Human-Entrained E-COSMIC: Embodied Communication System for Mind Connection

Tomio Watanabe

Faculty of Computer Science and System Engineering, Okayama Prefectural University
111 Kuboki, Soja, Okayama, 719-1197 Japan
watanabe@cse.oka-pu.ac.jp

Abstract. The embodied communication system for mind connection (E-COSMIC) has been developed by applying the entrainment mechanism of the embodied rhythms of nodding and body movements to physical robots and CG characters. E-COSMIC comprises an embodied virtual communication system for human interaction analysis by synthesis and a speech-driven embodied interaction system for supporting essential human interaction and communication based on the analysis that uses the embodied virtual communication system. Human-entrained embodied interaction and communication technology for an advanced media society is introduced through some applications of E-COSMIC.

Keywords: Human Interface, Human Communication, Human Interaction, Enbodied Communication, Enbodied Interaction, Enbodied Interface.

1 Introduction

In a human face-to-face conversation, embodied rhythms between speech and body motions such as nodding are mutually synchronized not only between talkers but also in a talker. The phenomenon is observed in mother-infant interaction as a primitive form of communication [1], [2]. This synchrony of embodied rhythms in communication, referred to as entrainment, generates the sharing of embodiment in human interaction, which plays an important role in human interaction and communication. Entrainment in communication is also observed in physiological indices such as respiration and heart rate variability [3]. This embodied communication closely related to behavioral and physiological entrainment is an essential form of communication that forms the relation of interaction between talkers through mutual embodiment. Hence, the introduction of this mechanism to human interface is indispensable to the realization of human-centered essential interaction and communication systems.

In this paper, by focusing on the embodied entrainment, the embodied interaction and communication technology through the development of the embodied communication system for mind connection (E-COSMIC) is proposed for supporting human interaction and communication [4]. E-COSMIC mainly comprises an embodied virtual face-to-face communication system and a speech-driven embodied

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2007, LNCS 4557, pp. 1008 – 1016, 2007.

interaction system, as shown in Fig. 1. The former is developed for human interaction analysis by synthesis and the latter, for supporting human interaction and communication based on the analysis that uses the former. The effectiveness of the system is demonstrated by some actual applications on CG/robot and human interactive communications.

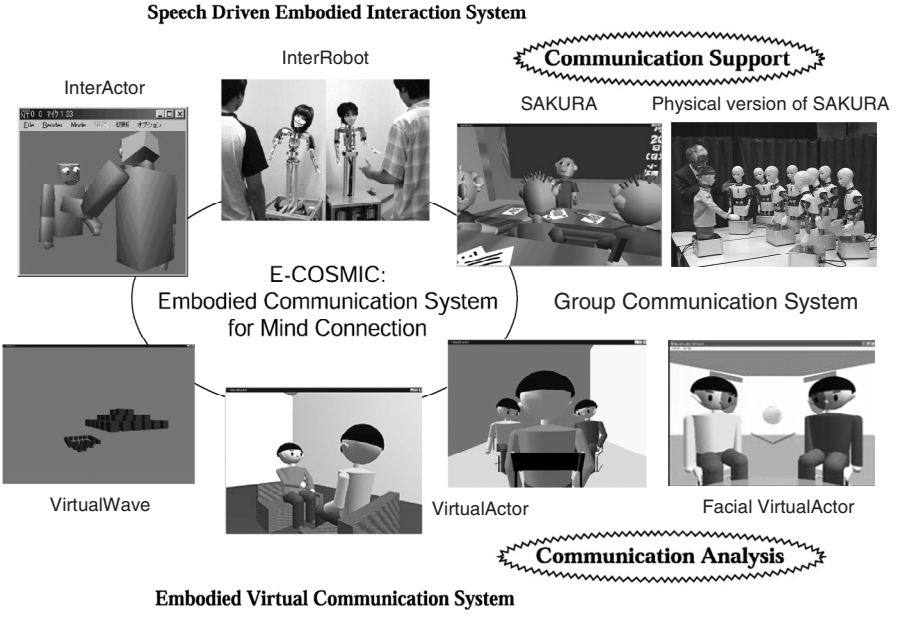

Fig. 1. E-COSMIC

2 Embodied Virtual Communication System

The concept of an embodied virtual face-to-face communication system is illustrated in Fig. 2. The figure presents a VirtualActor (VA), an interactive avatar, that represents the talker's interactive behavior such as gesture, nodding, blinking, facial color and expressions, paralanguage, respiration, etc., based on one's verbal and nonverbal information as well as physiological information in a virtual face-to-face communication environment. Fig. 3 provides an example of a virtual face-to-face scene with two VAs from the diagonal backward viewpoint of one's own VA. The motions of the head, arms, and body for each VA are represented based on the positions and angles measured by four magnetic sensors that are placed on the top of the talker's head, both wrists, and the back of the body [5]. Two remote talkers can communicate through their VAs and become aware of the interaction through the embodied interaction of VAs in the same virtual communication environment from any viewpoint. The analysis by synthesis for interaction in communication is performed by processing the behavior of VAs, such as cutting or delaying the motion and voice of VAs in various conditions of their spatial relations and positions. For

example, to examine the effects of only nodding on interaction, it is possible for a VA to represent just nodding without body motion even if the talker nods with body motion. Thus, the characteristics and relations of the embodied interaction between talkers are systematically clarified through the analysis by synthesis of interaction in communication by using the system in which talkers are the observers of interactions as well as the operators of interaction through their VAs. Further, physiological measurements such as respiration, heart rate variability, and facial skin temperature, as indices of emotional states in communication are utilized not only for quantitatively evaluating the interaction but also for transmitting the change in talkers' emotions through the VA affect display in which facial color and expressions are synthesized based on the measurement. An embodied virtual group communication system is also developed for three human interaction supports and analyses by synthesis, as indicated in Fig. 1 [6].

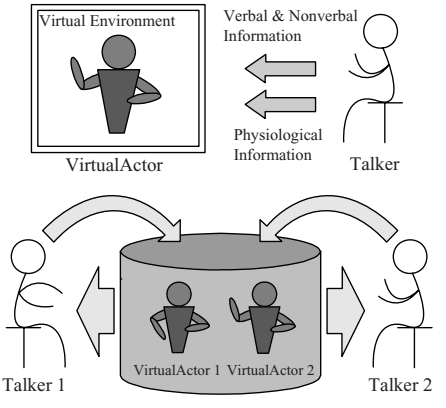

Fig. 2. Concept of the embodied virtual communication system

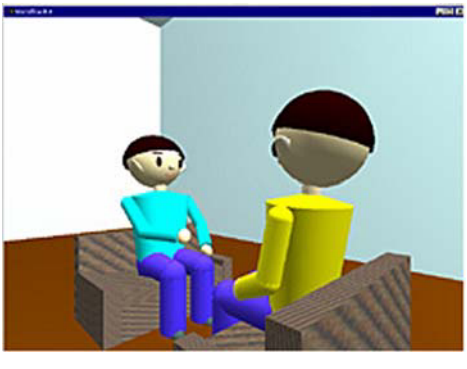

Fig. 3. Example of a virtual face-to-face scene with two VirtualActors representing the talker's own self and his/her partner

3 Speech-Driven Embodied Interaction System

Based on the human interaction analysis that uses the embodied virtual communication system, a speech-driven embodied interaction system is developed for supporting human interaction by generating the communicative motions of a CG character known as InterActor or a physical robot referred to as InterRobot; these communicative motions are coherently related to speech input [7]. The concept and system are presented in Fig. 4. The system comprises two InterActors (InterRobots) that have both the functions of a speaker and a listener based on speech input. When Talker 1 speaks to InterActor 2, InterActor 2 responds to Talker 1's utterance with an appropriate timing through its entire body motions, including nodding, blinking, and actions, in a manner similar to the body motions of a listener. Thus, Talker 1 can talk smoothly and naturally. Subsequently, the speech is transmitted via a network to the remote InterActor 1. InterActor 1 can effectively transmit Talker 1's message to Talker 2 by generating the body motions similar to those of the speaker based on the time series of the speech and by simultaneously presenting both the speech and the entrained body motions. This time, Talker 2 in the role of a speaker achieves communication in the same way by transmitting his/her speech via InterActor 1 as a listener and InterActor 2 as the one talking to Talker 1. Thus, in this manner, two remote talkers enjoy a conversation via InterActors. The information transmitted and received by this system is only through speech. Of significance is the fact that it is a human who transmits and receives the information; the InterActor/InterRobot merely generates the entrained communicative movements and actions based on speech input and supports the sharing of mutual embodiment in communication.

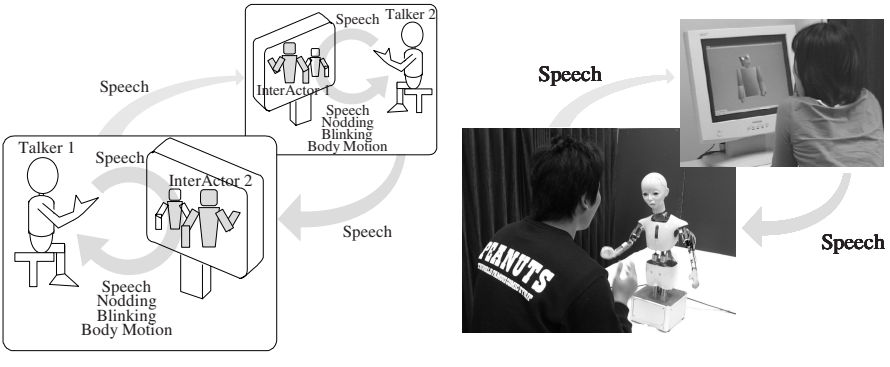

Fig. 4. Concept and system of the speech-driven embodied interaction system

Fig. 5 illustrates a speech-driven embodied group-entrained communication system referred to as SAKURA [8]. SAKURA activates group communication in which InterActors are entrained to one another as a teacher and some students in the same virtual classroom. By using SAKURA, talkers can communicate with a sense of unity through the entrained InterActors by using only speech input via the network. Fig. 6 depicts a physical version of SAKURA with InterRobots. Their entrained movements and actions based on speech can activate and assist human embodied interaction and

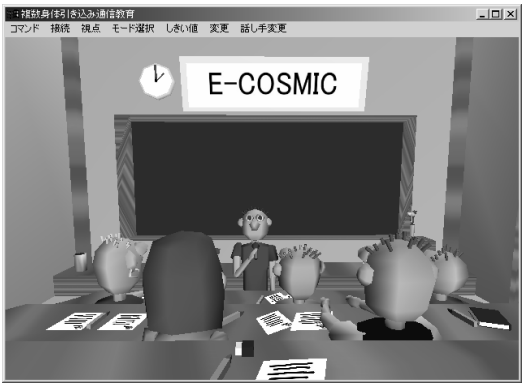

Fig. 5. SAKURA: The speech-driven embodied group-entrained communication system

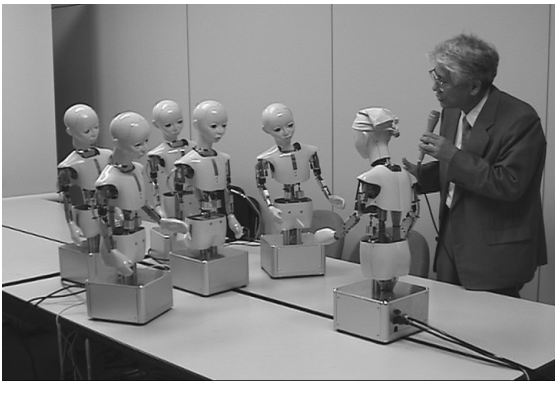

Fig. 6. Physical version of SAKURA with InterRobots

Fig. 7. Speech-driven embodied interaction system with InterRobots and an InterActor in the National Museum of Emerging Science and Innovation

communication. Fig. 7 indicates another physical version of SAKURA with four InterRobots and one InterActor, which is exhibited in the National Museum of Emerging Science and Innovation where visitors can enjoy a dynamic experience of embodied communication. They perceive the effects of group-entrained communication environment intuitively and recognize the importance of embodied communication.

4 Embodied Interaction and Communication Technology

In this section, some actual applications of InterActor/InterRobot to human interface are introduced. Fig. 8 depicts an interaction scene between an InterRobot and children in a kindergarten. This InterRobot is commercialized and marketed for kindergarten. Children enjoy and are excited about having conversations with the InterRobot, while the teacher standing behind the InterRobot enjoys talking and encouraging children in a new communication mode from a completely different standpoint, just changing to a friend and so forth. By focusing on an animal character that is most preferred by children, an animal-type InterActor/InterRobot known as InterAnimal is developed in

Fig. 8. Interaction scene between an InterRobot and children in a kindergarten

Fig. 9. InterAnimal: Animal-type InterActor

order to encourage and cheer up children, as depicted in Fig. 9. The bear-type InterAnimal shown in Fig. 10 was popular in the 2005 exhibit EXPO. Fig. 11 illustrates a toy version of InterRobot with the function of a listener, which generates the listener's actions of nodding, tilting his/her head, and moving his/her arms up and down, based on speech input. The stuffed toy bear is eager to listen without ever uttering a word. It is commercialized and marketed under the name of Unazukikun.

Fig. 10. InterAnimal in EXPO 2005

Fig. 11. Toy version of InterRobot

The InterActor, as indicated in Fig. 12, is also commercialized under the name of InterCaster through which news and media contents are effectively and cordially transmitted in a commercial program. By superimposing InterActors as listeners on the video images of a lecture such as an education program, the InterActor-superimposed learning support system is developed, as illustrated in Fig. 13 [9]. The system provides group-entrained interaction effects for audiences who watch the video, in which two InterActors at the bottom of reduced images are entrained with the lecturer's speech. The InterActor is a speech-driven CG-embodied interaction character that can generate

communicative movements and actions for an entrained interaction. An InterPuppet, on the other hand, is an embodied interaction character that is driven by both speech input, similar to the InterActor, and hand motion input, like a puppet. Therefore, humans can use the InterPuppet to communicate effectively by using deliberate body movements as well as natural communicative movements and actions. An advanced InterPuppet with a cellular phone-type device is developed as indicated in Fig. 14. This interaction character can be used in a mobile environment.

Fig. 12. InterCaster

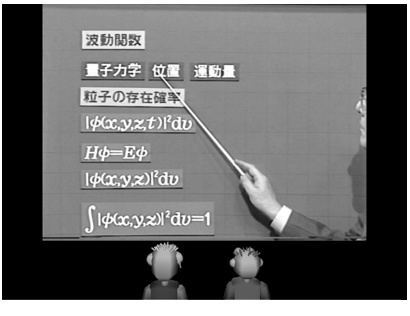

Fig. 13. InterActor-superimposed learning support system

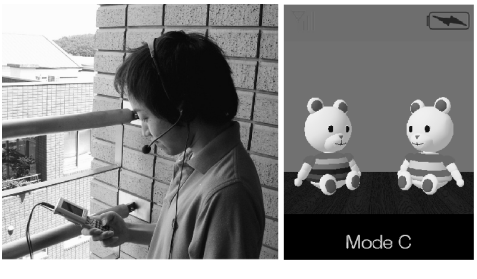

Fig. 14. InterPuppet with a cellular phone-type device

5 Conclusion

The human-entrained embodied interaction and communication technology for an advanced media society was proposed through the development of E-COSMIC for supporting essential human interactive communication based on the entrainment mechanism of the embodied rhythms between speech and body movements such as nodding. Some actual applications pertaining to CG/robot and human interactive communication were also demonstrated. In particular, the speech-driven embodied interaction system, such as InterActor and InterRobot, is a robust and practical communication support system for everyday living, which activates embodied interaction and communication in a new communication mode by using only speech input. The speech-driven entrainment technology for enhancing interaction and communication would be expected to form the foundation of mediated communication technologies as well as the methodology for the analysis and understanding of human interaction and communication, and to develop a new embodied communication industry for supporting essential human interactive communication.

Acknowledgments. This work has been supported by Core Research for Evolutional Science and Technology (CREST) of Japan Science and Technology Agency (JST).

References

[1] Condon, W.S., Sander, L.W.: Neonate Movement is Synchronized with Adult Speech: Interactional Participation and Language Acquisition. Science 183, 99–101 (1974)
[2] Kobayashi, N., Ishii, T., Watanabe, T.: Quantitative Evaluation of Infant Behavior and Mother-Infant Interaction. Early Development and Parenting 1, 23–31 (1992)
[3] Watanabe, T., Okubo, M.: Evaluation of the Entrainment Between a Speaker's Burst-Pause of Speech and Respiration and a Listener's Respiration in Face-to-Face Communication. In: Proc. of the 6th IEEE International Workshop on Robot-Human Interactive Communication (RO-MAN'97), pp. 392–397 (1997)
[4] Watanabe, T.: E-COSMIC: Embodied Communication System for Mind Connection? In: Proc. of RO-MAN 2004, pp. 1–6 (2004)
[5] Watanabe, T., Ogikubo, M., Ishii, Y.: Visualization of Respiration in the Embodied Virtual Communication System and Its Evaluation. International Journal of Human-Computer Interaction 17(1), 89–102 (2004)
[6] Shintoku, T., Watanabe, T.: An Embodied Virtual Communication System for Three Human Interaction Support and Analysis by Synthesis. In: Proc. of the 5th IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA2003), pp. 211–216 (2003)
[7] 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)
[8] Watanabe, T., Okubo, M.: SUKURA: Voice-Driven Embodied Group-Entrained Communication System. In: Proc. of HCI International 2003, vol. 2, pp. 558–562 (2003)
[9] Watanabe, T., Yamamoto, M.: An Embodied Entrainment System with InterActors Superimposed on Images. In: Proc. of HCI International 2005, vol. 4(2045), pp. 1–6(2005)