# Three Key Challenges in ARM-COMS for Entrainment Effect Acceleration in Remote Communication

Teruaki Ito[1] and Tomio Watanabe[2]

[1] The University of Tokushima, 2-1 Minami-Josanjima, Tokushima 770-8506, Japan
tito@tokushima-u.ac.jp
[2] Okayama Prefectural University, 111 Tsuboki, Souja, Okayama 719-1197, Japan
watanabe@cse.oka-pu.ac.jp

**Abstract.** Remote communication systems, which are getting popular these days, allow us to enjoy the benefit of audio/video communication over the network. However, communication based on these systems is still not identical to face-to-face meetings. For example, open issues include lack of tele-presence, lack of entrainment in communication, etc. In order to tackle these issues, this study proposes an idea of remote individuals' connection through augmented tele-presence systems called ARM-COMS: ARm-supported eMbodied COmmunication Monitor System. ARM-COMS is composed of a tablet PC as an ICT (Information and Communication Technology) device and a desktop robotic arm which manipulates the tablet. Two types of modes, or intelligent tablet mode (IT-mode) and intelligent avatar mode (IA-mode), play a key role in ARM-COMS to implement the three functions; namely, autonomous positioning (AP), autonomous entrainment movement (AEM), and autonomous entrainment positioning (AEP). This paper proposes the basic concept of ARM-COMS to accelerate the entrainment effect in remote communication.

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

## 1 Introduction

Thanks to the development of ICT (Information and Communication Technology) technologies and the expanding internet connection services, remote communication systems are now one of the popular applications today. Apart from the high quality commercial systems at the top, many of the application software for remote communication are freely downloadable. [Abowdm et al. 2000]. Even though remote communication is getting popular over the network, several drawbacks are still unsolved, such as lack of tele-presence and lack of relationship in communication [Greenberg et al. 1996].

As for tele-presense issues, an idea of mobile robot-based remote communication proposes one solution. Using a mobile robot for remote communication, the experimental results in several studies show the effectiveness of these remote controlled robots in communication [Tariq et al. 2011; Kashiwabara et al. 2012]. Embodiment of

S. Yamamoto (Ed.): HIMI 2014, Part I, LNCS 8521, pp. 177–186, 2014.

an agent using anthropomorphization of an Object [Osawa, et al. 2012] is also an interesting idea to show the presence [Sirkin et al 2012].

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 support distance communication using several critical functions such as face image display of the operator [Otsuka et al. 2008], drivability to move around, tele-manipulation on remote objects as well as basic communication functions including talk/listen/see [Kim et al. 2012]. However, there are still a gap between robot-based video conferences/meetings and face-to-face ones.

A new challenge was undertaken by a robotic arm type system with mobile function [Wongphati et al. 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. An idea of enhanced motion display [Otsuka et al. 2011] has been reported. The effectiveness of dynamic motion of display to represent the physical object has been reported [Yakuyama et al, 2011]. However, it has not been applied to the movement of human body. Therefore, non-verbal movement of the remote person is still an open issue.

This study focuses on the critical aspect of entrainment in communication [Watanabe 2011]. This paper proposed an idea for connecting remote individuals through augmented tele-presence systems called ARM-COMS(ARm-supported eMbodied COmmunication Monitor System), focusing on the two issues; lack of tele-presence and lack of relationship in communication [Ito et al. 2013]. However, considering the advantages of tablet PC as one of the mobile ICT devices, the idea of ARM-COMS has been upgraded by integrating IT-mode in addition to IA-mode. Therefore, this paper proposes an updated idea of ARM-COMS to cover the whole idea.

## 2 A Proposal of an Idea of Active Monitor Arm for Augmented Tele-presence – ARM-COMS

This study proposes an idea of augmented tele-presence systems, which is called ARM-COMS. ARM-COMS is designed to integrate the two components (a tablet PC and a robotic arm) with two manipulation modes: intelligent tablet mode (IT-mode) and intelligent avatar mode (IA-mode). The tablet PC is integrated into ARM-COMS for video communication manipulated by an active monitor arm as well as for information retrieval by using it as a general information device. IT-mode and IA-mode are interchangeable based on the situation of the user. This section covers the two basic modes of ARM-COMS.

### 2.1 IT-Mode in ARM-COMS

The tablet PC is one of the very popular mobile ICT devices today. As a typical situation in using a tablet PC, a user holds the devise in left hand and manipulates it on the touch screen with right fingers. However, if both hands are not free, it is not suitable to use it this way. If a user is in bed, for example, it would not be comfortable to use it this way, either. Considering the characteristic feature of tablet PC as a mobile device, it would lose this feature if a tablet PC is placed on a desktop holder, which is often seen these days. Since the tablet PC is a convenient mobile tool, it would be an ideal situation that our own tablet autonomously and automatically approaches to us when we need it and where we need it even if we do not do anything. This is what IT-mode of ARM-COMS is aiming at.

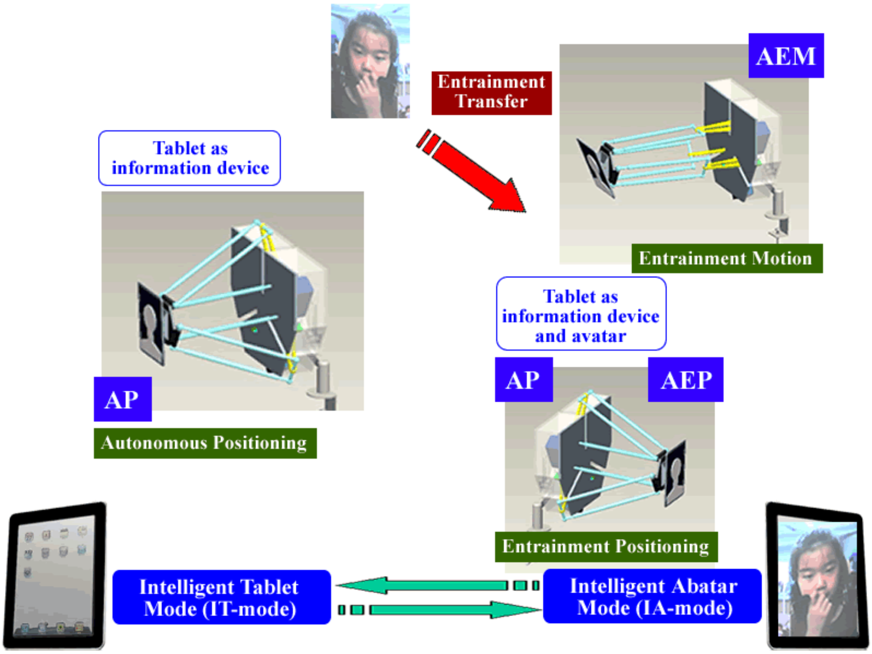


**Fig. 1.** Overview of ARM-COMS to support remote communication

## 2.2 IA-Mode in ARM-COMS

As mentioned above, the tablet PC is one of the very popular mobile ICT devices, it allows us not only to retrieve information, but also to communicate with others over the network. As a typical situation in using a tablet PC for communicating with others, the user holds the devise in left hand and manipulates the touch screen with right fingers, which is quite similar to that of IT-mode operation mentioned above. However, when we compare video communication with face-to-face communication, there is a significant difference from the two points of view. When we talk with somebody in a face-to-face meeting, what we share is not merely the same physical space, but also an invisible communication space or atmosphere. As a result, entrainment will occur between the participants of the conversation. However, when we talk with somebody over the network, we can only see the face on the screen and cannot share the same physical space. As a result, this kind of entrainment is different from that of a face-to-face meeting. Since the entrainment is associated with physical movement of a person, a dynamic movement of a tablet PC during remote communication will make some effect on entrainment acceleration. In addition to the physical movement of a person, the physical distance between things or people implicitly expresses a relationship of people. This is what IA-mode of ARM-COMS is aiming at.

## 3 Three Key Challenges in ARM-COMS

This study proposes an idea of ARM-COMS, which behaves like the human neck and mimics a person's movements to play a role as an avatar. In addition to that, ARM-COMS understands the intention of users so that it behaves like the human hand to offer something when they want it and where they want it in a timely manner. In order to implement this idea, this research is tackling the three key challenges as mentioned below. As mentioned in section 2, ARM-COMS operates on two types of operation mode, or IT-mode and IA-mode. Therefore, the three challenges are based on these two modes.

### 3.1 Challenge 1: Autonomous Position Control

Since the tablet PC is a convenient tool, it would be an ideal situation that our own tablet autonomously and automatically approaches to us when we need it and where we need it as if the tablet PC understands what we want. For example, suppose a user is working at a desk and receives an incoming Skype call. Considering what the use is doing, ARM-COMS autonomously takes the table PC in front of the user to urge the acceptance of the connection. Challenge 1 pursues that ARM-COMS dynamically locates the tablet PC autonomously at a convenient and comfortable position to the user when they need it and where they need it.

### 3.2 Challenge 2: Autonomous Entrainment Movement Control

Challenge 1 does not directly relate to video communication. However, Challenge 2 and 3 are directly related to video communication usage. It has been reported that entrainment among participants emerges during conversation if the participating subjects share the same physical space and engage in the conversation [Okada et al. 1994; Watanabe et al. 2004]. However, this kind of entrainment in a face-to-face meeting is different from that of remote communication. Tracking the head movement of a speaking person in a remote site, ARM-COMS manipulates the tablet PC [Tomotoshi et al. 2012; Wongphati et al. 2012] as an avatar to mimic the head movement of the remote person so that entrainment emerges as if the local person interacts with the remote person locally.

### 3.3 Challenge 3: Autonomous Entrainment Position Control

In a face-to-face meeting, each person takes a meaningful physical position to represent the relationship with the others, or to send non-verbal messages 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, by sending non-verbal messages. For example, the tablet PC would be approaching to the speaking person to show that the remote person is interested in the talk.

# 4 Design of AEM in ARM-COMS

## 4.1 Basic Motion Control for ARM-COMS AEM

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-titling 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).

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.

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

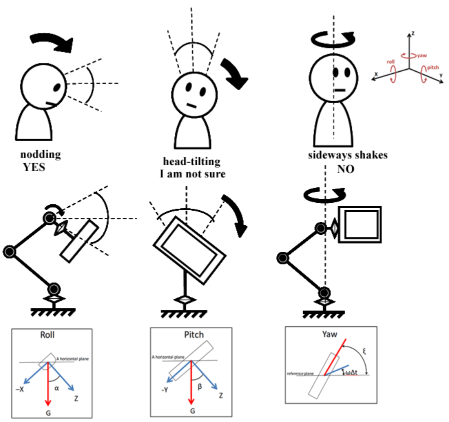


**Fig. 2.** Three types of target head motion

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

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

For sideway shaking movement, yaw angle ξcan be calculated by scheme (3), whereξis the rotation angle around z-axis, ω is the angle subtracted from the horizontal angle andΔt is the sampling interval, which was 100[ms] here.

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

### 4.2 Prototype System for ARM-COMS AEM

A prototype robotic arm system for AEM in ARM-COMS was configured as shown in Fig.3. 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 a serial cable.

First, this robotic arm was controlled by a PC using a remote controller (Wii Remote + Wii Motion Plus) through Bluetooth connection. The combination of Wii Remote and Wii Motion plus made it possible to trace the acceleration of the neck at the three axis and three rotation angle around these three axis. Control software for the arm was developed by Visual C++ with library Wiimotelib v.1.8. As a result, the robotic arm was wirelessly controllable by Wii Remote manipulation.

According to the feasibility test of the first prototype, it was recognized that the prototype could mimic the head motion if the Wii remote was attached to the head. However, Wii remote was not appropriate to attach to the body of human. Therefore, a 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. Then, another type of sensor (TSND-121, ATR Promotion) was applied instead, in order to reduce the number of sensors. TSND-121 is an integrated sensor composed of InvenSence MPU6050 which covers acceleration and angular velocity, as well as AMI306 which covers gyro motion. By modifying the control software applicable to TSND-121, the ARM-COMS prototype enabled the target to control in three-motion only by a single TSND-121.

Fig.3 also shows the nodding, head-tilting and lean forward motion, all of which were mimicked by the prototype ARM-COMS, where a pseudo-display attached to the arm follows the head’s motion.

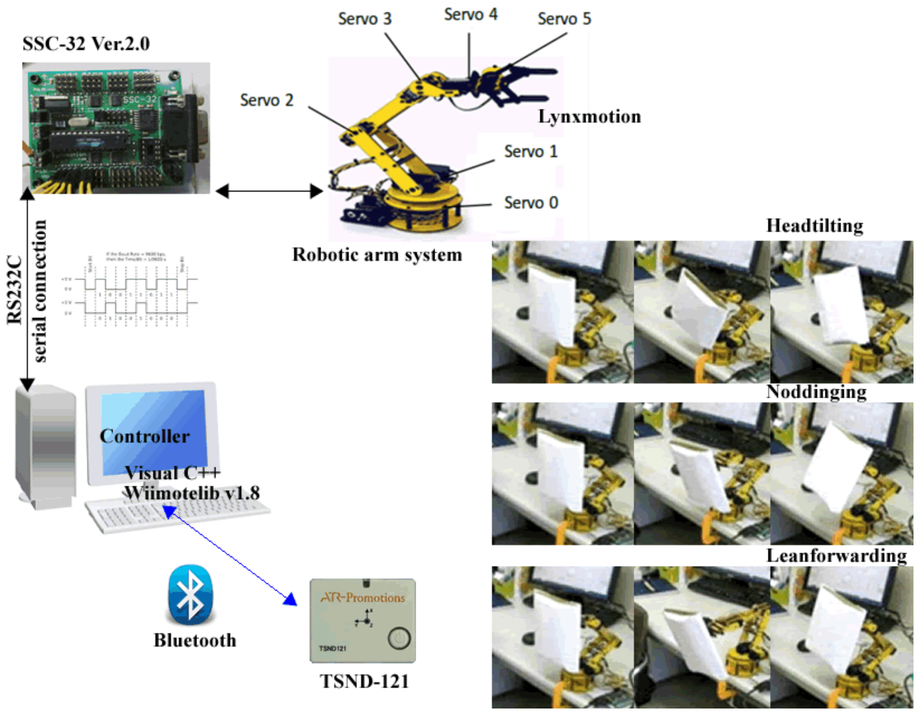


**Fig. 3.** Overview of ARM-COMS AEM prototype and its motion

Fig. 4 shows the flow diagram of server-client control software which was developed to control the robotic arm for ARM-COMS AEM.

A server process is started on a PC as a detached process of socket at remote site where ARM-COM AEM is connected. When the client PC starts the client program, it sends a request to the socket for connection. When the server program receives the request signal, it will accept the connection. Once the connection is established between the server and the client, the controlling command is sent to the remote PC. The command is generated based on Scheme (1) – (3) using the sensing data obtained from the integrated sensor, TSND-121. The remote PC controls the ARM-COMS based on the command received from the local PC. Three types of head movement were shown in Fig.3.

### 4.3 Feasibility Study of ARM-COMS AEM

Feasibility tests of ARM-COM AEM were conducted to compare the video communication with and without ARM-COMS to make clear the effectiveness of the idea of ARM-COMS between the two different places, or Site-A and Site-B.

Site-A was regarded as a local site where ARM-COMS was installed equipped with a smart phone as a pseudo-active display. Subject-A communicates with Subject-B via Skype on the smart phone. A magnetic sensor (Fastrak, POLHEMUS) was

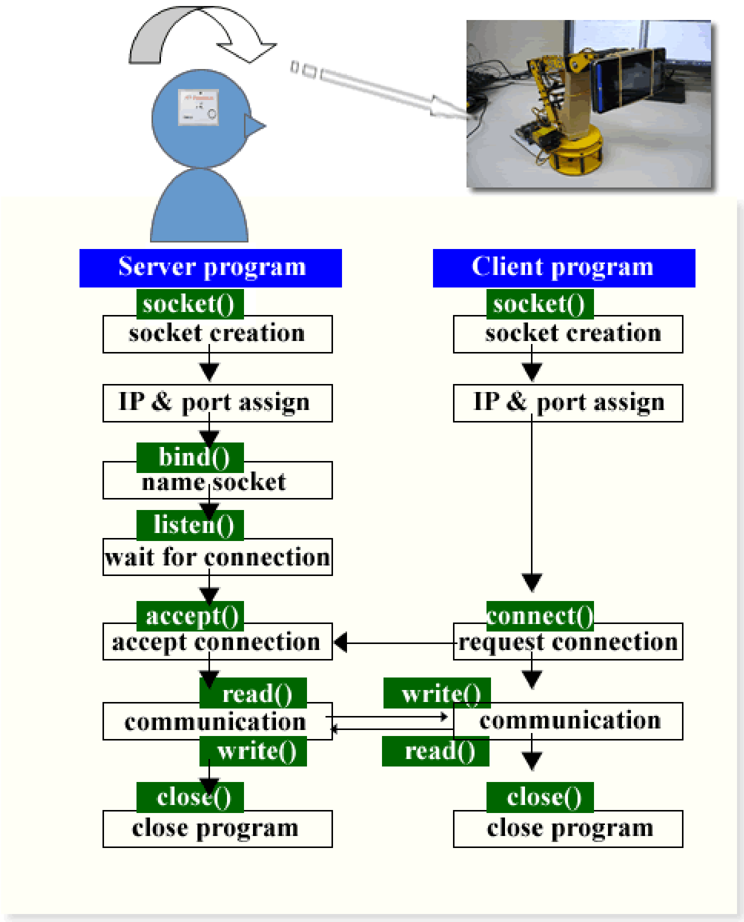


**Fig. 4.** Overview of client-server program to control ARM-COMS

attached to the head of Subject-A during conversation to detect the head motion of subject-A. Site-B was regarded as a remote site where Subject-B communicates with Subject-A in Site-A via 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. In this way, a remote communication environment was set up to conduct the feasibility tests.

Feasibility experiments in remote communication with/without ARM-COM AEM were conducted. Based on the video recording data for the movement of subjects, head movement data during the conversation, synchronization data between the subjects, etc, feasibility of ARM-COMS AEM was recognized as generally positive.

## 5 Concluding Remarks

The paper proposed an idea of active monitor arm named ARM-COMS with the two types of modes in ARM-COM system and described the three challenges based on these modes. Namely, autonomous position control, autonomous entrainment movement control, and autonomous entrainment movement control. 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 entrainmental 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.

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. Ito, T., Watanabe, T.: ARM-COMS: Arm-supported embodied communication monitor system. In: Yamamoto, S. (ed.) HIMI/HCII 2013, Part III. LNCS, vol. 8018, pp. 307–316. Springer, Heidelberg (2013)
4. 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)
5. 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)
6. Kubi, http://revolverobotics.com/meet-kubi/
7. 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)
8. Osawa, T., Matsuda, Y., Ohmura, R., Imai, M.: Embodiment of an agent by anthropo morphization of a common object. Web Intelligence and Agent Systems: An International Journal 10, 345–358 (2012)
9. Ohtsuka, S., Oka, S., Kihara, K., Tsuruda, T., Seki, M.: Human-body swing affects visibility of scrolled characters with direction dependency. In: 2011 Symposium Digest of Technical Papers, Society for Information Display (SID), pp. 309–312 (2011)

10. 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)
11. 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)
12. Tariq, A.M., Ito, T.: Master-slave robotic arm manipulation for communication robot. Japan Society of Mechanical Engineer, Proceedings of 2011 Annual Meeting 11(1), S12013 (2011)
13. Tomotoshi, M., Ito, T.: A study on awareness support method to improve engagement in remote communication. In: First International Symposium on Socially and Technically Symbiotic System (STSS 2012), Okayama, vol. 39, pp. 1–6 (August 2012)
14. Watanabe, T.: Human-entrained Embodied Interaction and Communication Technology. In: Emotional Engineering, pp. 161–177. Springer (2011)
15. 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)
16. Wongphati, 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)
17. Yakuyama, H., Tsunami, Y., Tadakuma, R.: Evaluation of an enhanced-motion display. In: SICE Tohoku, ch. 267-17, pp. 1–5 (2011) (in Japanese)