A Robotic Introducer Agent Based on Adaptive Embodied Entrainment Control

Mutsuo Sano[1], Kenzaburo Miyawaki[1],
Ryohei Sasama[2], Tomoharu Yamaguchi[2], and Keiji Yamada[2]

[1] Osaka Institute of Technology, Fuculty of Information Science and Technology
1-79-1 Kitayama, Hirakata City, Osaka, Japan
{sano,miyawaki}@is.oit.ac.jp
[2] NEC Corporation, C&C Innovation Research Laboratories
8916-47 Takayama-Cho, Ikoma, Nara, Japan
{r-sasama@cp,yamaguchi@az,kg-yamada@cp}.jp.nec.com

Abstract. The necessity for the individual and the individual's tying in real space increases while the age of piece progresses. The robot is requested in that and it is requested what role be able to be played. Here, we pursue the research on the robot design to expand and to promote the group conversation. It proposes the technique to advance the conversation of couple 1 of the first meeting smoothly as the first stage. Especially, it is confirmed that nonverbal interactions are more important than language interactions by a lot of researches so that the individual and the individual may tie. We newly define the communications activity based on embodied entrainment, and propose the method to control the behavior of the robot dynamically according to the state of communications. The active control method uses interaction timing learning which depends on nonverbal communication channels. Our mechanism selects an appropriate embodied robotic behavior by changing the communication strategy based on the state transition of an introduction scene, and increases the communication activity measured by sensing data. The action timing is learned and controlled by a decision-tree. As the result, the real agent robot could control communication situations similarly to a human. We became "Yes" by the evaluation value of 82% for the question that communications had risen as a result of doing the questionnaire survey to 20 university students. Moreover, familiarity became a first meeting introduction robot with "Yes" for the question about whether being possible to have it by the evaluation value of 85%. Therefore, the effectiveness of this proposal technique was verified. Finally, the possibility that the circle of communications can be expanded to N person's group is discussed based on this result.

Keywords: Embodied Entrainment Nonverbal Communication Robotic Introducer Agent, Group communication, Human's Action Learning.

1 Introduction

Recently, the necessity for an individual mind and the mind's tying in real space increase in the human society while the age of piece progresses. What is the role

J.A. Jacko (Ed.): Human-Computer Interaction, Part II, HCII 2009, LNCS 5611, pp. 368–376, 2009.

that the robot that has the sociality can be accomplished in that? We pursue the robot design to expand the circle of communications at the same time as smoothly doing the group conversation by a home space and a public space, etc. Humans communicate with each other via many channels, which can be categorized into verbal and nonverbal. It has been confirmed that the nonverbal communication channels, such as an intonation, an accent, a gaze, a gesture, nodding or an emotional expression, encourage synchronization, embodied entrainment and a friendship [1]–[7]. Watanabe et al. [8] has investigated such a robotic agent, which includes embodied entrainment. The research has developed the speech-driven interactive actor, the Inter Actor, which could predict an appropriate timing of a nod and a gesture with voice information of a communication partner. Kanda et al. [9] [10] adopts a joint gaze function to a robot and has investigated more adaptive robot through design-based approach. However, we cannot find so many researches about an adaptive embodied entrainment, which aims to increase group communication activity by real time feedback on a situation of the group communication. This paper presents adaptive control of the embodied entrainment using learning of interaction timing in group communication. The control mechanism is based on communication activity measurement and uses nonverbal communication channels effectively. First of all, we start from the research that the robot supports the case with a first meeting couple 1 mutually shown in Fig.1, and install the system into a robotic introducer agent and confirm its effectiveness. In this paper, we show our robotic agent design. Next, we explain our method for adaptive control of embodied entrainment and show an experimental result of our method. Finally, the effectiveness is verified according to result of the questionnaire, and the N person's development with the group communications is discussed.

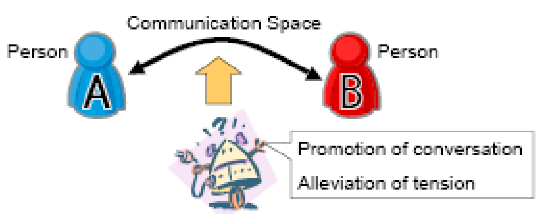

Fig. 1. Introducer robot of a first meeting couple

2 Basic Strategy to Promote Group Conversation

Fig. 2 shows our robot agent. The purpose of this introducer agent is support of conversations of 2 participants who have not known each other. For the purpose, we designed various basic strategies of agent behaviors. The details of the behaviors are as follows. (Fig.3)

1. **Leading utterances**
   The agent asks a question to participants for establishment of a binary reliable relation between herself and each participant.
2. **Gaze lead**
   For transmission of information of a participant A to the other participant B, the agent moves her gaze to A and asks him/her about the information. After that, the agent leads A's gaze to B to accord A's gaze and B's gaze.
3. **Gaze distribution**
   When all participants talk each other, the agent moves her gaze to them equally.
4. **Synchronizing nod**
   The agent adjusts her nod timing to participants' nod.
5. **Dynamically Synchronizing**
   Synchronizing method is dynamically changed according to the state of two persons as follows.
   a) Natural synchronizing to a listener
   When the listener listens and tunes to the talker by the nod and the back-channel feedback operation, the robot tunes it according to listener's timing to pile up the communication space further.
   b) Synchronizing to a speaker to invite the other participant to the conversation
   When the listener is not tuning it to talker's timing, the robot tunes to talker's rhythm while turning one's gaze to the listener, invites the listener's tune, and creates the rhythm of the communications space.

Fig. 2. Robotic Introducer Agent

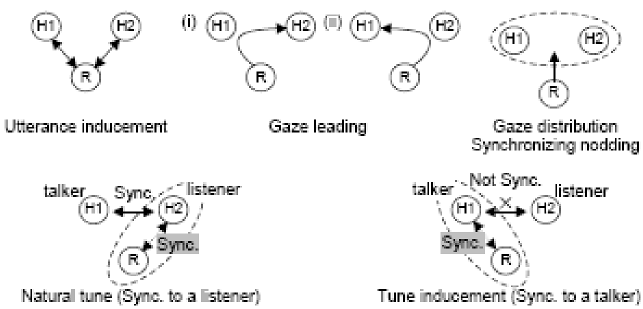

Fig. 3. Basic strategies to promote group conversation

3 Adaptive Control of Embodied Entrainment

3.1 Macro Strategy to Promote Group Conversation

We think about the state transition model to promote the group conversation by using the basic strategies. It establishes a rapport by grounding process, and communications of the group are activated by enhancement process in Fig.4. We segmented an introduction scene into 5 states based on preliminary observations. The agent moves among these states according to the situation of the participants' communication. Fig.4 shows the 5 state transitions. The details of the states and the agent behaviors are as follows.

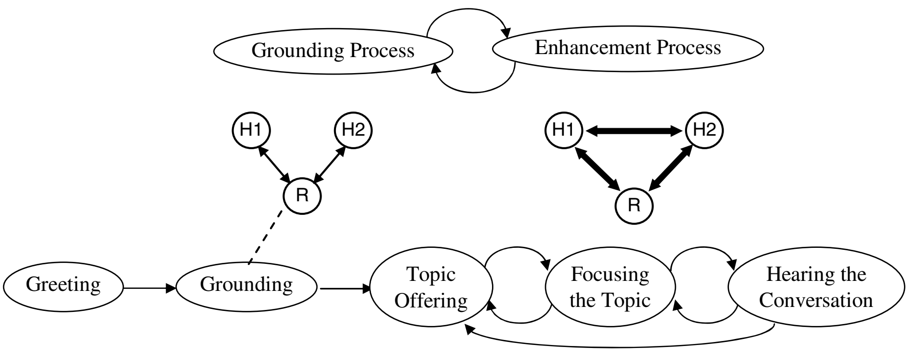

Fig. 4. Macro strategy and state transition of an introduction scene

3.2 Communication Activity Measurement and the State Transition Condition

The robot agent monitors the situation of participants' communication, and estimates the current state on Fig.4. For the estimation, we define the communication activity measurement. The agent calculates the activity in every time slice. It is defined by the average of speech power, gaze direction, and gaze corresponding degree of both. When it satisfies a condition of a state transition, the agent moves to the next different state, which is directed by the transition.

Next we show the behaviors at each state.

1. **Greeting**
   The agent introduces himself and offers a brief explanation of the situation. Then, she simply introduces the participants' name.
2. **Grounding**
   In this state, the agent tries to make a reliable relationship between the robot and each participant. For this, the agent cites the participants' profile and asks simple questions to them. It aims to entrain them in to the rhythm of the robot by using the strategy of utterance leading. Here, the robot assumes that it knows the data of the hobby etc. of the person who is introduced.
3. **Topic Offering**
   This state encourages conversations between the participants. The agent offers the information and profiles about a participant to the other participant, or asks a

simple question to the participants in order to make a favorable situation for conversations of the participants. By such behaviors, the agent controls the participants' gaze to make them communicate face to face each other.

4. **Focusing on a Specified Topic**
   In this state, the agent tries to join in a conversation of the participants. The agent focuses on a topic, which was offered at the previous state, Topic Offering.
5. **Hearing Conversations**
   After a success in making a close and friendly relationship between the participants, the agent keeps hearing their conversations quietly. In this state, the agent nods and looks at a speaker using basic embodied entrainment strategies, with a proper timing.

3.3 Learning for Adaptive Control of Embodied Entrainment

Our agent performs some actions to encourage embodied entrainment. The agent retrieves the information of the participants by sensors, and determines its action. We built a decision-tree, which is shown in Fig.5, by inductive learning for this interaction. The following 3 rules are extracted from the tree.

1. **Rule #1 Reactive nodding**
   This rule represents the reactive nodding to a participant's nod. The agent nods when someone of the participants nods.
2. **Rule #2 An utterance**
   This rule explains the following situation. When a participant keeps quiet and the other participant terminates his/her speaking, it can be supposed that the conversation got into a break. Therefore, the agent recognizes this situation and she offers a topic to avoid silence.
3. **Rule #3 An utterance and a nod to a participant's speech**
   This rule is very similar to the Rule #2 but represents the situation that a participant stopped his/her utterance without terminating his/her speech. In this situation, the agent speaks in 70% probability and nods in 30%. The criterion to determine

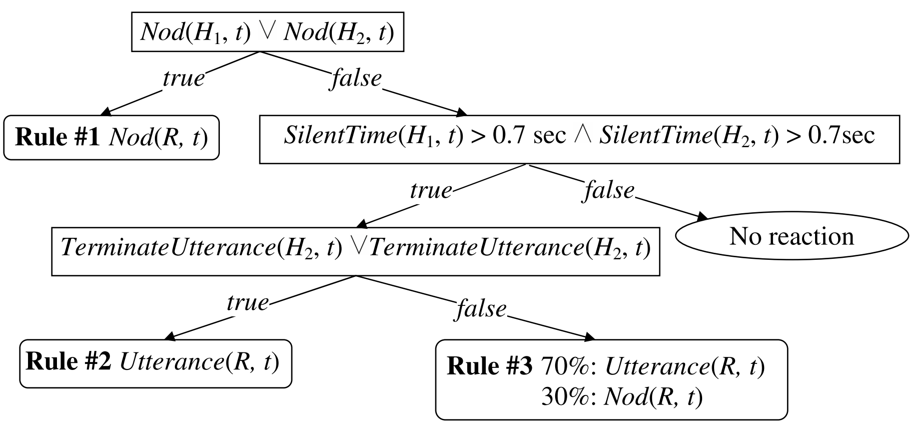

Fig. 5. Decision tree for embodied entrainment

whether a participant terminated his/her utterance or not, depends on the final part of the utterance. We extract the final part by Japanese morphological analysis.

4 Experiments

4.1 Experimental System

Information of current situation is recognized by the situation cognitive system which is composed of voice recognition, voice power detection, gaze detection and nodding detection. After that the interaction generator retrieves the information, and determines the agent behavior with the participants' profiles.

4.2 Sensing Environment

Detection modules in the cognitive system retrieve raw sensor data and extract the participants' state and action. The details of each module and sensors are listed below.

1. **Gaze Detection**
   Multiple cameras are used for this module. We set 2 cameras for each participant. Each camera detects the participant's face, and the gaze detector estimates that the participant is turning his/her gaze to the camera, which can detect the participant's face.
2. **Nodding Detection**
   We set acceleration sensors to the participants' head, and measure the movement of heads. Based on the movement, this module detects his/her nodding.
3. **Voice Recognition, Voice Power Detection**
   A pin microphone is set on each participant.

4.3 Experimental Result

Fig.6 shows an experiment with our introducer agent. We asked 2 participants to talk each other during 6 minutes. The participants had never known each other. In this situation, we checked whether our method could control the agent behavior appropriately.

Fig.7 plots the *Eval*(*t*) and the state transition of a communication situation. *Eval(t)* is the communication activity degree at time slice, and is defined as follows.

$$E(t) = \sqrt{\left(\alpha * \frac{PA(t)}{PA(ave)}\right)^2 + \left(\alpha * \frac{PB(t)}{PB(ave)}\right)^2 + \left(\frac{1}{\beta} Gw\right)^2} \quad (1)$$

$$Eval(t) = \frac{E(t) + Eval(t-1)}{2}$$

PA(*t*) or *PB*(*t*) denotes the speech power of person A or B at time t, and P*A(ave)* or *PB*(*ave*) is the average of *PA*(*t*) or *PB*(*t*). *Gw* shows the probability (%) of gaze

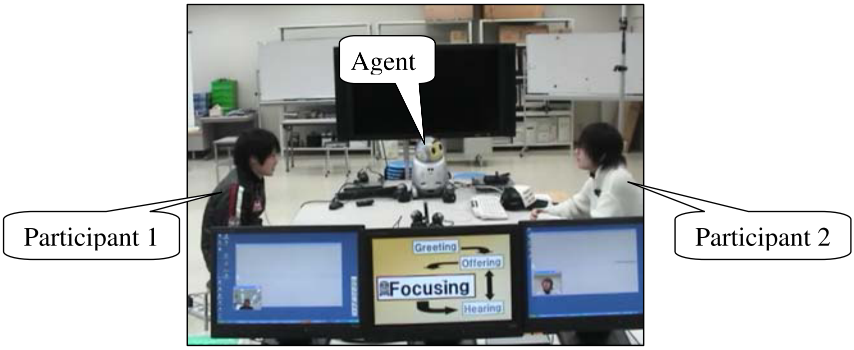

Fig. 6. Experiments with the agent

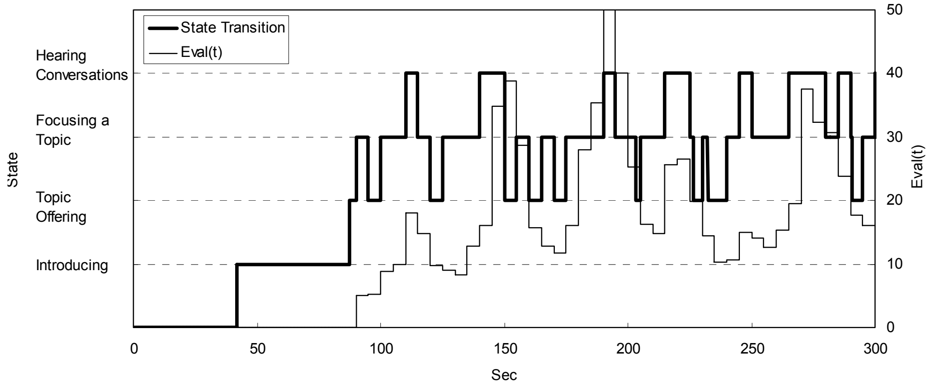

Fig. 7. *Eval*(*t*) and state transition

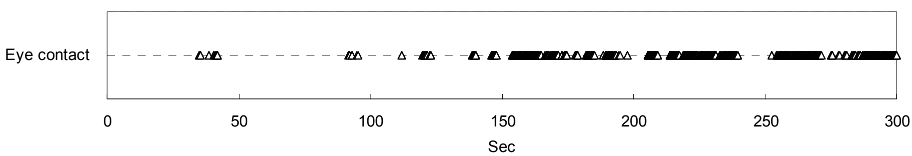

Fig. 8. Participants' eye contact

corresponding. α and β are a weighting factors. Then we define *Eval*(*t*) as the average of *E(t)* on the time axis.

We can see an appropriate movement of the state and the *Eval*(*t*) on the graph. Fig.8 shows a frequency of participants' eye contact. We could also observe the situation that the participants had intensified friendship through many natural conversations, which had been led by our agent.

We did the questionnaire survey to 10:20 university students. As a result, a negative evaluation was not at all, and became 82% for the question that communications had risen by the evaluation value in which four or more was converted by 100% by the evaluation value of five stage average. Against the first meeting introduction robot and familiarity was confronted to the question about whether being possible to have it, and moreover, a negative evaluation was not at all, and became 85% by the evaluation value in which four or more was converted by 100% by the evaluation value of the stage average. Thus, the effectiveness of this proposal technique was verified by the questionnaire survey.

5 Future Works

The mechanism of the group communications activation of couple 1 proposes with this paper can be enhanced also to N person's group communications basically. However, an important point is to have to select the strategy for activation appropriately in enhancing responding to group member's personality. It doesn't mediate in consideration of the character though a hobby each other and the concern, etc. are understood beforehand and communications were supported in the actual experiment. For instance, it might be effective to have the person with a high communications skill to the group of the facile personality the coupling. Moreover, the scene that mediates communications is not only a first meeting. We assume application to the home for the aged and the waiting room, etc. shown in Figure 9, and are researching the framework to activate the group communications. The number also of cases to shut one's mind voluntarily increases for the elderly person, and the achievement of the heart proxy that the robot ties to the mind the mind in real space is requested.

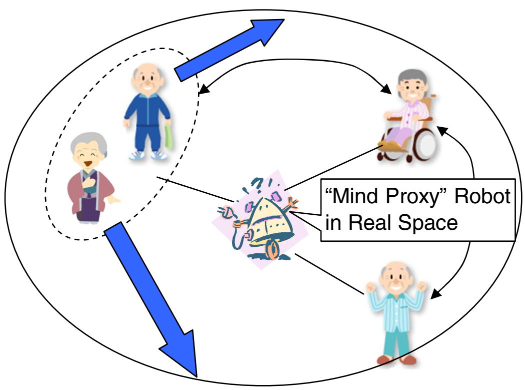

Fig. 9. Enhancing of activation of N person's group communications to research

6 Conclusion

We proposed the mechanism of the group communications activation of couple 1. We showed the control mechanism of adaptive embodied entrainment and implemented the mechanism and installed it into the real robotic introducer. The effectiveness of our proposal was confirmed by the communication experiments and the questionnaire survey. This mechanism was based on the communication activity measurement, and the learning of timing for the nonverbal actions such as gaze, nodding or back-channeling. In the aspect of embodied action control, the system selects the most appropriate action, which is suitable to each communication state, such as greeting, grounding, offering topics, focusing the topic and hearing a conversation. Add to the action control, the proper changes of a speaker and a listener increased the communication activity degree. We focused on the gaze lead, gaze distribution and synchronizing nod for the embodied action. The timing of robotic actions and utterances was based on the rules, which was extracted from a decision-tree built through induction learning of human behaviors. We installed our control system into the introducer robot and tested it on the communication experiments. As the result, the agent could control the communication situation similarly to the human.

We start to enhance this method to N person's group communications schema. And we will apply it to the application of the home for the aged and the waiting room.

References

[1] Matarazzo, J.D., Weitman, M., Saslow, G., Wiens, A.N.: Interviewer influence on durations of interviewee speech. Journal of Verbal Learning and Verbal Behavior 1, 451–458 (1963)

[2] Webb, J.T.: Interview synchrony: An investigation of two speech rate measures in an automated standardized interview. In: Pope, B., Siegman, A.W. (eds.) Studies in Dyadic Communication, pp. 115–133. Pergamon, New York (1972)

[3] Kendon, A.: Movement coordination in social interaction: Some examples described. Acta Psychologica 32, 101–125 (1970)

[4] Condon, W.S., Sander, L.W.: Neonate movement is synchronized with adult speech: interactional participation and language acquisition. Science 183, 99–101 (1974)

[5] Bernieri, F.J., Reznick, J.S., Rosenthal, R.: Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in motherinfant interactions. Journal of Personality and Social Psychology 54, 243–253 (1988)

[6] Takeuchi, Y., Katagiri, Y.: Emotional feedback by synthesized voice and facial expressions. In: Proccedings of the 3rd international conference of cognitive science, pp. 697–698 (2001)

[7] Kojima, H., Nakagawa, C., Yano, H.: Can a robot empathize with people? International Journal of Artificial Life and Robotics 8, 83–88 (2004)

[8] Watanabe, T.: E-COSMIC: Embodied Communication System for Mind Connection. In: Proc. of the 13th IEEE International Workshop on Robot-Human Interactive Communication (RO-MAN 2004), pp. 1–6 (2004)

[9] Kanda, T., Ishiguro, H., Ono, T., Imai, M., Maeda, T., Nakatsu, R.: Development of 'Robovie' as platform of everyday-robot research, The transactions of the Institute of Electronics. Information and Communication Engineers. D-I J85-D-I(4), 380–389 (2002)

[10] Kanda, T., Ishiguro, H., Imai, M., Ono, T.: Development and Evaluation of Interactive Humanoid Robots. Proceedings of the IEEE (Special issue on Human Interactive Robot for Psychological Enrichment) 92(11), 1839–1850 (2004)