Available online at www.sciencedirect.com
SCIENCE DIRECT®
Int. J. Human-Computer Studies 62 (2005) 247–265
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

Cooperative embodied communication emerged by interactive humanoid robots

Daisuke Sakamoto[a,b,*], Takayuki Kanda[b], Tetsuo Ono[a,b], Masayuki Kamashima[b,c], Michita Imai[b,c], Hiroshi Ishiguro[b,d]

[a] *Department Media Architecture, Future University-HAKODATE, 116-2 Kamedanakano, Hakodate, Hokkaido, 041-8655, Japan*

[b] *ATR Intelligent Robotics and Communication Laboratories, 2-2-2 Hikaridai, Seikacho, Sorakugun, Kyoto, Japan*

[c] *Keio University Graduate School of Science and Technoloty, 3-14-1 Hiyoshi Yokohama, Kanagawa, Japan*

[d] *Osaka University Graduate school of Engineering, 2-1 Suita, Osaka, Japan*

Received 9 February 2004; received in revised form 22 May 2004
Available online 16 December 2004

Abstract

Research on humanoid robots has produced various uses for their body properties in communication. In particular, mutual relationships of body movements between a robot and a human are considered to be important for smooth and natural communication, as they are in human–human communication. We have developed a semi-autonomous humanoid robot system that is capable of cooperative body movements with humans using environment-based sensors and switching communicative units. Concretely, this system realizes natural communication by using typical behaviors such as: "nodding," "eye-contact," "face-to-face," etc. It is important to note that the robot parts are NOT operated directly; only the communicative units in the robot system are switched. We conducted an experiment using the mentioned robot system and verified the importance of cooperative behaviors in a route-guidance situation where a human gives directions to the robot. The task requires a human participant (called the "speaker") to teach a route to a "hearer" that is (1) a human, (2) a developed robot that performs cooperative movements, and (3) a robot that does not move at

*Corresponding author. Department Media Architecture, Future University-HAKODATE, 116-2 Kamedanakano, Hakodate, Hokkaido, 041-8655, Japan. Tel.: +81 138 34 6132; fax: +81 138 34 6301.

E-mail addresses: g2104017@fun.ac.jp (D. Sakamoto), kanda@atr.jp (T. Kanda), tono@fun.ac.jp (T. Ono), kamasima@ayu.ics.keio.ac.jp (M. Kamashima), imai@ayu.ics.keio.ac.jp (M. Imai), ishiguro@ams.eng.osaka-u.ac.jp (H. Ishiguro).

1071-5819/$ - see front matter
doi:10.1016/j.ijhcs.2004.11.001

all. This experiment is subjectively evaluated through a questionnaire and an analysis of body movements using three-dimensional data from a motion capture system. The results indicate that the cooperative body movements greatly enhance the emotional impressions of human speakers in a route-guidance situation. We believe these results will allow us to develop interactive humanoid robots that sociably communicate with humans.

Keywords: Human–robot interaction; Entrainment; Subjective experiments; Environment-based sensing

1. Introduction

We believe that subtle expressivity is an effective technology for embodied agents such as robots and character interfaces. The legacy of CUI and poor GUI-based interaction is that only simple information could be transmitted in the 'as is' condition. On the contrary, embodied agents with a head, arms and body can communicate with humans non-verbally by using gestures as well as verbal information such as voice and text. In this paper, we report subtle expressivity for emotional communication that has been made possible by a robot's human-like body movements.

Over the past several years, many humanoid robots such as Honda's (Hirai et al., 1998) have been developed. We believe that in the near future, humanoid robots will interact with humans in our daily lives. These robots' human-like bodies enable humans to intuitively understand their gestures and cause people to unconsciously behave as if they were communicating with humans (Kanda et al., 2002). In other words, if a humanoid robot uses its body effectively, people will be able to communicate naturally with it. This allows robots to perform such communicative roles in human society as route guides.

Previous research works proposed various types of communicative behaviors made possible by human-like robots. For instance, the eye is a very important communicative body part; eye gaze and eye contact are therefore often implemented in robots. Nakadai et al. (2001) developed a robot that tracks a speaking person, and Matsusaka et al. (1999) developed such a robot that uses eye contact. These works demonstrated that the eyes play an important role in conveying communicative intention to humans.

Furthermore, the eyes allow us to share attention with other people. This phenomenon is widely known as the joint-attention mechanism in developmental psychology (Moore and Dunham, 1995). Scassellati (2000) developed a robot, called Cog, as a testbed for a joint-attention mechanism. In this work, the robot follows the others' gaze in order to share attention. Other research groups have also developed robots that have a joint-attention mechanisms (Kozima and Vatikiotis-Bateson, 2001; Breazeal and Scassellati, 2000). Imai et al. (2003) used a robot's arms as well as its eyes to establish joint attention and verified its effectiveness.

These research works show that mutual relationships of body movements between a robot and a human are important for smooth and natural human-like communication. Ono et al. (2001) verified the importance of eye contact, arm gestures, and appropriate positional relationships (orientation of body direction) in a route-guide robot. In this research, it was found that body movements are not only used for visually understanding what the speaker says but also for synchronizing communication. The speaker's body movements entrain hearers to establish a relationship with him or her. Such an unconscious synchronization of body movements is called "entrainment."

In robotics, two directions have been taken to exploit cooperative body movements based on entrainment: one is to use it for a human interface. Ogawa and Watanabe (2001) developed a robot that induces entrainment by using cooperative body movements such as nodding, which supports human–human telecommunication. The other direction is a robotic partner that autonomously interacts with humans. We are trying to develop such an autonomous robot that uses its human-like body effectively in communication with humans. We found that such cooperative body movements as eye-contact and synchronized arm movements are mutually related to one's subjective evaluation of the robot (Kanda et al., 2003). Furthermore, we believe that cooperative body movement is essential for humanoid robots that entrain humans into communication with it.

Here, we describe a point of difference between the research of embodied conversational agents and our research. Cassell et al. (2000) have investigated the planning needed to realize a personified character agent that can behave in human-like cooperative body movement and utterances in "virtual space". Similarly, Nakano et al. (2003) considered embodied movements such as eye contact and nodding as a "signal" that help to construct relationship between humans and agents. However, we expect that there is the difference of international property between a robot in the real-world and a software agent in the virtual one. Although this problem is a matter of controversy, in this paper, we try to investigate the body movement in three-dimensional real-world, such as body orientation and pointing, spatial and temporal synchronization that will not able to realize by two-dimensional one.

In this paper, we investigate the effect of cooperative body movements in human-robot communication. We have developed a humanoid robot that is capable of performing various cooperative body movements such as eye contact and synchronized body movements. A software module we have named a "communicative unit" produces each cooperative body movement, and one communicative unit controls each body part (head, left arm, right arm, body orientation, and utterance). These communicative units are currently selected by the Wizard of Oz (WOZ) method; however, we have recorded the operation logs for future use in implementing autonomy.

We performed an experiment to verify the effect of the implemented embodied behaviors. In the experiment, human participants taught the robot a route. Our hypothesis was that humans would feel comfortable and thus easily teach the robot the route due to entrainment through cooperative body movements. Since the robot's task is to listen to the guidance given by a human, the robot only performed reactive body movements toward the human speaker instead of performing symbolic gestures such as sign language.

2. Robot for cooperative embodied communication

We developed a humanoid robot that is capable of cooperative embodied behaviors by using a motion-capturing system and the WOZ method. In this section, we introduce the robot's system configuration.

2.1. Interactive humanoid robot ''robovie

Fig. 1 (left) shows our interactive humanoid robot ''Robovie,'' which is characterized by its human-like body expression and various sensors (Ishiguro et al., 2003). The human-like body consists of eyes, a head, and arms, which generate the complex body movements required for communication. The various sensors, such as auditory, tactile, ultrasonic, and visual, enable Robovie to behave autonomously and to interact with humans. Furthermore, the robot satisfies the mechanical requirements of autonomy; plus it includes all computational resources needed for processing the sensory data and for generating behaviors.

2.2. Environment-based sensing

We employed an optical motion-capturing system as an environment-based sensor that allows the humanoid robot to perform cooperative body movements such as eye-contact and synchronized arm movements. Fig. 2 illustrates the software configuration of the robot system with the environment-based sensor. There are three software components executed in parallel: position calculator, communicative units, and robot controller.

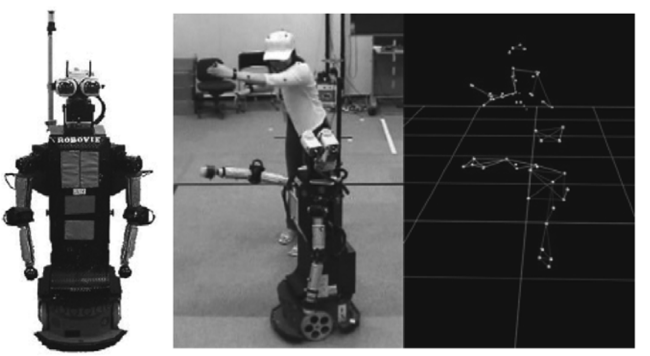

Fig. 1. Robovie(left) and motion capturing system(right).

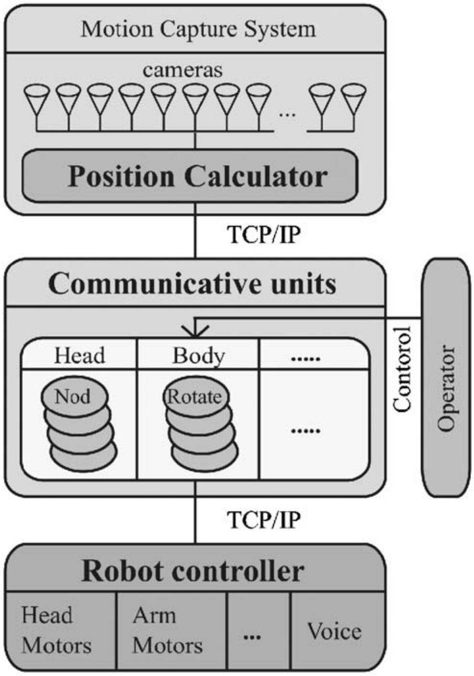

Fig. 2. Software configuration.

2.2.1. *Position calculator*

The motion-capturing system was used to measure the human's and robot's body movements. This system comprises 12 pairs of infrared cameras and infrared lights and markers that reflect infrared signals. These cameras were set around the room. The system calculates each marker's three-dimensional position from all camera images, and it features high resolution in both time (120 Hz) and space (accuracy of 1 mm in the room).

As shown in Fig. 1 (right), we attached 23 markers to both a human and the robot on the following places: head (the human wore a cap with attached markers), shoulders, neck, elbows, and wrists. By using the attached markers at corresponding parts on the robot and humans, the robot can perform cooperative body movements. (Some of the markers were used only for tracking humans or robot movement with kinematics constraints, which was performed by the motion-capturing system.)

The position calculator utilized the motion-capturing system to obtain three-dimensional position data on each marker and to transmit the data to the communicative units. The delay in calculation was about 30–50 ms.

2.2.2. *Communicative units*

We prepared communicative units for controlling head, utterance, right arm, left arm, and body orientation. Each communicative unit controls the corresponding body part based on the body movements of the human who is communicating with

Table 1
Implemented communicative units

Head	Arm (left/right)	Body (locomotion)	Utterance
Eye contact	Synchronized arm movement (left/right)	Face-to-face	Please teach me the way.
Gaze in pointed direction	Mirrored synchronized arm movement (left/right)	Standing side-by-side	Hum, uh-huh, well,
Nod	Arm down	No movement	Excuse me, once more,
Tilt head doubtfully	No movement		More slowly, please,
			Thank you, I understand.

the robot. In other words, these communicative units realize cooperative body movements. Table 1 lists all of the implemented communicative units. We also explain typical ones below:

2.2.2.1. Eye contact. This unit controls the robot's head toward the human's head so that the robot maintains eye contact with the human. The position of the human face is ascertained by determining the positions of the three markers attached to the human head.

2.2.2.2. Nod. This unit controls the robot's head to perform a nodding gesture. It does not require any sensory information related to the markers.

2.2.2.3. Synchronized arm movement (right/left). This unit controls each of the robot's arms to imitate the human's right or left arm motions. It realizes synchronized arm movements as a human hearer does (Ono et al., 2001) when he or she points to an object or direction to guide someone along a route. Using the marker positions of shoulder, elbow, and wrist, it calculates the position of a human hand relative to the shoulder, after which it calculates the joint angles of the robot arm from the calculated relative position of the human hand.

2.2.2.4. Mirrored synchronized arm movement (right/left). This unit is similar to the "synchronized arm movement" except that the imitation is mirrored (for example, if the human points to the right, the robot will point to the left). It is to be used when a human and the robot are face-to-face, whereas the normal type is used when a human and the robot stand side-by-side.

2.2.2.5. Standing side-by-side. A previous study (Ono et al., 2001) also identified the importance of body orientation in route guidance. In that study, standing side-by-side was a more suitable body orientation than face-to-face. This communicative unit realizes the body orientation of standing together in a row. It calculates the human's relative position and orientation from the robot and then moves the robot to a preferable place in the environment. The positions and orientations are retrieved from the shoulder markers.

We have also implemented communicative units for utterances such as "hum," "well," "uh-huh," and "I understand." These are also used as responses to the route guidance given by humans. Each of them is so simple that the robot speaks a phrase using only one of these utterances.

2.2.3. Robot controller

This unit simply receives commands from the communicative units and controls the robot's body. In addition, it transmits current joint angles and positions to the communicative units.

2.3. System configuration with WOZ settings

Fig. 3 is a scene illustration giving an example of when the software is used with the WOZ settings. Currently, this robot system is semi-autonomous: the robot autonomously moves its body in reaction to human movements with certain communicative units, and human operators need only control the switching of the communicative units. Details of the switching rule are described in Section 3.1.5. Meanwhile, their operations are recorded for future use.

We believe that it will be possible to implement an autonomous switching mechanism instead of the WOZ settings in the future; however, we have no knowledge of how to switch them, so we need to collect data to obtain the implicit rules of switching.

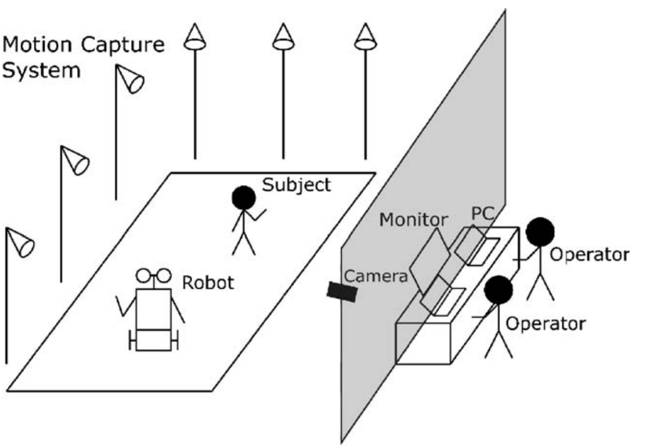

Fig. 3. Selection of communication unit with WOZ method.

3. Experiment

An experiment was conducted to verify the effects of the implemented cooperative body movements.

3.1. Method

A human participant (called the "speaker") teaches a route to a hearer (the developed robot or another human). The robot performed the cooperative body movements with the WOZ method while being guided along the route. We investigated how the body movement of the hearer (especially the robot) affects the speaker.

3.1.1. Conditions

Three conditions were prepared for the hearer: human participant (named "H condition"), a robot with body movements (Rr condition), and a robot without body movements (Rn condition).

3.1.2. Participants

Fifty university students participated in the experiment (23 males, 27 females). They did not know the route that they would teach prior to the experiment. They participated in experiments under the H condition and one of the R conditions (either Rr or Rn). In the H condition, they joined as a speaker or a hearer. The experiment was performed with a counterbalanced design (that is, the experiment was conducted in the order of either H–R or R–H). The assignment of participants was completely random.

3.1.3. Environment setting

Fig. 4 shows the experimental environment. A speaker taught one of two routes, either B (S1…S4) or C (T1…T6), to a hearer in the room (shown as A).

3.1.4. Procedure

The procedure of the experiment was as follows:

1. A speaker walks the route that he/she is going to teach once.
2. The speaker freely teaches the route to the hearer in the room.
3. When the hearer says "I understand," the experiment ends.
4. The speaker answers a questionnaire.

3.1.5. Robot operation (WOZ policy)

3.1.5.1. Switching rule of communicative units. Here we explain the switching rules for an operator when he/she switches communicative units.

First, once an experiment has started, the operator chooses the "Eye contact" unit.

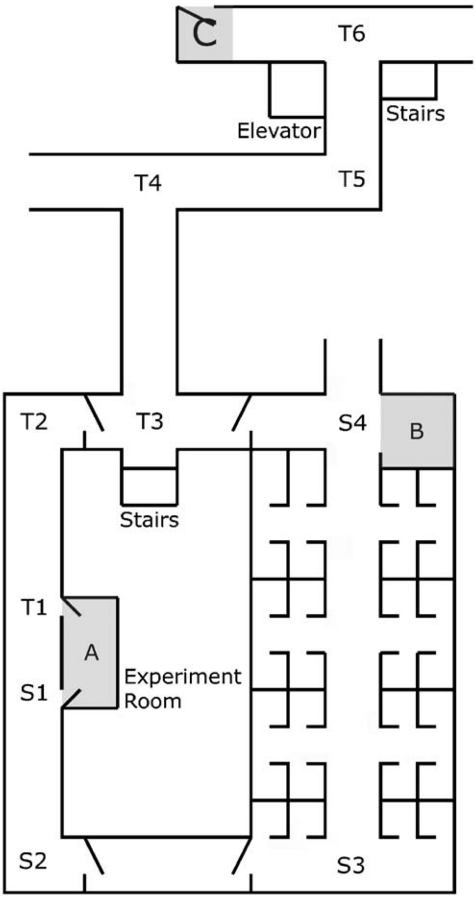

Fig. 4. Environment for experiments.

Second, the operator chooses the "Standing side-by-side" unit to adjust the robot's body to suit the speaker's. When the speaker points to an object or direction to guide the hearer along the route by arm movement, the operator chooses either the "Synchronized arm movement (right/left)" unit or "Mirrored synchronized arm movement (right/left)" unit. At this time, the operator chooses the robot's arm which distant from the speaker. In other words, the operator does not control robot's arm movement directly but only chooses a unit. With the speaker's route guidance, the operator chooses the "Nod" unit and the "Gaze at pointed direction" unit.

Following the completion of an experiment, the operator chooses the "Arm down" and "Face-to-Face" units. At this time, however, the "Eye contact" unit remains selected.

Fig. 5 shows two samples of cooperative body movement. These pictures show the robot's synchronized arm movement and standing points. Figs. 6–8 show the flow charts that are switching rule of each robot parts.

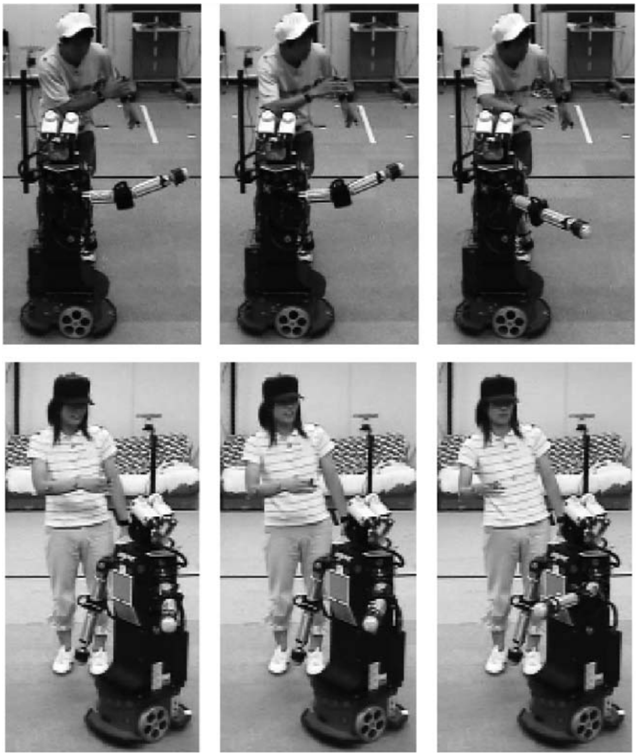

Fig. 5. Examples of cooperative body movements.

3.1.5.2. Utterance rules of communicative units. Here we explain the switching rules for an operator when he/she selects robot utterance units.

First, once an experiment has started, the operator chooses the ''Please teach me the way'' unit.

After that, the operator chooses the ''hum'' or ''hum-hum'' units with speaker's route guidance. If the speaker says ''Do you understand?'', the operator chooses the ''yes'' unit. However, if the route guidance is too difficult or too fast to understand, the operator chooses the ''Once more,'' ''Please more slowly'' units. In this situation, the difficulty is judged by the operator (the experimenter).

Once an experiment is finished, the operator chooses the ''I understand'' and ''thank you'' units.

Fig. 9 shows the flow chart that is the switching rule of robot's utterance.

3.1.6. Evaluation

We asked each speaker the six items shown in Table 2. They rated these items on a 1–7 scale, where 7 is the most positive.

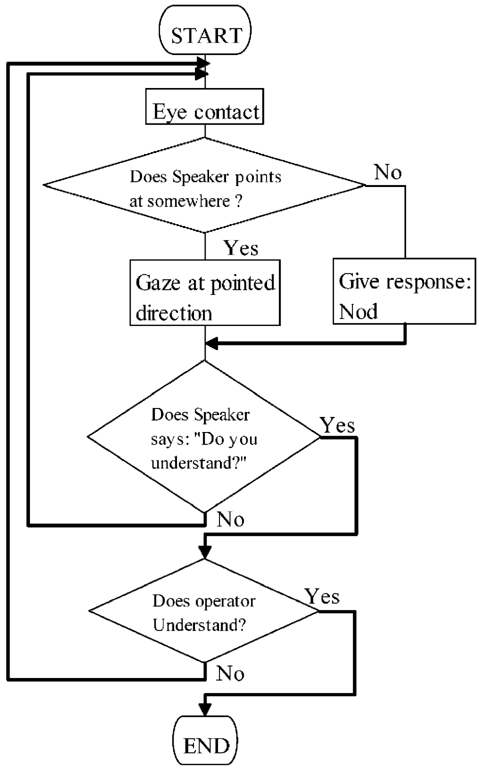

Fig. 6. Switching rule of head movements.

3.2. Hypothesis and expectation

Our hypothesis for the experiment was as follows:

The cooperative body movements of a hearer (made by the developed robot) allow a speaker to smoothly and naturally teach a route as if they were teaching it to other humans.

In addition to the above hypothesis, we predicted that the subjective evaluation for the Rr condition would be better than that for the Rn condition, but similar to the H condition.

4. Results

We show the results of the experiment by analysing both the questionnaire responses and the recorded body movements.

4.1. Analysis of questionaires

Table 3 shows the means, standard deviations, and results of an ANOVA (analysis of variance) of the questionnaire. Fig. 10 also illustrates the results of the

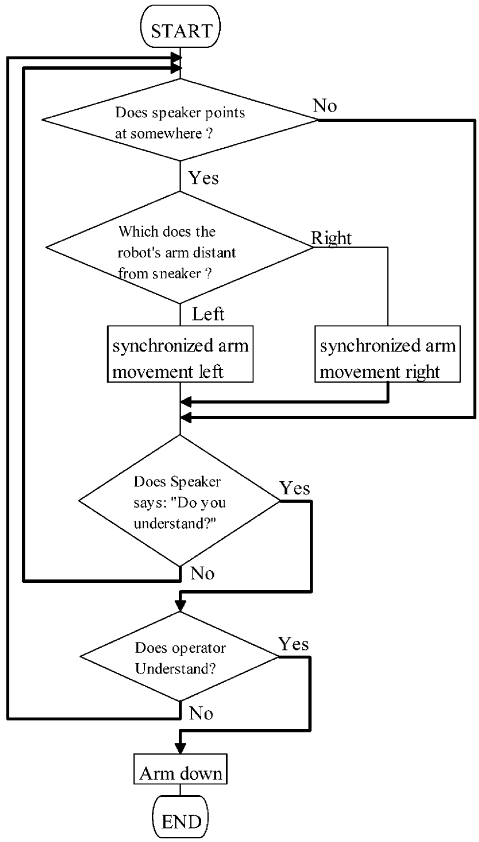

Fig. 7. Switching rule of arm movements.

questionnaire. Since several experiments failed due to hardware failure, we omitted nine sets of data (as a result, the numbers of valid participants were H: 25, Rr: 21, Rn: 20). The ANOVA shows that there are significant differences in Q.3, 5, and 6, and nearly a significant difference in Q.1, 4. An LSD method proved that the results of H and Rr are significantly better than those of Rn (Q.3 ($MSe = 1.7708, p < .05$), Q.5 ($Mse = 1.9231, p < .05$), and Q.6 ($Mse = 1.8316, p < .05$)). It also suggests that the result of H is better than those of Rn (Q.1 ($Mse = 3.0192, p < .05$), Q.4 ($Mse = 1.7303, p < .05$)).

These results also highlight the positive effects of the developed robot system. The participants tended to assume that the robot, displaying cooperative body movements, seemed to hear their guidance, and they also tended to feel empathy with it. Thus, the cooperative body movements contributed to emotional aspects during route instruction. As a result, our hypothesis was verified.

Furthermore, we normalized the questionnaire averages to compare each condition. In this situation, H's condition was 1, Rn's condition was 0, and Rr's

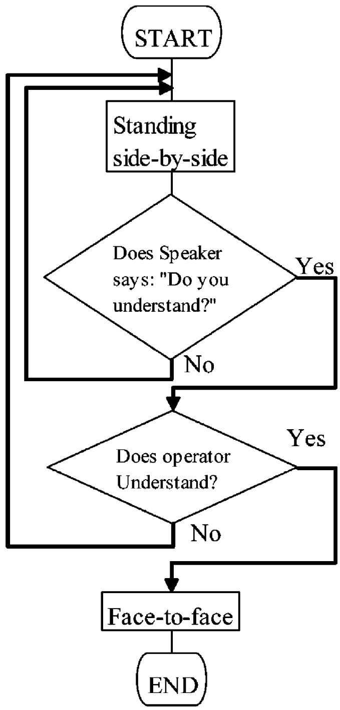

Fig. 8. Switching rule of body movements.

condition was 0.78. This result shows that the participants would treat the robot as if they had communicated with humans; they felt comfortable and thus could easily teach the robot the route. Therefore, we are sure that this robot was appropriately operated and correctly switched between communicative units.

4.2. Analysis of body movements

The results of the analysis of humans' body movements are given below. We analysed the three-dimensional data from the motion capture system and behaviors from the video taken during the experiment.

First, we compared each condition's results as the calculated amounts of body movement using the three-dimensional data from the motion capture system. We then calculated body movements as AVE, which was average finger movement per second. The AVE measure was calculated by this formula:

$$AVE = \left(\left(\sum_{t=1}^{n} \sqrt{(P_{\text{right}}(t) - P_{\text{right}}(t+1))^2} \right) \bigg/ n + \right. \\ \left. + \left(\sum_{t=1}^{n} \sqrt{(P_{\text{left}}(t) - P_{\text{left}}(t+1))^2} \right) \bigg/ n \right) \bigg/ 2.$$

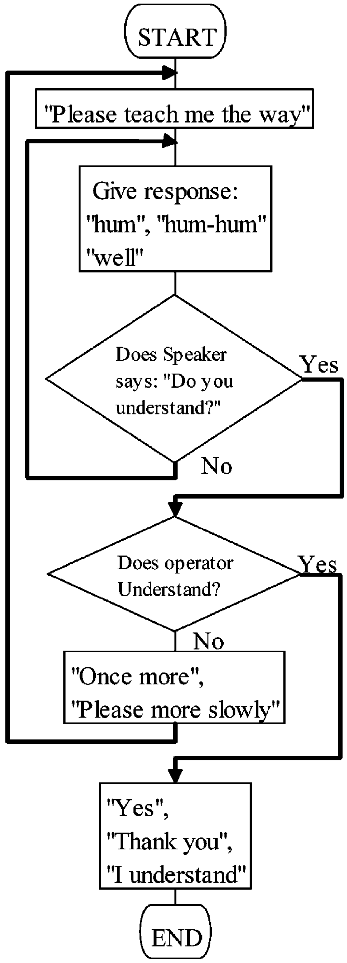

Fig. 9. Switching rule of utterances.

Table 2
Questionnaire for participants

Question #	Question
1	Did you easily recall the route to teach? (Recallability)
2	Did you easily teach the route? (Easiness)
3	Do you think the hearer heard your guidance? (Hearing)
4	Do you think the hearer understood the route? (Understanding)
5	Did you think you shared the route information with the hearer? (Sharedness)
6	Did you understand the hearer's feelings? (Empathy)

Table 3
Results from questionnaire

Condition	Q. 1	Q. 2	Q. 3	Q. 4	Q. 5	Q. 6
H	3.84 (1.67)	4.28 (1.61)	5.96 (0.92)	4.68 (1.22)	5.16 (1.22)	4.80 (1.10)
Rr	3.38 (1.75)	4.10 (1.69)	5.67 (1.13)	4.57 (1.53)	4.90 (1.38)	4.24 (1.38)
Rn	2.65 (1.71)	3.30 (1.52)	4.10 (1.51)	3.70 (1.68)	3.25 (1.41)	3.20 (1.40)
ANOVA results ($F_{(2,63)}$)	$F = 2.58$ $p = .083$ (+)	$F = 2.13$ $p = .127$(n.s.)	$F = 14.33$ $p < .001$ (***)	$F = 2.69$ $p = .076$ (+)	$F = 12.28$ $p < .001$ (***)	$F = 8.31$ $p = .001$ (***)
Multiple comparison	(H > Rn)		H, Rr > Rn	(H > Rn)	H, Rr > Rn	H, Rr > Rn

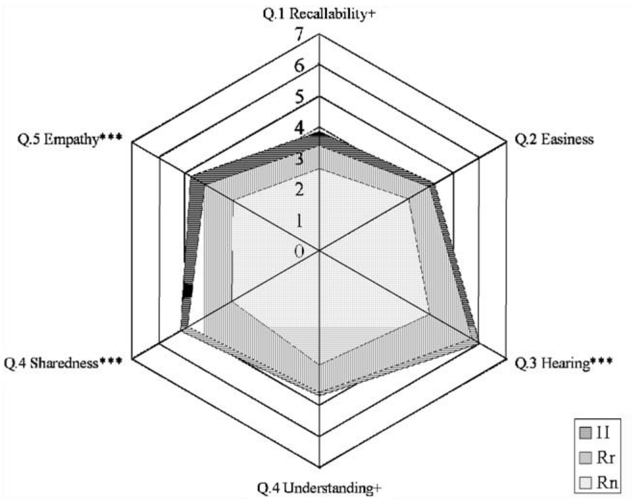

Fig. 10. Comparison of subjective evaluation.

P_{right}, P_{left} are relative three-dimensional points where the right and left finger points differ from both shoulder points, and t is a second. We calculated each condition's average as AVE_{H}, AVE_{Rr} and AVE_{Rn} from this formula.

Table 4 shows the averages for each condition. These results show that there are no significant differences among the three conditions ($F_{(2,63)} = 0.38$, $p = .685418$). However, these results may have individual differences, so we tried to subtract the H condition average from the Rr or Rn condition average for the same subject. This shows the differences in the increase and decrease of Rr and Rn conditions from the H condition. We call these conditions H–Rr condition and H–Rn condition. If this value is 0, the body movement is the same amount in the H condition and Rr or Rn condition. If this value is positive, the body movement in the H condition is larger than in the Rr or Rn condition. If this value is negative, the body movement in the Rr or Rn condition is larger than in the H condition.

Table 4
Averages of finger movements

Condition	H	Rr	Rn
Average/variance	168.52 (104.52)	177.09 (103.04)	148.73 (105.13)
ANOVA results $F_{(2,63)}$	$F = 0.38$, $p = 0.685418$		

Table 5
Result of body movement analysis

Condition	H–Rr	H–Rn
Ave(Var) mm/s	−14.20 (83.00)	84.95 (85.72)
ANOVA results ($F_{(1,16)}$)	$F = 5.52$, $p = .032$ (*)	

Table 5 shows the averages, standard deviations, and results of an ANOVA of each condition's body movements. The ANOVA shows that there is a significant difference in the H–Rr and H–Rn conditions ($F_{(1,16)} = 5.52$, $p = .032$ (*)). This shows that the amount of Rr condition's body movement is larger than in the Rn condition. Furthermore, the Rr condition is almost the same amount of H condition's body movement. This suggests that the robot's cooperative body movement leads to speaker's body movement.

Next, we conducted the video analysis. The video analysis is an evaluation of two indexes by an observer who does not know the experiment's hypothesis. One index is an evaluation of the body direction. We evaluate body direction by the shoulder movement. In this analysis, we classified the subjects into three categories: The speaker's shoulder not moving is rated in 0. The speaker's shoulder moving going along with turns in a different direction is 1. The speaker's shoulder moving along with turns and leg(s) movements is 2. Another index is an evaluation of the appearance of gesture. In this analysis, we also classified the subjects into three categories: Speakers who do not move at all are rated as 0. The speakers who only move their hands are 1. The speakers who move and raise their hands are 2.

Table 6 shows the averages, standard deviations, and results of an ANOVA of the video analysis. The ANOVA shows that there is significant difference in the body directions ($F_{(2,63)} = 3.79$, $p = .028$ (*)). However, we did not confirm that there is significant difference in the appearances of gestures. An LSD method proved that the results of H are significantly better than those of Rn ($\mathrm{Mse} = 0.3809, p < .05$). We believe that the H condition obtains a better value than the Rn condition from the result and tendency of average to be H > Rn. This suggests that the Rr condition may obtain a better value results above those of the Rn condition are the body direction. It is the gesture that increases Rr condition rather than Rn condition is body movement and the amount of arm movement in this experiment. This analysis allows

Table 6
Results of video analysis

	Body direction	Appearances of gesture
H	0.72 (0.78)	1.32 (0.733)
Rr	0.43 (0.58)	1.38 (0.778)
Rn	0.20 (0.40)	1.00 (0.837)
ANOVA results ($F_{(2,63)}$)	$F = 3.79$, $p = .028$ (*)	$F = 1.35$, $p = .267$
Multiple comparison	H > Rn	

us to conclude that the robot's cooperative body movement stimulates speakers to perform cooperative body movement as done in human–human conversation.

5. Discussions

Our results confirm that cooperative body movements performed by a humanoid robot affect humans from an emotional aspect, indicating that cooperative body movements improve understanding of human-robot communication and jointly held information.

This analysis of body movements suggests a relationship between cooperative body movements and the emotional factor. This is because the robot's cooperative body movements stimulate the speakers to perform gestures that are equivalent to those of human–human conversation after evaluating the robot's body movements. On the other hand, a significant difference between humans and the developed humanoid robot was revealed during subjective evaluation and analysis of body movements. Accordingly, to improve the effectiveness of cooperative body movement, we must implement many communicative units and autonomous switching modules to achieve efficient human-like communication.

We did not confirm the effect of cooperative body movement from the aspect of information transmission, through there were nearly a significant differences in Q.4 (Understanding) and almost significant differences in Q.2 (Easiness). Consequently, we cannot conclude that there is no meaning to the Q.2 ($p = .127$) value. That is to say, it is possible that the remaining cooperative body movement is related to the aspect of understanding each other. Another study (Ono et al., 2001) concluded that cooperative body movement affects to aspect of conversationalists understanding each other. They also confirmed that cooperative body movement occurs between humans and the robot when the robot performed suitable body movements. Furthermore, they suggested that the most important factor in communication is the relationship between humans and the robot, wherein cooperative body movement guides this relationship and embodies communication. We consider that this relationship occurred in our experiment, although we need to verify the effects of mutual body movement by producing a robot system that can more effectively perform cooperative body movements.

6. Conclusions

We developed an interactive humanoid robot system that performs cooperative body movements and investigated the effect of the implemented body movements in embodied communication with humans in a route-guidance situation. The experimental results indicated that the cooperative body movements greatly enhanced the emotional impressions of human speakers during route guidance. We believe that our findings will lead to sociable humanoid robots that can engage in talk with humans smoothly and naturally by using their human-like bodies effectively. While these early findings show promise for the usage of a robot's body in embodied communication, it remains one of our future works to develop a completely autonomous mechanism for selecting the appropriate cooperative body movements.

Acknowledgement

This research was supported by the National Institute of Information and Communications Technology. The authors would like to thank anonymous reviewers for giving us very valuable comments and suggestions for revising our paper, and the special issue editors Dr. Noriko Suzuki and Dr. Christoph Bartneck for providing clear guidance.

References

Breazeal, C., Scassellati, B., 2000. Infant-like social interactions between a robot and a human caregiver. Adaptive Behavior 8 (1), 49–74.

Cassell, J., Bickmore, T., Campbell, L., Vilhjalmsson, H., Yan, H., 2000. Human conversation as a system framework: designing embodied conversational agents. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), Embodied Conversational Agents. MIT Press, Cambridge.

Hirai, K., Hirose, M., Haikawa, Y., Takenaka, T., 1998. The development of honda humanoid robot. Proceedings of the IEEE Int. Conf. on Robotics and Automation, 1321–1326.

Imai, M., Ono, T., Ishiguro, H., 2003. Physical relation and expression: joint attention for human-robot interaction. IEEE Transaction on Industrial Electronics 50 (4), 636–643.

Ishiguro, H., Ono, T., Imai, M., Kanda, T., 2003. Development of an interactive humanoid robot "Robovie" -An interdisciplinary approach. In: Jarvis, R.A., Zelinsky, A. (Eds.), Robotics Research. Springer, Berlin, pp. 179–191.

Kanda, T., Ishiguro, H., Ono, T., Imai, M., Nakatsu, R., 2002. Development and evaluation of an interactive humanoid robot "Robovie". IEEE International Conference on Robotics and Automation, 1848–1855.

Kanda, T., Ishiguro, H., Imai, M., Ono, T., 2003. Body movement analysis of human-robot interaction. Proceedings of International Joint Conference on Artificial Intelligence, 177–182.

Kozima, H., Vatikiotis-Bateson, E., 2001. Communicative criteria for processing time/space-varying information. Proceedings of the IEEE International Workshop on Robot and Human Communication.

Matsusaka, Y., Kubota, S., Tojo, T., Furukawa, K., Kobayashi, T., 1999. Multi-person conversation robot using multi-modal interface. Proceedings of the World Multiconference on Systems, Cybernetics and Informatics 7, 450–455.

Moore, C., Dunham, P.J. (Eds.), 1995, Joint Attention: Its Origins and Role in Development. Lawrence Erlbaum Associates, London.

Nakadai, K., Hidai, K., Mizoguchi, H., Okuno, H.G., Kitano, H., 2001. Real-time auditory and visual multiple-object tracking for robots. Proceedings of the International Joint Conference on Artificial Intelligence, 1425–1432.

Nakano, Y.I., Reinstein, G., Stocky, T., Cassell, J., 2003. Towards a model of face-to-face grounding. Proceedings of the Association for Computational Linguistics, 553–561.

Ogawa, H., Watanabe, T., 2001. InterRobot: speech-driven embodied interaction robot. Advanced Robotics 15 (3), 371–377.

Ono, T., Imai, M., Ishiguro, H., 2001. A model of embodied communications with gestures between humans and robots. Proceedings of the Twenty-third Annual Meeting of the Cognitive Science Society 8, 732–737.

Scassellati, B., 2000. Investigating Models of Social Development Using a Humanoid Robot, Biorobotics. MIT Press, Cambridge.