ARM-COMS: ARm-Supported eMbodied COmmunication Monitor System

Teruaki Ito¹ and Tomio Watanabe²

¹ The University of Tokushima, 2-1 Minami-Josanjima, Tokushima 770-8506 Japan

² Okayama Prefectural University, 111 Tsuboki, Souja Okayama 719-1197 Japan
ito@me.tokushima-u.ac.jp, watanabe@cse.oka-pu.ac.jp

Abstract. Remote communication systems are getting popular these days, which makes it possible to enjoy audio/video communication over the network in high quality. However, remote communication using these systems is still not identical to face-to-face meeting due to several reasons which are still open issues. This study focuses on two of the issues. One is lack of tele-presence and the other one is lack of connection in communication. In order to tackle these issues, this study proposes an idea of connecting remote individuals through augmented tele-presence systems called ARM-COMS: ARm-supported eMbodied COmmunication Monitor System. ARM-COMS is composed of a desktop robotic arm and a tablet PC which is attached to it. The tablet PC presents audio/video images of a remote person just as typical video conference. However, ARM-COMS controls the physical movement of the tablet PC during the video conference to be consistent with physical motion of the remote participant as if the remote person were there and behaved in a face-to-face conversation. ARM-COMS also considers the meaningful physical position in space to show the connection with other person or topics. The tablet PC approaches to the speaking person as if embodied communication occurs. This paper shows the idea of ARM-COMS and presents some of the on-going work of the study to show the feasibility of the idea.

Keywords: Embodied communication, augmented tele-presence robotic arm manipulation, human interface.

1 Introduction

Network-based communication has been popular these days according as the development of video conference systems [Abowdm 2000]. Even though remote communication is getting popular over the network using these systems, several drawbacks are still regarded as open issues as opposed to a face-to-face meeting, such as lack of tele-presence and lack of relationship in communication [Greenberg 1996]. Using a mobile robot for remote communication with multimodal information, the experimental results in several studies show the effectiveness of these remote controlled robots in communication [Kashiwabara 2012]. Embodiment of an agent using anthropomorphization of an Object [Osawa 2012] is also an interesting idea to

show the presence. These robots could provide tele-presence of the operator in the remote site and even enables some kinds of tele-operating tasks from distance. These robots provide basic function to suppose distance communication using several critical functions such as face image display of the operator [Otsuka 2008], drivability to move around, tele-manipulation on remote objects as well as basic communication functions including talk/listen/see [Kim 2012]. Even a smaller robot is now available on the market for these purposes as well. However, there are still a gap between robot-based video conferences/meetings and face-to-face ones, which is an open issue to be solved for better communication. A new challenge is undertaken by a robotic arm type system with mobile function [Tariq 2011; Wongphathi 2012]. For an example of non-mobile arm type system, Kubi [Revolve Robotics] allows the remote user to “look around” during their video call by commanding Kubi where to aim the tablet using intuitive remote controls over the web. However, non-verbal movement of the remote person is still an open issue.

This study proposes an idea for connecting remote individuals through augmented tele-presence systems called ARM-COMS (ARm-supported eM-bodied COmmunication Monitor System), focusing on the two issues; lack of tele-presence and lack of relationship in communication. ARM-COMS is composed of two sub-systems: a tablet PC for video communication and a robotic arm to control physical motion of the tablet PC as an avatar. As opposed to the manual control of Kubi, ARM-COMS makes the tablet PC physically behave as if the remote person were actually there by way of movement and positioning of the tablet PC. Fig.1 shows an illustrative image of ARM-COMS, which enables the behavior of a tablet PC as an avatar of a remote person.

Fig. 1. Overview of ARM-COMS application in vide conference

This paper shows the basic idea of ARM-COMS and presents the on-going work of this study.

2 A Proposal of an Idea ARM-COMS

This study proposes an idea of ARM-COMS, which is capable of movements similar to a human neck and is used for mimicking a person's movements. ARM-COMS traces the head movement of a person during conversation, understands the meaning of motions and makes the tablet PC physically behave as if the person were actually there. As opposed to the conventional tele-presence robots, ARM-COMS pursues the realization of entrainment in video conversation with a remote person. The idea of ARM-COMS challenges two types of issues. One is motion-control to mimic the movement of a remote person and to enable entrainment in conversation. The other one is position-control to show the dynamic relationship with participants in conversation [Okada 1994] or to show the status of interest in topics under discussion.

Fig. 2. An overview of ARM-COMS to support remote communication

— Challenge 1: Entrainment movement control

It has been reported that entrainment among participants emerges during conversation if the participating persons get together in the same physical space and engage in the conversation. [Okada 1994; Watanabe 2011] Tracking the head movement of a speaking person, ARM-COMS mimics the movement using a tablet PC by a robotic arm and challenges to enable this control function in remote communication.

— Challenge 2: Entrainment position control

In a face-to-face meeting, each person takes a meaningful physical position to represent the relationship with the other participants, or to send a non-verbal message to others. A closer position would be taken for friends, showing close relationship, whereas a non-closer position would be taken for strangers, showing unfriendly relationship [Osawa 2012]. ARM-COMS controls a tablet PC to dynamically locate an appropriate position in space and to explicitly represent the relationship with other participants, sending a non-verbal message. The tablet PC would be approaching to the speaking person to show that the remote person is interested in the talk.

This study tackles these two types of challenges to implement the proposed idea as ARM-COMS. However, this paper mainly focuses on the challenge 1 and presents the on-going work. The challenge 2 will be presented in a separate paper.

3 ARM-COMS: Motion Control Experiment

3.1 Basic Structure of Motion Control

During conversation, various types of body/head movements can be observed. In order to mimic some of these movements, this study focuses on three types of head movements, namely, nodding, head-tilting, and head shaking movements. All of these are very typical non-verbal expression in Japan during conversation. Nodding means affirmative, agree, listening, etc. Head-tilting means ambiguous, not sure, impossible to answer, etc. Head shaking means negative, disagree, no way, etc. Fig.3 shows the corresponding physical motions implemented by the robotic arm control. If the monitor behaves like these in conversation, it is assumed that the physical movements could send a non-verbal message. Technically speaking, these three types of movements can be regarded as the rotation around each axis as shown in Fig. 3. Therefore, the rotation angles of these three motions can be calculated as in Scheme (1), (2) and (3).

Fig. 3. Three types of target head motion

For nodding movement, roll angle can be calculated by scheme (1), where α is the rotation angle around y-axis, and X,Z are the acceleration value for each direction.

$$\alpha = \sin^{-1} \left(\frac{-X}{\sqrt{Y^2 + Z^2}} \right) (3-1) \quad (1)$$

For “I am not sure” movement, pitch angle β can be calculated by scheme (2), where β is the rotation angle around x-axis, and Y,Z are the acceleration value for each direction.

$$\beta = \sin^{-1} \left(\frac{-Y}{\sqrt{Y^2 + Z^2}} \right) (3-2) \quad (2)$$

For sideways shaking movement, yaw angle ξ can be calculated by scheme (3), where ξ is the rotation angle around z-axis, ω is the angle subtracted from the horizontal angle and Δt is the sampling interval, which was 100[ms] here.

$$\xi = \sum \omega \Delta t \quad (3-3) \quad (3)$$

3.2 Prototype System

A prototype robotic arm system was built to mimic the head motions based on the scheme (1), (2) and (3). Fig. 4 shows the overview of the system architecture of the prototype.

Fig. 4. An overview of ARM-COMS prototype and its motion by Wii remote

The robotic arm system is composed of a table top robotic arm (Lynxmotion) with motor controller board (SSC-32 Ver.2.0) which is connected to PC (Windows 7) by serial cable. The robotic arm is controlled by PC using a remote controller (Wii Remote + Wii Motion Plus) through Bluetooth connection. The combination of Wii Remote and Wii Motion plus makes it possible to trace the acceleration for three axis

and three rotation angle around these three axis. Software to control the arm was developed by Visual C++ with library Wiimotelib v.1.8. As a result, the robotic arm is wirelessly controllable by Wii Remote manipulation.

3.3 Study 1: Embodied Proxy Platform Using a Robotic Arm

According to the feasibility test of the prototype, it was recognized that the prototype could mimic the head motion if the Wii remote is attached to the head. However, Wii remote is not suitable to attach to the body of human. Therefore, another type of wireless acceleration sensor (WAA-001, ATR-Promotion, Bluetooth type) was used. Since this sensor detects only acceleration value for three axes, a pair of sensors was used to cover the target three motions. One was attached to the ear portion of the head, while the other one was attached to the neck. Fig.5 shows the nodding, head-tilting and lean forward motion, all of which were mimicked by the prototype ARM-COMS.

Fig. 5. Master-slave motion of ARM-COMS system to demonstrate the physical behavior

ARM-COMS mimicked the basic motions as shown in Fig.5, where a pseudo-display attached to the arm follows the head's motion. However, head-shaking motion was not available. Therefore, another type of sensor (TSND-121, ATR Promotion)

was applied. TSND-121 is an integrated sensor composed of InvenSense MPU6050 which covers acceleration and angular velocity, as well as AMI306 which covers gyro motion. By modifying the control software to be applicable to TSND-121, the ARM-COMS prototype enabled the target three-motion only by a single TSND-121, was used for the experiment in Study 2 presented in the next section.

3.4 Study 2: Feasibility Study of Video Communication Using ARM-COMS Idea

Study 2 was designed to compare the video communication with and without ARM-COMS to make clear the effectiveness of the idea of ARM-COMS.

4 pairs of participants were recruited from students and two sets of video conversation were conducted by those paired students. One conversation is just a regular Skype communication and the other one is based on ARM-COMS idea. Fig.6 shows the overview of the experimental setup. The video communication was conducted between Site-A and Site-B by a pair of Subject-A and Subject-B, who were located in separate places.

Site-A is regarded as a local site where ARM-COMS is installed. For a tentative purpose, a smart phone is equipped to the robotic arm to be used as a pseudo-active display. Subject-A communicates with Subject-B via Skype on the smart phone. A magnetic sensor (Fastrak, POLHEMUS) is attached to the head of Subject-A during conversation to detect the head motion of subject-A.

Site-B is regarded as a remote site where Subject-B communicates with Subject-A in Site-A through video communication (Skype) on a laptop PC. A multi-sensor TSND-121 was attached to the head of Subject B to trace the head movement during conversation, which was also used to control ARM-COMS in Site-A. The sensing data from the multi-sensor was transmitted to the client program in the laptop through Bluetooth. The socket program communicates with the server program in desktop PC on Site-A, and controls the ARM-COMS via Wi-Fi network. The collecting data from the multi-sensor was also used to analyze the head movement of Subject-B during video conversation.

Subject-A and Subject-B were video-recorded in both Site-A and Site-B during the whole conversation, which was used to synchronize the head movements of Subject-A and Subject-B in the conversation.

Several types of topics for video conversation were prepared from daily conversation. For example, “Which animal do you like?”, “What did you eat for the breakfast?”, “How did you come to the university today?”, “Which convenience store do you often go? ”, etc. Each pair of subject was asked to perform one-minute video conversation about one of the selected topics twice. First conversation was without the robotic arm manipulation, which was similar to regular Skype conversation, and the second one was with arm manipulation, which was identical to the idea of ARM-COMS.

Fig. 6. An overview of ARM-COMS to support remote communication

3.5 Results and Discussion

Video conversation was smooth in any of the four pairs with or without AMR-COMS. It was observed that most of the four pairs conducted their conversation without using dynamic head movement simply because it was what they normally do in regular conversation.

Figure 7 shows an example of comparison data between with and without ARM-COMS during conversation. As opposed to the dynamic head movement of Subject B, that of Subject A was stable. Entrainment of the head was not significantly observed. However, synchronization of the two subjects was observed in the bowing at the beginning and at the ending of the conversation.

For smooth conversation, a topic memo was given to each subject. As a result, each subject quite often looked down the memo in the conversation, rather than looking at the counterpart in the video screen. Subject did not use the memo at the greetings in the beginning and in the end. This made it possible to perform the conversation in a smooth manner. However, this might have hindered the entrainment with the counterpart. Future works needs further discussion on the experimental setup.

Fig. 7. Comparison data of video communication with and without ARM-COMS

4 Concluding Remarks

The original idea of this study is aiming at how to control the physical position and movement of a tablet PC in such a way that a remote participant would most likely behave if he/she were actually there. Active display presents the tele-existence of a remote object shown in the display by physical movement. However, ARM-COMS not only presents the tele-presence of a remote person, but also explicitly shows the relationship between the remote person and the local participants by way of the entrainment behavior of a table PC.

ARM-COMS employs only a general tablet PC attached to the sub-system of robotic arm, which will be specifically designed and built for this purpose. ARM-COMS not only presents a new idea for remote communication system, but also opens a potential new market for non-industrial robotic arm design and products.

This paper proposed the basic idea of ARM-COMS, which is based on the combination of entrainment movement and entrainment positioning, and showed the preliminary result of feasibility in entrainment movement. The future works include the design and manufacturing of sub-system for ARM-COMS robotic arm, development of the entrainment movement/positioning algorithm and its feasibility study.

References

1. Abowdm, D.G., Mynatt, D.E.: Charting past, present, and future research in ubiquitous computing. *ACM Transactions on Computer-Human Interaction (TOCHI)* 7(1), 29–58 (2000)
2. Greenberg, S.: Peepholes: low cost awareness of one's community. In: *Conference Companion on Human Factors in Computing Systems: Common Ground*, Vancouver, British Columbia, Canada, pp. 206–207 (1996)
3. Kashiwabara, T., Osawa, H., Shinozawa, K., Imai, M.: TEROOS: a wearable avatar to enhance joint activities. In: *Annual Conference on Human Factors in Computing Systems*, pp. 2001–2004 (May 2012)
4. Kim, K., Bolton, J., Girouard, A., Cooperstock, J., Vertegaal, R.: TeleHuman: Effects of 3D Perspective on Gaze and Pose Estimation with a Life-size Cylindrical Telepresence Pod. In: *Proc. of CHI 2012*, pp. 2531–2540 (2012)
5. Kubi, <http://revolverobotics.com/meet-kubi/>
6. Okada, K., Maeda, F., Ichikawa, Y., Matsushita, Y.: Multiparty videoconferencing at virtual social distance: MAJIC design. In: *SCW 1994 Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work (CSCW 1994)*, pp. 385–393 (1994)
7. Osawa, T., Matsuda, Y., Ohmura, R., Imai, M.: Embodiment of an agent by anthropomorphization of a common object. *Web Intelligence and Agent Systems: An International Journal* 10, 345–358 (2012)
8. Otsuka, T., Araki, S., Ishizuka, K., Fujimoto, M., Heinrich, M., Yamato, J.: A Realtime Multimodal System for Analyzing Group Meetings by Combining Face Pose Tracking and Speaker Diarization. In: *Proc. of the 10th International Conference on Multimodal Interfaces (ICMI 2008)*, Chania, Crete, Greece, pp. 257–264 (2008)
9. Sirkin, D., Ju, W.: Consistency in physical and on-screen action improves perceptions of telepresence robots. In: *HRI 2012, Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, pp. 57–64 (2012)
10. Tariq, A.M., Ito, T.: Master-slave robotic arm manipulation for communication robot. In: *Proceedings of 2011 Annual Meeting on Japan Society of Mechanical Engineer*, vol. 11(1), p.S12013 (September 2011)
11. Tomotoshi, M., Ito, T.: A study on awareness support method to improve engagement in remote communication. In: *The first International Symposium on Socially and Technically Symbiotic System (STSS 2012)*, Okayama, vol. 39, pp. 1–6 (August 2012)
12. Watanabe, T.: *Human-entrained Embodied Interaction and Communication Technology, Emotional Engineering*, pp. 161–177. Springer (2011)
13. Watanabe, T., Okubo, M., Nakashige, M., Danbara, R.: InterActor: Speech-Driven Embodied Interactive Actor. *International Journal of Human-Computer Interaction* 17(1), 43–60 (2004)
14. Wongphathi, M., Matsuda, Y., Osawa, H., Imai, M.: Where do you want to use a robotic arm? And what do you want from the robot? In: *International Symposium on Robot and Human Interactive Communication*, pp. 322–327 (September 2012)