Android tele-operation through Brain-Computer Interfacing: A real-world demo with non-expert users

Christoforos Christoforou
R.K.I Leaders Limited
Agias Triados 26A,
7100, Aradippou, Cyprus
cchristoforou@krileaders.com

Nikolaos Mavridis, Emerson Lopes Machado
Interactive Robots and Media Lab
United Arab Emirates University
Al Ain, UAE

George Spanoudis
Department of Psychology
University of Cyprus
Nicosia, Cyprus

***Abstract*— Towards our vision of natural robotic tele-presence by thought, a real-world system which was demonstrated in public with non-expert users is presented. The system consists of an EEG-based brain computer interface module (BCI), located in Cyprus, and relying on machine learning methods for recognizing brain patterns and translating them into control-commands. The BCI module is controlling Ibn Sina, an android robot which is located 1400 miles away in the United Arab Emirates, in real-time, through an IP connection. Visual feedback from the robot is provided through a video link to the users. Previously untrained human users operated the system in multiple occasions (during public demos). We describe the demonstrations, as well as the structure of our system, and report quantitative measures of our system's performance. This is up to our knowledge the first time that non-expert humans have controlled a remote android through thought during a public demonstration.**

I. INTRODUCTION

Telepresence is defined as action at a distance or the projection or presence where you physically aren't. Robotic Telepresence is a subset of telepresence whereby a robot acts as a surrogate for a remote user. Its aim is to allow the remote user to perform physical action in remote places, or potentially hazardous environments. A number of potential applications for such technology, have been identified. For example, doctors with difference expertise could connect to the remote robot and provide diagnosis and treatment to the patient, or salesmen could provide product-specific information in a shopping mall. The ability of multiple human operators to share a single physical presence couldprovide the right expertise (for example, medical) to the right place at the right time. In addition, security and surveillance robots could be used to check up on pets or family members from far away. Such robots could also allow health-care professionals to monitor elderly people taking medication at home to ensure the dosage and routine are correct.

Such robotic telepresence systems have been demonstrated in the past. For example, Hiroshi Ishiguro created an android he called Geminoid HI-1 [1]. The android is a look-alike of its creator and performs with lifelike movements such as blinking, "breathing" and fidgeting. It can be remotely controlled via a motion capture system that tracks the user's mouth movements and allows the robot to speak. Another, example is NASA robots Spirit and Opportunity, which are roaming the surface of Mars on behalf of researchers here on the planet earth, giving researchers the ability to see and manipulate on the Martian surface.

Current robotic telepresence technologies, rely on conventional interfaces for controlling such robots. Advances in machine learning and brain-computer interface research allow for non-invasive decoding of human thought and translate them into commands [2]. Such interface could provide a more natural and direct control over robotic telepresence.

Towards our vision of a natural thought-controlled robotic telepresence, this paper presents a real-world BCI enabled robot telepresence system, which was demonstrated in two occasions. We will start with a description of the demonstrations, followed by the methods used and modules of our system. Then, we will proceed with quantitative results on the performance of our proposed thought-controlled robotic telepresence system. Finally, we will derive a conclusion and present our plans for future extensions.

II. MODULES, METHODS AND DEMONSTRATIONS

The initial demonstration of our system was held in July 2009, during the our lab's summer school, and was based on a simple left/right movement imagination task BCI [2], driving our Ibn Sina android [3] through action mapping mechanisms, providing a correspondence to two predetermined movements. During this demonstration, the human subject was connected to the BCI module, and after a short calibration session, could issue one of two commands to the android.

Our public demo took place on September 25th 2009 as part of "e-night" researchers' night event organized in Cyprus. Two non-expert humans operated the system, which had a P300 speller front-end [4], mapping to 6 discrete movements of the robot:

Say Hello, Move the head Left, Move the head Right, Move the head Up, Move the head Down, and Smile

In figure 1 and figure 2, we present scenes of the live demonstration of our thought-controlled robotic telepresence. The video from the live demo, as well as other videos and information regarding our activities and robots, are available in [5],[6].

Our system is composed of five main components: the *stimulus-presentation* module, the *EEG analysis engine*, the *robot-control* module, the live video-*feedback* module, and the Ibn Sina *android robot*. The users of the system, would observe the stimulus-presentation on an LCD monitor, where a number of stimuli were presented. The EEG analysis engine, after temporal integration of evidence, would then translate their brain responses into control commands, which were transmitted via a TCP/IP connection to the robot-control module. The robot would be in the UAE, while the humans wearing the BCI in Cyprus. The live video-feedback module transmitted live images of the robot performing the actions. For the live-video feedback, we used Skype's [7] video conference features, while for the robot-control module, we implemented our own interface in Java. The architecture of the overall system is presented in figure 3, while the details of the stimulus presentation module, the signal analysis engine, the robot-control module as well as the robot are presented in the section that follows.

A. Stmulus Presentation

For the stimulus presentation module, we implemented a P300 paradigm [4]. The specific paradigm uses visual stimulation enabling to the selection of commands; the method displays a 5x6 matrix composed of characters, (letters and digits). In our case, we used the 24 letters of the Greek alphabet, and the digits one through six. Characters are briefly highlighted in a random order. The user of the system is asked to observe the matrix, and recognize when a character of choice is highlighted. The paradigm relies on the P300 observation, a signal in the EEG that is generated when the user realizes that the letter of his choice is highlighted.

Figure 1. Live demonstration of Brain-Computer Interface

Figure 2. Video Feedback from Remote Android Robot

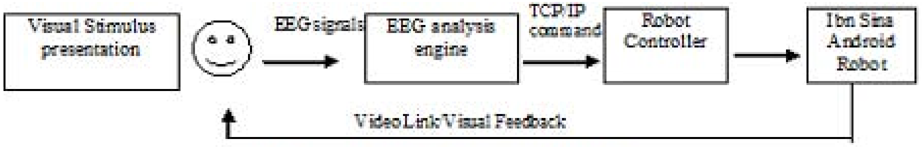

Figure 3. Modular System Architecture

B. *Signal Analysis Engine*

The signal analysis engine is responsible for decoding the EEG signal and translates it into commands towards the robot control module. In order to decode which character the user intents to choose, our system exploits the P300 signal. This corresponds to an increase in the amplitude of the EEG (on average) at about 300ms after a stimulus onset. In our paradigm, the P300 is expected to be present when the character the user wants to select is highlighted and not-presented when any other character is highlighted. By identifying the presence or absence of the P300 signal in the EEG segments right after each character is highlighted, we can pinpoint accurately the user's choice.

There are a number of challenges in identifying the P300 signal in EEG, on a single-trial basis (i.e. using only a short window segment following the onset of the character highlighted). Challenges include, the low signal-to-noise ratio; typically, less than -20dB, the variability in EEG across different users and across sessions, and the high dimensional space of the EEG observation. To address these issues, our signal analysis engine uses machine learning techniques to identify the P300 signal.

C. *Robot Control*

The robot control module receives commands over the TCP/IP link, and translates them to motion sequences to be played back by the robot. The motion sequences contain desired angles for the robot's motors, for each particular time step. These are handcrafted using the VSA software suite [8], and the can be accompanied by text-to-speech produced phrases, using a speech synthesizer by Acapella [9].

D. *The Ibn Sina Android and Center*

The Ibn Sina Robot (figure 4) [3] is the world's first android with Arabic-language conversational abilities. The robot is part of the IbnSina Center, a novel augmented reality interactive theatre installation, named after the famous polymath of the 10th century, known as Avicenna in the West.

Ibn Sina (Abu Ali Ibn Sina, 980-1037) was the foremost philosopher and physician of his times. He was also an astronomer, chemist, geologist, logician, mathematician, poet, psychologist, scientist, and teacher. He wrote more than 450 treatises on a wide range of subjects, and his "Canon of Medicine" [10] was a standard medical text at many European medieval universities, and his work built upon Aristotle and Galen, as well as other ancient and contemporary tradition.

The installation of the Ibn Sina *Interactive Theatre*, (fugure 5) consists of a ten-meter stage, multiple stage sensors, a screen, a pseudo-3D holograph transparency, and a seating area for the audience. The stage is populated by a custom-designed humanlike humanoid robot (the "IbnSina" robot) and by humans and other entities. The screen and the holograph can display static and moving images, virtual environments as well as online virtual worlds populated by characters, or a windowed / blended mix of the above. The robotic and virtual characters could be autonomous, partially-autonomous, puppetteered, scripted, or real-timecontrolled by imitation of human body movements (embodied telepresence). Through the system described in this paper, they can also be remote controlled by thought, via brain-computer interfacing.

Furthermore, multiple other modes of *participation of distant humans* can be supported: not only through videoconference, but also through control of robots and/or virtual characters. An example of teleparticipation through controlling avatars in Twinity (a Second Life-like online virtual world) is shown in figure 6. A part of the projection area forms a window looking into the virtual world, and thus visibility of the virtual world from the real stage is achieved (left part of figure 6). On the other hand, a videofeed is placed on a wall of a virtual room, and thus visibility of the real stage from the virtual world is achieved (right part of figure 6). Thus, essentially a bidirectional window between the real stage and the online virtual world is formed, which together with an audio feed, enables interaction, and thus teleparticipation of remote humans through control of avatars.

Figure 4. Photo of Ibn Sina Android

Figure 5. Ibn Sina Center Interactive Theatre: Stage, Robot, Hologram

Figure 6. Participation of Remote Humans through Online Worlds

The IbnSina center serves as a platform for multiple purposes: artistic, research as well as educational; and most importantly, the centrality of a character such as IbnSina hopefully catalyzes the reconnection of the wider region of the UAE to a past during which scientific inquiry and the arts had flourished; and thus, enables the creation of a future for the region which will emphasize such cultural values.

The Ibn Sina robot has been constructed through an interesting mixture of engineering and art. To achieve humanlike expressivity in the face, a porous elastomer called Frubber was used. This material, being a fluid-filled porous elastomer, better matches the physics in the human facial soft tissues than do conventional elastomer materials. Thus the resulting facial expressions crease and wrinkle better, and require considerably less force to actuate. This allows a full set of simulated facial muscles to be included in the special envelope of the robotic head. In addition to facial expressions, the Ibn Sina robot includes robotic neck action and arms. The arms and neck are actuated using Dynamixel/Maxon motors. The arms contain 14 degrees of freedom each, the neck includes 3, and the face contains 28, for a total of 59 degrees of freedom. The robot contains Sony CCD imaging embedded in the eye. The Ibn Sina hardware allows a high fidelity simulation of a full range of human facial expressions, and the majority of gestural function of human neck and arms. The figure has been decorated with clothing that historically resembles garb of the time of Ibn Sina, and with facial hair and skin coloration that mimics the appearance of Ibn Sina, as known through skull data, miniatures, and descriptions, as closely as possible.

III. RESULTS

In order to evaluate the performance of our proposed thought-controlled robotic telepresence, we used measurements from three subjects. EEG data was acquired using a Biosemi amplifier, from 32 active electrodes mounted on a standard electrode cap, at location based on the international 10/20 system. Data was sampled at 1000Hz. A software based 0.5 Hz high pass filter was used to remove the DC drifts and a 50Hz notch filter was applied to minimize line noise artifacts. These filters were design to be linear-phase to prevent delay distortions. Stimulus events recorded on a separate channel were delayed to match latencies introduced by the digital filtering of EEG.

Each subject, used our thought-controlled robotic teleprsence to issue commands to the robot. First, a 20 minute calibration session was used to train the system's classifier. Then, in the application session, each user issued 92 commands using our system. Subjects could take a break between commands.

We show results on the performance of our system for each subject. Specifically, in figure 5 we show the performance of each subject in terms of correctly issued commands, as a function the trial sequence. All subjects. achieve 100 percent in the probability of correct selection of the commands at the beginning of the application session and for a number of consecutive trials. Subject one, issues 75 consecutive commands correctly before its first error, subject two issues 63 consecutive commands correctly, and subject 3 issues 38 consecutive commands correctly before its first error.

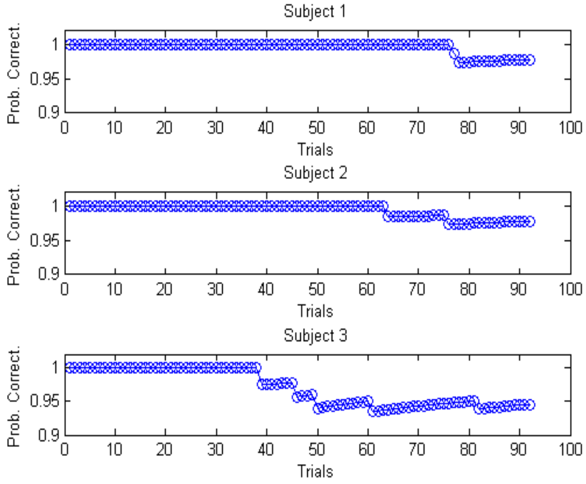

Figure 7. Performance measures for the three subjects. Horizontal axis: shows the order in which the commands where issued. Vertical axis: variation in probability of issuing a correct command over trials.

For all three subjects, the longest consecutive commands' sequences are those beginning from the initiation of the application session. This suggests that, errors might have been due to human factors, such as the subject gets tried or looses attention.

The overall performance of the system in terms of percentage of correctly issues commands is at 97%. Specifically, subjects one and two issued 90 out of 92 commands correctly (or 98 percent), and subject 3, issued 87 out of 92 commands correctly (or 95% percent).

IV. CONCLUSIONS AND FUTURE WORK

We have presented a novel thought-controlled robotic telepresence system, demonstrated in public. The system consists of an EEG-based brain computer interface module (BCI), located in Cyprus, and relying on machine learning methods for recognizing brain patterns and translating them into control-commands. The BCI module is controlling an android robot which is located 1400 miles away in the United Arab Emirates. Control is taking place in real-time, through an IP connection. In the paper, we started by introducing the need for such systems as well as some existing relevant work. We presented the system's methods and modules, as well as the demonstrations, and provided a preliminary evaluation of the system's performance.

In future work, we will evaluate our system on a larger subject set, and open events at malls and public places are being planned at the moment. We also want to investigate opinions and attitudes towards potential real-world applications of relevant technologies, through suitably designed questionnaires, administered to the event participants. Furthermore, we will work on optimizing the brain computer interface to maximize command transfer rates, and minimize training time, as well as retraining time. Also, another important issue is choice of correspondence between thought categories and actions / behaviors of the robot – as well as choice and tuning of feedback channels. This is also highly dependent on the target application – and thus, at a later stage, we would like to investigate this question for a number of selected specific applications.

Finally, we believe that brain-computer interfacing, can also become invaluable not only towards many hard and directly beneficial real-world applications, but when coupled with robots and interactive characters, can also become very useful towards novel and exciting interactive participatory arts. In that respect, we plan to integrate BCI-controlled teleparticipation in a real-world public artistic performance with our robots, in order to expose this interesting potential to a wider audience, and thus provide stimulus for unleashing imagination and creativity.

V. REFERENCES

[1] http://www.irc.atr.jp/en/research_project/humanoid/geminoid/

[2] B. Blankertz et al., "The BCI Competition 2003: Progress and Perspectives in Detection and Discrimination of EEG Single Trials," IEEE Trans. Biomedical Eng., vol. 51, no. 6, 2004, pp. 1044–1051.

[3] N. Mavridis and D. Hanson, "The IbnSina Center: An Augmented Reality Theater with Intelligent Robotic and Virtual Characters", IEEE Ro-MAN 2009

[4] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain-computer interfaces for communication and control. Clin Neurophysiol, 113(6):767-791, June 2002

[5] R.K.I Leaders Limited company website (Cyprus), http://www.rkileaders.com

[6] Interactive Robots and Media Laboratory (IRML) website http://irml.uaeu.ac.ae , and IRML video demo collection at http://www.youtube.com/irmluaeu

[7] Skype videoconferencing software, http://www.skype.com

[8] Visual Show Animation (VSA) from brookshire software LLC, http://www.brookshiresoftware.com

[9] Acapela Group speech and language technologies, http://www.acaperl-group.com

[10] Ibn Sina, The Canon of Medicine. Chicago: Kazi Publ., 1999