of x. The most responsive powers at the n_i^{th} harmonic frequency for the i^{th} SSVEP stimulus target $(i = 1,2,3,4,5)$ are normalized for extracting the SSVEP features by (5).

$$F_i = \frac{P(2\pi n_i f_i)}{\underset{3Hz \le f \le 25Hz}{\mathrm{mean}}\left(P(2\pi f)\right)} \quad (i = 1,2,3,4,5) \tag{5}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2016.2541162, IEEE Transactions on Cognitive and Developmental Systems

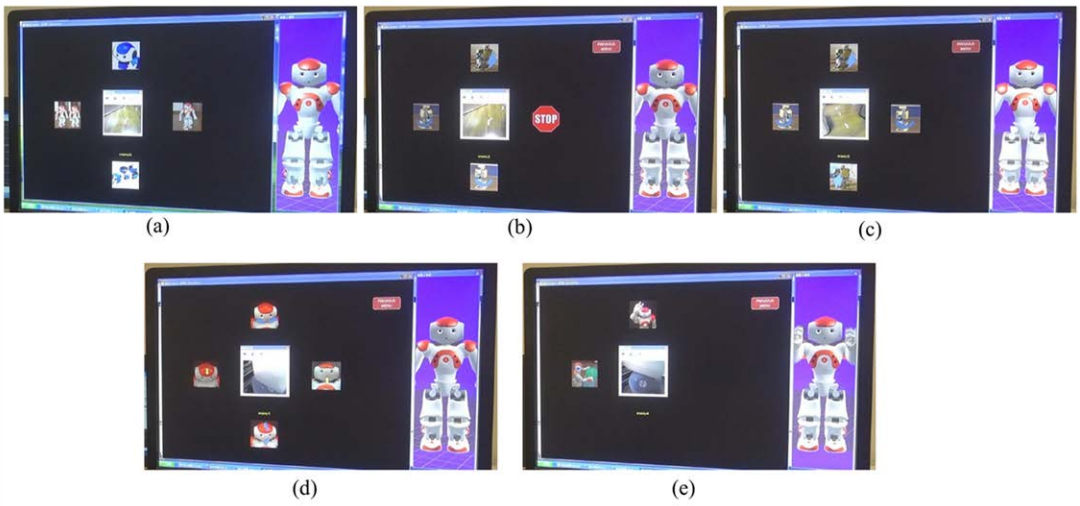

Fig. 8. Pictures of the adaptive content-coding SSVEP interface taken during a closed-loop control experiment. (a) Top-level menu; (b) Submenu of the WGM set; (c) Submenu of the WFT set; (d) Submenu of the EI set; (e) Submenu of the MO set.

where $\underset{3Hz \le f \le 25Hz}{\text{mean}}\left(P(2\pi f)\right)$ denotes the average of the power spectrum between 3 and 25Hz. The normalized feature F_i of the five frequencies is detected when it is above a given threshold σ_i. n_i and σ_i need to be calibrated for each subject during the off-line training process because their corresponding power feature F_i heavily depends on individuals and inherent variability.

IV. Behavior-based SSVEP Hierarchical Architecture

The hierarchical architecture coordinates the fourteen visual stimuli that encode the human mental intentions corresponding to the robot behaviors, which are needed to achieve full-body movement of the humanoid robot. It consists of three layers, as shown in Fig. 7. The SSVEP Model layer presents the visual stimuli using an adaptive content-coding SSVEP interface, classifies the SSVEP brain signals, and displays environmental information from live video feedback. The Behavior Mapping layer maps the five brainwave patterns elicited by the visual stimuli to commands for controlling the fourteen robot behaviors. The Humanoid Controller layer translates each command into a series of discrete motions to activate the corresponding robot behavior. The proposed architecture is developed by multiple programming packages, such as VC++, MATLAB, OpenViBE, Choregraphe and embedded NAOqi, as shown in Fig. 7.

We propose a two-level menu based behavior mapping layer that allows a subject to select a behavioral set on a top-level menu and to further select a desired robot behavior on an appropriate submenu. On the top-level menu, only four visual stimuli represent the four behavioral sets, i.e., WGM, WFT, EI, and MO, as shown in Fig. 8a. The subject is able to switch from the top-level menu to one of the relevant submenus representing the four behavioral sets. A submenu flickers five visual stimuli. Stimuli 1-4 around the live video feedback represent the four robot behaviors defined in the submenu, and stimulus 5 located in the top-right corner flashes a red square with the words “PREVIOUS MENU” to switch the interface back to the top-level menu. The subject activates a robot behavior by staring at its corresponding stimulus in the submenus.

To help the subject understand better the robot behaviors represented by the visual stimuli, we implement an adaptive content-coding SSVEP interface in the SSVEP Model layer to flicker a group of robot images as the visual stimuli, the contents of which represent the corresponding control commands and change status when the menus are switched. Fig. 8 shows the screenshots of the interface, including the top-level menu and the four submenus during a closed-loop control experiment. The stimulation properties such as color, brightness and contrast affect the elicited SSVEP responses [27]. For this study, we select nineteen robot images with moderate brightness and contrast to represent varieties of commands, i.e., selecting the four behavioral sets, switching back to the top-level menu, and activating the fourteen types of robot behaviors.

V. Experiments and Evaluation

A. Subjects

The experiments were carried out in an office environment without electromagnetic shielding. The five healthy subjects (four male and one female, aged 23-29, have normal or corrected-to-normal vision) participated in the experiments. All the subjects understood the experimental procedures very well and were the first time participants of the closed-loop operation task as described in Section II-B. The subjects subj1, subj2 and subj3 among them were proficient in the SSVEP experiments, but only the subject subj3 had prior experience on a relative easy operation task in a spacious room [17]. The subjects subj4 and subj5 had never participated in the closed-loop control experiment using SSVEPs. In the experiments, the subjects sat in a comfortable armchair, 50 cm away from the visual stimuli presented on a 22-inch LCD monitor with a 60Hz refresh rate. This project was reviewed and approved by Tianjin medical university general hospital ethics committee, and all subjects

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2016.2541162, IEEE Transactions on Cognitive and Developmental Systems

TABLE I
EXPERIMENTAL PROCEDURES

Off-line calibration	Calibration of the SSVEP model
	Selection of 5 frequencies from the 6 available ones
On-line open-loop control evaluation of the SSVEP model	Evaluation of success rates, response times and ITRs
	Allocation of the 5 frequencies to the visual stimuli
On-line closed-loop control evaluation of performing the operation task	Manual control session
	WFT-only × 1
	Hybrid WFT-WGM × 2
	WGM-only × 2
	Hybrid WFT-WGM × 1
	WGM-only × 1

gave written consent.

B. Experimental Procedure

The experimental procedures included the off-line calibration, the on-line open-loop control evaluation of the SSVEP model, and the on-line closed-loop control evaluation of performing the operation task.

The off-line calibration included recording the brain signals of each subject, establishing his/her feature vectors, and calibrating the SSVEP model. Each subject was trained how to shift his/her focal point on an image and to adjust his/her mental status. For each subject, the five frequencies with the most responsive SSVEPs were selected from the frequencies of 4.615, 6.667, 8.571, 12, 15 and 20 Hz.

The on-line open-loop control asked the subjects to output randomly a sequence of commands by staring at the corresponding stimuli. To evaluate success rates, response times, and ITRs yielded by each subject, the BRI system only used the SSVEP Model layer to flicker images at the five frequencies used in the off-line calibration process. Each subject underwent a total of 50 trials, i.e., 10 trials for each of the five stimuli. In each trial, a stimulus was selected randomly as the target. The user interface displayed a yellow arrow over the target stimulus that the subject should focus on. The system processed the real-time EEG signals, and provided audio feedback of the classification results using beep sounds with five different tones. The subject adjusted his/her mental status to evoke the desired SSVEP responses within a period of 7 s after the yellow arrow appeared. The EEG data along with the stimulation markers, which indicated when the corresponding stimulus was activated as the target, were recorded to calculate the success rates, the response times, and the ITRs. The success rate, p, represents a percentage of trials in which the desired commands are successfully activated within the period of 7 s. The response time, T, represents the time elapsed after the subject received an instruction until he/she successfully outputted the command. The ITR in bits/min is defined in [30].

$$ITR = \frac{p\log_2(p) + (1-p)\log_2\left(\frac{1-p}{N-1}\right) + \log_2(N)}{T} \quad (6)$$

where N is the number of defined commands. As the five visual stimuli on the user interface, shown in Fig. 5, are used to encode different robot behaviors as well as menu selection commands, we defined the following stimuli with the five frequencies by considering their success rates and the response times:

Rule 1: Stimulus 3, defined by the visual stimulus with the shortest response time, was used to stop the robot promptly when one of the WF, TLSI, or TRSI behaviors defined in WGM was active.

Rule 2: Stimulus 1, defined by the visual stimulus with the highest success rate and a relatively short response time, was used to encode the WF or SF behavior, as it was most frequently activated during performing the operation task.

Rule 3: Stimulus 5, defined by the visual stimulus with the longest response time and the lowest success rate, was used to associate with the menu selection command of switching back to the top-level menu, as it was the least frequently activated.

Rule 4: Stimulus 4, defined by the visual stimulus with the lower success rate and response time of the two remaining visual stimuli, was used to encode the SB behavior because SB was less frequently activated during performing the operation task.

Rule 5: Stimulus 2 was defined by the visual stimulus that was not allocated through Rules 1-4.

The on-line closed-loop control evaluation asked the subjects to telepresence control the full-body movement of the humanoid robot via SSVEP responses to accomplish the operation task discussed in Section II-B. In doing so, the subjects practiced performing the operation task in two sessions.

First, they used a keyboard manually to control the robot to accomplish the operation task several times to become familiar with the control procedure. The manual control session allowed the subjects to manipulate the hierarchical architecture-based BRI system using five keys on the keyboard, i.e., “up,” “down,” “left,” “right” and “Q,” which replaced the five flickering stimuli of the BCI interface.

Second, they compared three types of brain-controlled strategies, namely WFT-only, WGM-only, and hybrid WFT-WGM, through the BRI system. WFT-only allowed the subjects to accomplish the operation task using only the four behaviors defined in the WFT set: SF, TRFR, TLFR, and SB. WGM-only allowed the subjects to accomplish the operation task using only the four behaviors defined in the WGM set: WF, TLSI, TRSI, and Stop. The hybrid WFT-WGM strategy allowed the subjects to accomplish the operation task using the behaviors defined in both the WFT and WGM sets.

The BRI control session trained the subjects how to use these three types of brain-controlled strategies to accomplish the operation task. Each subject conducted seven trials to complete the brain control session. In the first trial, the subject practiced how to use the WFT-only strategy to perform the operation task. In the second and third trials, the subject practiced how to use the hybrid WFT-WGM strategy to perform the operation task, i.e., the subject activated the behaviors in WGM to navigate the robot from the start position to the position A, as shown in Fig. 2, and then the behaviors in WFT to pass through a narrow area to avoid object collisions. In the fourth and fifth trials, the subject practiced how to use WGM-only to perform the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2016.2541162, IEEE Transactions on Cognitive and Developmental Systems

operation task. The sixth trial was an additional hybrid WFT-WGM trial, and the seventh trial was an additional WGM trial. Table I shows the sequence of performing all the seven trials. The major difference of these trials lies in the method of activating the walking behaviors; thereby, we evaluated navigation performances using the following metrics.

1) Total navigation time: The time was taken to navigate the humanoid robot from the start point to reach the target position.

2) Output commands for navigation: The number of commands was outputted for controlling the humanoid robot to reach the target.

3) Sideswipe collisions: The number of sideswipe collisions was recorded, in which the humanoid robot grazed an obstacle but did not stop walking.

4) Obstacle collisions: The number of obstacle collisions was recorded, in which the humanoid robot hit an obstacle and had to be repositioned to continue with the subsequent navigation task.

VI. Results

Table II lists the on-line open-loop control results of the five subjects yielded by the SSVEP model. For each subject, the frequency with the lowest responsive SSVEPs among the six frequencies was excluded from encoding his/her mental intentions. We measured the success rates and response times for the selected five frequencies and averaged them to evaluate ITR. The results show that all the subjects achieve significant success rates and response times but with different individual amplitudes yielded by the same stimulus targets. These individualities are caused by differences of inherent human sensitivities to the frequencies that influence the amplitudes of the evoked SSVEP responses. Subj5, who participated in the SSVEP experiments the first time, achieved the highest average success rate of 96%, the shortest average response time of 3.29 s, and the best ITR of 36.5 bits/min. Subj4, who had no prior experience with the on-line SSVEP experiments, delivered the lowest average success rate of 82% and ITR of 21.9 bits/min. The other three subjects subj1, subj2 and subj3, who were experienced with the on-line SSVEP experiments, achieved average success rates of 92.0%, 84.0%, and 86.0%, average response times of 3.39 s, 3.53 s, and 3.66 s, and average ITRs of 31.1 bits/min, 23.2 bits/min, and 23.9 bits/min, respectively.

TABLE II
On-line Testing Results of the Five Subjects Yielded by the SSVEP Model

Subject		subj1	subj2	subj3	subj4	subj5	Mean
Success rate (%) for each frequency	4.615Hz	N/A	100^{1}	100^{1}	N/A	90^{4}	N/A
	6.667Hz	100^{1}	100^{3}	60^{4}	100^{1}	100^{3}	
	8.571Hz	90^{5}	90^{2}	70^{5}	100^{3}	N/A	
	12Hz	90^{2}	N/A	N/A	50^{5}	90^{5}	
	15Hz	90^{3}	40^{5}	100^{3}	60^{4}	100^{2}	
	20Hz	90^{4}	90^{4}	100^{2}	100^{2}	100^{1}	
	Average	92	84	86	82	96	88
Response time (s) for each frequency	4.615Hz	N/A	2.88^{1}	3.25^{1}	N/A	3.24^{4}	N/A
	6.667Hz	3.57^{1}	2.66^{3}	4.47^{4}	3.29^{1}	2.94^{3}	
	8.571Hz	4.22^{5}	3.16^{2}	5.00^{5}	2.85^{3}	N/A	
	12Hz	3.34^{2}	N/A	N/A	3.72^{5}	3.86^{5}	
	15Hz	2.41^{3}	4.75^{5}	2.66^{3}	4.46^{4}	3.25^{2}	
	20Hz	3.39^{4}	4.22^{4}	2.95^{2}	3.28^{2}	3.18^{1}	
	Average	3.39	3.53	3.66	3.52	3.29	3.48
ITR (bits/min)		31.1	23.2	23.9	21.9	36.5	27.3

The superscripts indicate the corresponding stimulus numbers selected for each frequency. We list the number for a frequency on both its success rate and response time.

Table II also shows the intra-individual variability of the SSVEP responses with respect to different flickering frequencies. We allocated these frequencies to the five visual stimuli, which encode different mental intentions, following the rules discussed in Section V-B. The superscripts listed in Table II indicate the corresponding stimulus numbers assigned by each frequency.

Table III lists the on-line closed-loop operation task performance performed by one WFT-only trial and three WGM-only trials for each subject. Please note that the performance yielded by the three WGM-only trials is represented by their averages over three trials. The metrics are explained in Section V-B. The following remarks on the results are addressed.

1) In the WFT-only experiments, all the subjects successfully accomplished the operation task through the BRI control with no obstacle collision. The average navigation time was 271.2 s, the average number of output commands was 51.8 Cmds/trial[1],

TABLE III
Closed-loop Control Performance Yielded by the WFT-only and WGM-only Strategies

Subject	WFT-only (1 trial)				WGM-only (3 trials)			
	Time (s)	Commands (Cmds/trial)	Sideswipe Collisions (times/trial)	Obstacle Collisions (times/trial)	Time (s)	Commands (Cmds/trial)	Sideswipe Collisions (times/trial)	Obstacle Collisions (times/trial)
subj1	284	53	0	0	164.7	22.7	2/3	2/3
subj2	285	54	1/1	0	166.7	24.7	4/3	3/3
subj3	250	52	0	0	136	21	2/3	1/3
subj4	267	52	0	0	161.7	18	6/3	5/3
subj5	270	48	0	0	207	27	3/3	0
Mean (±SD)	271.2 (±14.3)	51.8 (±2.3)	0.2 (±0.4)	0	167.2 (±25.5)	22.7 (±3.4)	1.1 (±0.6)	0.7 (±0.6)

[1] The abbreviation "Cmds" indicates the number of output commands.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2016.2541162, IEEE Transactions on Cognitive and Developmental Systems

TABLE IV
CLOSED-LOOP CONTROL PERFORMANCE OF THE HYBRID TASK

Subject	Navigation Time (s)			Sideswipe Collisions (times/trial)	Obstacle Collisions (times/trial)
	Total	WGM	WFT		
subj1	262.7	94.3	147.3	1/3	0
subj2	223	83	124.7	3/3	0
subj3	242.7	90.3	134.3	0	0
subj4	210.7	86.3	110.7	3/3	1/3
subj5	241.3	86.3	134.3	0	1/3
Mean (±SD)	236.1 (±20.0)	88.1 (±4.4)	130.3 (±13.6)	0.5 (±0.5)	0.1 (±0.2)

and the average number of sideswipe collisions was 0.2 times/trial. In the WGM-only experiments, all the subjects spent less time and outputted fewer commands for completing the operation task. The average navigation time was 167.2 s, and the average number of output commands was 22.7 Cmds/trial. However, all of the subjects had more collisions (the average number of sideswipe was 1.1 times/trial, and the average number of obstacle hitting was 0.7 times/trial) in the WGM-only experiments because of difficulties in promptly estimating the distance to obstacles through monocular video feedback and stopping the humanoid robot immediately through the BRI control.

2) In the WFT-only experiments, the average navigation time and average number of output commands achieved by the subjects subj4 and subj5, who participated in the SSVEP experiments the first time, were 268.5 s and 50 Cmds/trial, whereas those achieved by the experienced subjects, subj1, subj2 and subj3, were 273 s and 53 Cmds/trial. These results indicate that the subjects with no prior experience can achieve performance comparable with those of the experienced subjects; thus, the WFT-only experiments demand less training of the subject.

In the WGM-only experiments, however, the subjects obtained diverse performances. Subj3, who had participated in a similar experiment before [17], achieved the best performance with the lowest navigation time of 136 s, 2 sideswipe collisions in the 3 trials, and 1 obstacle collision in the 3 trials. Subj1 and subj2, who were proficient with on-line SSVEP experiments but had not performed this type of complicated operation task, achieved performances with average navigation times of 164.7 s and 166.7 s, 2 and 4 sideswipe collisions in the 3 trials, and 2 and 3 obstacle collisions in the 3 trials. Subjects subj4 and subj5, who had no prior experience with the on-line SSVEP experiments, achieved performances with average navigation time of 161.7 s and 207.0 s, 6 and 3 sideswipe collisions in the 3 trials, and 3 and 0 obstacle collisions in the 3 trials. These results indicate that training a subject to get familiar with using the BRI control system and conducting the operation task is critical to improve his/her performance in the WGM-only experiments.

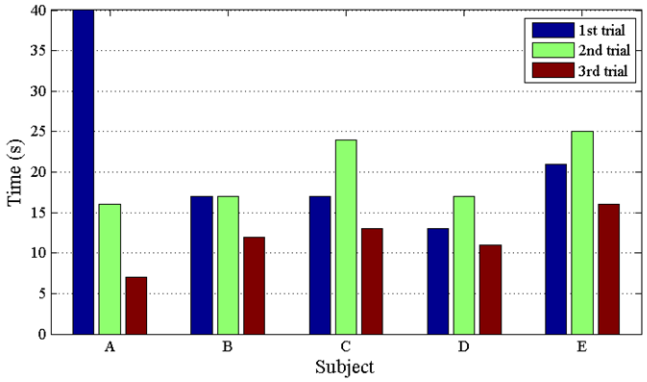

Fig. 9. Time cost of switching the menu by each subject in the three trials of hybrid experiments.

3) Both subj4 and subj5, who were participating in the WGM-only experiments for the first time, delivered very different performances. Subj5 spent the longest navigation time (207 s), but had only 3 sideswipe collisions and 0 obstacle collisions in the 3 trials, while subj4 spent less navigation time (161.7 s), but delivered the largest number of collisions, including 6 sideswipe and 5 obstacle collisions in the 3 trials. The reason is that subj5 spent more time perceiving the environment and adjusting the position of the humanoid robot by outputting many more control commands with 27 Cmds/trial. Subj4, who delivered a lower accuracy rate and longer response time than subj5 did in the on-line open-loop experiments, caused 5 more obstacle collisions in the WGM-only experiments, and thus spent less time adjusting the robot's position through BRI control. When an obstacle collision occurred, the robot was manually repositioned for continuing the operation task. This action saved time to re-adjust the robot position and orientation for completing the operation task.

Table IV summarizes the performances of completing the operation task documented in the hybrid WFT-WGM trials, in which the robot behaviors in both WFT and WGM sets could be flexibly combined. We compare the performances with those achieved in the WFT-only and WGM-only trials as follows.

1) All the subjects spent the less average navigation time (236.1 s) in the hybrid WFT-WGM trials than in the WFT-only trials (271.2 s), caused fewer collisions (0.5 sideswipe collision/trial and 0.1 obstacle collision/trial) than in the WGM-only trials (1.1 sideswipe collisions/trial and 0.7 obstacle collisions/trial). The WGM-only experiments indicated that training helped reduce the number of collisions, but almost did not affect the hybrid WFT-WGM trials which were performed before the WGM-only trials.

2) In the hybrid WFT-WGM trials, the hierarchical architecture-based BRI system had to be switched from the WGM set to the WFT set at position A, as shown in Fig. 2. The subject needed to activate three commands sequentially: switching back to the top-level menu, selecting the WFT set, and activating the desired WFT behavior. We measured the time from the end of the last WGM behavior to the start of the first WFT behavior. Fig. 9 shows these results from the five subjects in the 3 trials. Subj1 spent the longest time (40 s) among all the subjects in his first trial, in which he switched the menu a total of 3 times. We excluded this trial to calculate the average time for all the subjects using the remaining 14 trials.

The results indicated that the five subjects spent an average time of 16.1 ± 4.9 s for switching between the WGM and WFT set. One of the possible reasons for spending so much time is that the frequency selected for stimulus 5, which corresponds to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2016.2541162, IEEE Transactions on Cognitive and Developmental Systems

the command of "switching back to the top-level menu," results in the longest response time and the worst success rate in on-line testing. In addition, all the subjects used the shortest amount of time (11.8±3.3 s on average) in their last trials. This can likely be attributed to increasing experience, which reduces their hesitancy in making decisions and activating the desired behavior.

VII. Discussion

A. Performance of the Hierarchical Architecture

This paper evaluated the performance of the proposed hierarchical architecture using both on-line open-loop experiments and on-line closed-loop experiments. The on-line open-loop testing results show that the 5 subjects achieved an average success rate of 88%, an average response time of 3.48s, and an average ITR of 27.3 bits/min (in particular, the ITR of subj5 was 36.5 bits/min). This performance is comparable to that of the existing SSVEP-based BRI systems for control of a humanoid robot [10, 12, 17], but full-body movement of the humanoid robot was achieved.

The experiments also reflect the individual impacts of subjects on the proposed BRI system. As listed in Table II, different subjects achieved different success rates and response times yielded by the same stimulus targets. To reduce the individual influence, this paper allocated different frequencies to the five visual stimuli, which encode different robot behaviors as well as the menu selection commands, by considering the success rates and the response times of each subject and following the rules discussed in Section V-B. Fig. 10 shows the statistical results of the response times evoked by the five stimuli. Stimulus 1, 2, and 3, encoding the robot behaviors that are frequently used in the task, resulted in better performance for all the subjects.

The on-line closed-loop evaluation experiments demonstrate that the hierarchical architecture enables a subject to use only five SSVEP stimuli to telepresence control the full-body movement of a humanoid robot to accomplish a complex operation task, which requires fourteen types of robot behaviors. The five subjects spent 16.1±4.9 s on average to switch between the WGM and WFT sets through activating three BRI commands, but they spent 11.8±3.3 s on average in their last trials. This result suggests that training helped reduce the time spent switching.

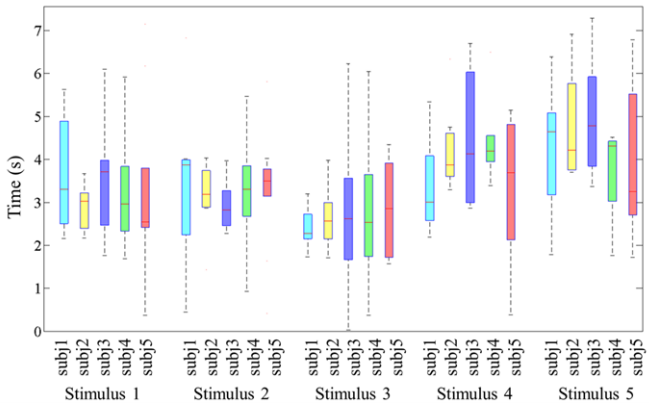

Fig. 10. Statistic results of the response time evoked by each stimulus.

B. WFT Behaviors versus WGM Behaviors

This paper compares navigation performance conducted using two different types of walking behaviors defined in the WFT and WGM sets. The WGM-only experiments achieved reduced average navigation time (167.2 versus 271.2 s) and fewer control commands (22.7 versus 51.8 Cmds/trial) to complete the on-line operation task, but more collisions were caused (sideswipe collisions: 1.1 versus 0.2 times/trial; obstacle collisions: 0.7 versus 0 times/trial), and more training was required. The WFT set is suitable for a new user or for an operation task in a cluttered environment, i.e., it is more appropriate for applications where collisions are more likely to be caused; the WGM set is suitable for an experienced user or for an operation task in an environment with a large free space. For real-world applications, such as the operation task described in this paper, we suggest the flexible use of both sets of robot behaviors according to environmental situations. The hybrid WFT-WGM experiments conducted in this study balanced navigation time and collisions better than did the WFT-only and WGM-only experiments.

C. Individual Issues

Table I lists the performances obtained by subjects under the different stimulation frequencies in the on-line open-loop control experiments, which confirmed the phenomena of inter-subject inconsistency that exists in every BCI study and often leads to BRI illiteracy, indicating some users are unable to attain effective control [31]. We diminish the individuality in two aspects. First, we set up the rules for allocating the five frequencies to the visual stimuli to encode the robot behaviors by considering the success rates and the response times achieved by individual subjects. The on-line closed-loop control results demonstrated that these rules enhanced the overall control performance. Second, we defined the four sets of robot behaviors and properly activated a robot behavioral set according to surrounding information. Usually, the WFT-only strategy (the average navigation time was 271.2±14.3 s) was more robust than the WGM-only strategy (the average navigation time was 167.2±25.5 s). Most likely, only either an experienced user or a subject with high ITRs can stop the robot rapidly and accurately after activating the WGM behavioral set. This result implies that optimizing robot behavior activation helps reduce inter-subject variability in the overall performance of the BRI system even with relatively low ITRs. As a result, the proposed behavior-based hierarchical architecture allows flexibly coordinating different behavioral sets for both BRI illiterates and experienced users.

VIII. Conclusions

This paper proposes a behavior-based SSVEP hierarchical architecture, consisting of an SSVEP Model layer, a Behavior Mapping layer, and a Humanoid Controller layer. The results of both the on-line open-loop experiments evaluating the SSVEP model and the on-line closed-loop experiments of performing the operation task demonstrated the effectiveness of the

behavior-based SSVEP hierarchical architecture in coordinating fourteen robot behaviors, which are encoded by corresponding human mental intentions, to achieve full-body movement of a humanoid robot. In our future research, we will improve the system performance to optimize time intervals needed for switching between the behavioral sets in two aspects. First, we will further improve the SSVEP model in its ITR and accuracy of classifying the five visual stimuli. Wang *et al.* (2015) and Allison *et al.* (2014) proposed novel hybrid methods to combine SSVEP and P300 models to improve classification accuracy [32, 33]. We will evaluate their performance of coordinating a large amount of behaviors in the proposed hierarchical architecture. Second, we will optimize the proposed hierarchical architecture by integrating the environmental perception and machine learning methods to coordinate behaviors more efficiently. The system will be developed to perceive the environment autonomously and display the optimal behaviors, which would most likely be used in the current situation, on the SSVEP interface. In addition, we will also report the performance of accomplishing the operation task using the behavior-based hierarchical architecture based on other brainwave-based models, e. g., ERPs, MI potentials and hybrid models [34].

ACKNOWLEDGMENT

The authors would like to thank Mr. Guoxing Zhao and Miss Mengfan Li for their help in conducting the experiments for this paper. The authors would also like to thank the reviewers for their suggestions, which have greatly improved the paper.

REFERENCES

[1] M. A. Lebedev and M. A. Nicolelis, "Brain-machine interfaces: past, present and future," *Trends Neurosci*, vol. 29, pp. 536-46, Sep 2006.

[2] D. J. McFarland and J. R. Wolpaw, "Brain-Computer Interface Operation of Robotic and Prosthetic Devices," *Computer*, vol. 41, pp. 52-56, Oct 2008.

[3] A. Ortiz-Rosario and H. Adeli, "Brain-computer interface technologies: from signal to action," *Rev Neurosci*, vol. 24, pp. 537-52, 2013.

[4] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin Neurophysiol*, vol. 113, pp. 767-91, Jun 2002.

[5] S. Bozinovski and A. Bozinovski, "Mental States, EEG Manifestations, and Mentally Emulated Digital Circuits for Brain-Robot Interaction," *IEEE T Auton Ment De*, vol. 7, pp. 39-51, Mar 2015.

[6] K. Hirai, M. Hirose, Y. Haikawa, and T. Takenaka, "The development of Honda humanoid robot," in *Proc. ICRA*, 1998, pp. 1321-1326.

[7] S. Nishide, J. Tani, T. Takahashi, H. G. Okuno, and T. Ogata, "Tool-body assimilation of humanoid robot using a neurodynamical system," *IEEE T Auton Ment De*, vol. 4, pp. 139-149, Jun 2012.

[8] C. J. Bell, P. Shenoy, R. Chalodhorn, and R. P. Rao, "Control of a humanoid robot by a noninvasive brain-computer interface in humans," *J Neural Eng*, vol. 5, pp. 214-20, Jun 2008.

[9] W. Li, C. Jaramillo, and Y. Li, "A brain computer interface based humanoid robot control system," in *Proc. IASTED*, 2011, pp. 390-396.

[10] M. Chung, W. Cheung, R. Scherer, and R. P. Rao, "A hierarchical architecture for adaptive brain-computer interfacing," in *Proc. IJCAI* 2011, pp. 1647-1652.

[11] Y. Chae, J. Jeong, and S. Jo, "Toward Brain-Actuated Humanoid Robots: Asynchronous Direct Control Using an EEG-Based BCI," *IEEE T Robot*, vol. 28, pp. 1131-1144, Oct 2012.

[12] B. Choi and S. Jo, "A low-cost EEG system-based hybrid brain-computer interface for humanoid robot navigation and recognition," *PLoS One*, vol. 8, p. e74583, 2013.

[13] J. Zhao, Q. Meng, W. Li, M. Li, F. Sun, and G. Chen, "An OpenViBE-based brainwave control system for Cerebot," in *Proc. ROBIO*, 2013, pp. 1169-1174.

[14] M. Li, W. Li, J. Zhao, Q. Meng, F. Sun, and G. Chen, "An adaptive P300 model for controlling a humanoid robot with mind," in *Proc. ROBIO*, 2013, pp. 1390-1395.

[15] K. Bouyarmane, J. Vaillant, N. Sugimoto, F. Keith, J. Furukawa, and J. Morimoto, "Brain-machine interfacing control of whole-body humanoid motion," *Front Syst Neurosci*, vol. 8, p. 138, 2014.

[16] W. Li, M. Li, and J. Zhao, "Control of humanoid robot via motion-onset visual evoked potentials," *Front Syst Neurosci*, vol. 8, p. 247, 2014.

[17] J. Zhao, Q. Meng, W. Li, M. Li, and G. Chen, "SSVEP-based hierarchical architecture for control of a humanoid robot with mind," in *Proc. WCICA*, 2014, pp. 2401-2406.

[18] R. C. Arkin and R. R. Murphy, "Autonomous navigation in a manufacturing environment," *IEEE T Robotic Autom*, vol. 6, pp. 445-454, Aug 1990.

[19] R. Brooks, "A robust layered control system for a mobile robot," *IEEE Journal of Robotics and Automation*, vol. 2, pp. 14-23, Mar 1986.

[20] W. Li, "Fuzzy-logic-based reactive behavior control of an autonomous mobile system in unknown environments," *Eng Appl Artif Intel*, vol. 7, pp. 521-531, Oct 1994.

[21] A. Corp. *Aldebaran Robotics* [Online]. Available: https://www.aldebaran.com/en

[22] B. M. Corp. *Cerebus - Animal* [Online]. Available: http://www.blackrockmicro.com/content.aspx?id=13

[23] Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, *et al.*, "OpenViBE: An Open-Source Software Platform to Design, Test, and Use Brain-Computer Interfaces in Real and Virtual Environments," *Presence-Teleop Virt*, vol. 19, pp. 35-53, Feb 2010.

[24] D. Regan, "Some characteristics of average steady-state and transient responses evoked by modulated light," *Electroencephalogr Clin Neurophysiol*, vol. 20, pp. 238-248, Mar 1966.

[25] F. B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: focus on essential paradigms and future perspectives," *Prog Neurobiol*, vol. 90, pp. 418-38, Apr 2010.

[26] I. Volosyak, H. Cecotti, D. Valbuena, and A. Gräser, "Evaluation of the Bremen SSVEP based BCI in real world conditions," in *Proc. ICORR* 2009, pp. 322-331.

[27] D. Zhu, J. Bieger, G. G. Molina, and R. M. Aarts, "A survey of stimulation methods used in SSVEP-based BCIs," *Comput Intell Neurosci*, vol. 2010, p. 1, Jan 2010.

[28] G. R. Muller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller, "Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components," *J Neural Eng*, vol. 2, pp. 123-30, Dec 2005.

[29] 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," *J Neural Eng*, vol. 6, p. 046002, Aug 2009.

[30] J. R. Wolpaw, H. Ramoser, D. J. McFarland, and G. Pfurtscheller, "EEG-based communication: improved accuracy by response verification," *IEEE Trans Rehabil Eng*, vol. 6, pp. 326-33, Sep 1998.
[31] 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 Trans Rehabil Eng*, vol. 18, pp. 107-116, Apr 2010.
[32] M. Wang, I. Daly, B. Allison, J. Jin, Y. Zhang, L. Chen, and Xingyu Wang, "A new hybrid BCI paradigm based on P300 and SSVEP", *J Neurosci Meth*, vol. 244, pp. 16-25, Apr 2015.
[33] B. Allison, J. Jin, Y. Zhang, and X. Wang, "A four-choice hybrid P300/SSVEP BCI for improved accuracy," *Brain-Computer Interfaces*, vol. 1, pp. 17-26, Jan 2014.
[34] F. Duan, D. Lin, W. Li, and Z. Zhang, "Design of a multimodal EEG-based hybrid BCI system with visual servo Module," *IEEE T Auton Ment De*, vol. 7, pp. 332-341, Dec 2015.

Jing Zhao received the B.S. in automation from Shandong University, Jinan, P.R. China in 2008 and the M.S. in measuring and testing technologies and instruments from 4th Academy of China Aerospace Science and Technology Corporation (CASC), Xian, P.R. China in 2011. Currently, he is a Ph.D. student at the School of Electrical Engineering and Automation, Tianjin University, Tianjin, P.R. China. His research interests include brain robot interaction and telepresence control of humanoid robot.

Wei Li received his B.S. (1982) and M.S. (1984) in electrical engineering from the Northern Jiaotong University, Beijing, P.R. China, and a Ph.D. (1991) in electrical and computer engineering from the University of Saarland, Germany. Currently, he is a Professor with the Department of Computer and Electrical Engineering and Computer Science at California State University, Bakersfield.

From 1996 to 2002, he was a full professor of Computer Science and Technology at Tsinghua University, Beijing, P.R. China. From 1999 to 2001, he was a Research Scientist of the Department of Electrical Engineering at the University of California, Riverside. He has been appointed a guest Professor at the State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, and an adjunct Professor in the School of Electrical Engineering and Automation, at Tianjin University, respectively in 2009 and 2012. His research interests include brain robot interaction, biologically inspired robots, chemical plume tracing and mapping, fuzzy control and modeling, and multisensor fusion and integration.

Dr. Li received the 1995 National Award for Outstanding Postdoctoral Researcher in China and the 1996 Award for Outstanding Young Researcher at Tsinghua University. He was an Alexander von Humboldt Foundation Research Fellow at the Technical University of Braunschweig, Germany (1997–1999).

Xiaoqian Mao received the B.S. and M.S. in measuring and testing technologies and instruments from Northeastern University, Shenyang, P.R. China in 2010 and 2012, respectively. Currently, he is a Ph.D. student at the School of Electrical Engineering and Automation, Tianjin University, Tianjin, P.R. China. His research interests include image processing and brain robot interaction.

Hong Hu received the B.S. in automation from Tianjin University, Tianjin, P.R. China in 2014. Currently, he is a postgraduate student at the School of Electrical Engineering and Automation, Tianjin University, Tianjin, P.R. China. His research interests include brain robot interaction and intelligent wheelchair.

Linwei Niu received the B.S. in computer science and technology from Peking University, Beijing, P.R. China in 1998, the M.S. in computer science from State University of New York at Stony Brook in 2001, and the Ph.D. in computer science and engineering from University of South Carolina in 2006. Currently, he is an associate professor in the Department of Math and Computer Science, West Virginia State University, USA. Dr. Niu is a member of IEEE (since 2005). His research interests include power-aware design for embedded systems, design automation, software/hardware co-design, and real-time scheduling.

Genshe Chen received the B. S. and M. S. in Electrical Engineering, Ph. D. in Aerospace Engineering, in 1989, 1991 and 1994 respectively, all from Northwestern Polytechnical University, Xian, China. Currently he is the Chief Technological Officer of Intelligent Fusion Technology, Inc., Germantown, MD, at where he directs the research and development activities for the Government Services and Commercial Solutions.

He was the CTO of DCM Research Resources LLC, Germantown, MD, and the program manager in Networks, Systems and Control at Intelligent Automation, Inc., leading research and development efforts in target tracking, information fusion and cooperative control. He was a Postdoctoral Research Associate in the Department of Electrical and Computer Engineering of The Ohio State University from 2002 to 2004. He worked at the Institute of Flight Guidance and Control of the Technical University of Braunshweig (Germany) as an Alexander von Humboldt research fellow and at the Flight Division of National

Aerospace Laboratory of Japan as a STA fellow from 1997 to 2001. He did postdoctoral work at the Beijing University of Aeronautics and Astronautics and Wright State University from 1994 to 1997. His research interests include SWAP-C sensor design, intelligent signal processing, multi INT fusion, space situational awareness, cognitive radio for satellite communication, cooperative control and optimization for military operations, guidance, navigation, and control, Cyber Security, C4ISR, Electronic Warfare, Game theory, Graphical theory, and Human-Cyber-Physical System.