Robovie: an Interactive Humanoid Robot

Hiroshi Ishiguro[1,2], Tetsuo Ono[1], Michita Imai[1],
Takeshi Maeda[1], Takayuki Kanda[1], Ryohei Nakatsu[1]
(1) ATR Media Integration & Communications Research Laboratories
(2) Department of Computer & Communication Sciences, Wakayama University,
930 Sakaedani, Wakayama 640-8510, Japan
E-mail: ishiguro@sys.wakayama-u.ac.jp

Abstract

We have developed a robot called "Robovie" that has unique mechanisms designed for communication with humans. Robovie can generate human-like behaviors by using the human-like actuators and vision and audio sensors. In the development, the software is a key. We have obtained two important ideas in human-robot communication through research from the viewpoint of cognitive science; one is importance of physical expressions using the body and the other is effectiveness of the robot's autonomy in robot's utterance recognition by humans. Based on these psychological experiments, we have developed a new architecture that generates episode chains in interactions with humans. The basic structure of the architecture is a network of situated modules. Each module consists of elemental behaviors to entrain humans and a behavior for communicating with humans.

1.Introduction

There are two research directions in robotics; one is to develop task-oriented robots that work in limited environments and the other is to develop interaction-oriented robots that collaborate with humans in open environments. Industrial and pet robots are the former ones. They perform particular tasks such as assembling industrial parts, behaving like an animal, and so on. On the other hand, the purpose of the robot that we are developing is not to perform particular tasks. We are trying to develop a robot that exists as our partner in our daily life. The fundamental requirement of humans in our daily life is to communicate and recognize the existence each other. Our robot supports such an aspect of our life and pro-

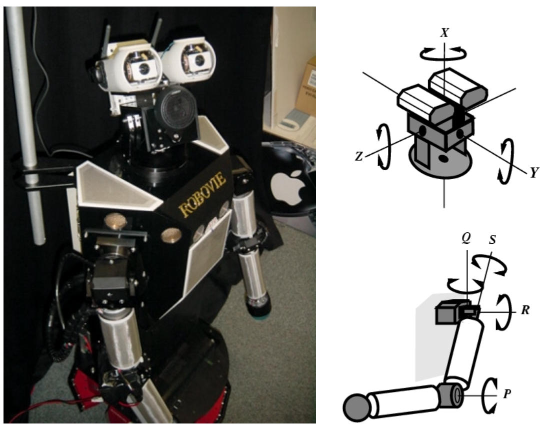

Fig. 1: Robovie

vides rich information to humans by using the communication functions. We consider, the robots existing as our partners will be a new information infrastructure for communication.

2.Hardware for interactions with humans

The robot that has a human-like appearance is designed for communication with humans. It equips with various sensors, such as vision, sense of touch, audition and so on. By using the human-like body and sensors, the robot can perform interactive-behaviors.

Fig. 1 shows the developed robot. It is a humanoid-type robot that moves with two driving wheels. The size is important as an interactive robot. Not to give an awful impression to humans, we have decided the size as 120 cm, which is same as a junior school student. The diameter is 40 cm and the weight is about 40 Kg. The robot has two arms (4*2 DOF), a head (3 DOF), two eyes (2*2 DOF for gaze control), and a mobile platform (2 driving wheels and 1 free wheel). Further, the robot has various sensors, skin sensors covering the whole body, 10 tactile sensors around the mobile platform, an omnidirectional vision sensor, two microphones to listen human voices, and 24 ultra-sonic sensors for detecting obstacles. The eye has pan-tilt mechanism with direct-drive motors and they are used for stereo vision and gazing control. The skin sensors are important for realizing interactive behaviors. We have developed a sensitive skin sensors using pressure sensitive conductivity rubber. Another important point in the design is the battery life. This robot can work 4 hours and charges the battery by autonomously looking for battery stations. With the actuators and sensors, the robot can generate almost all behaviors needed for communication with humans.

Fig. 2 shows several characteristic hardware components of Robovie. Generally, the robot needs 7 DOF for the arm to generate human-like gestures. However, it is possible to generate almost all of them by attaching 3 DOF arms with an angle of 45 degrees on the shoulder as shown in Fig. 2(a). The mobile plat form has two driving wheels and a free wheel. It is compact enough for moving in office and home environments as shown in Fig. 2(b). The unique sensor of the robot is the tactile sensor as shown in Fig. 2(c). The binocular vision shown in Fig. 2(d) is also unique. The camera has two direct-drive motors for generating human-like eye motions. The omnidirectional camera shown in Fig. (e) is different from the

(a) Arm and shoulder (b) Mobile platform (c) Tactile sensor

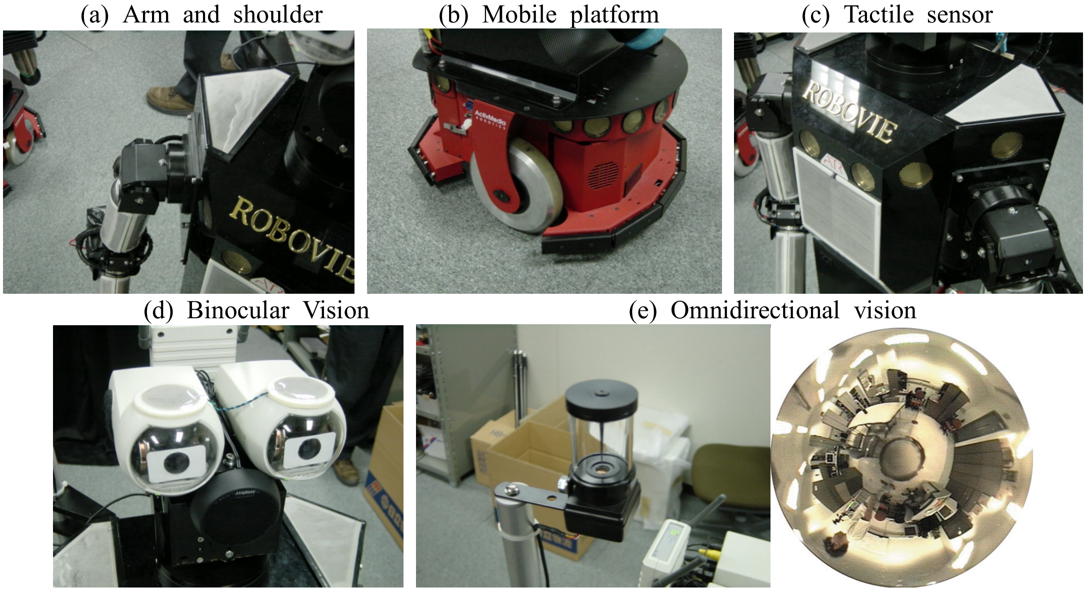

Fig. 2: Hardware components

human vision, but it plays an important role in the visual recognition. We, human, can recognize the surrounding environment quickly by using not only the vision but also rich knowledge and reasoning ability. However, it is not easy to implement such a function to the robot. The omnidirectional camera directory observes around the robot for memorizing the location and finding moving objects.

With respect to the computing resource, Robovie is a self-contained autonomous robot. It has a Pentium III PC on board for processing sensory data and generating gestures. The operating system is Linux. Since the Pentium III PC is sufficiently fast and Robovie does not require precise real-time controls like a legged robot, Linux is the best solution for easy and quick development of Robovie's software modules.

3. Basic ideas for designing interactive behaviors

Mutual entrained gestures are important for smooth communications between Robovie and a human. We have performed psychological experiments to ensure it. The aim of the experiments was, concretely speaking, to investigate correlations between body movements and utterance understanding in human-robot communications. We focused on the interactions between a subject and a robot while it teaches a route direction to the subject, and investigated the appearance of the subject's gestures and the level of the utterance understanding by using several different gestures in the teaching.

We conclude the experimental results as follows:

1. Many and various behaviors of the robot induce various human communicative gestures. In other words, the subject's gestures are increased by entrainment and synchronization with the robot; and a relationship between the robot and the subject is established from the mutual gestures.
2. The emerged mutual gestures help to understand robot's utterance.
3. The joint viewpoint represented by the robot gestures allows the subject to understand the utterance.

The experiment clarified the importance to share a joint viewing point in human-robot communication. The results suggest proper robot behaviors in the development of the robot. The concept of the joint viewing point can be extended as the concept of joint attention and it gives more proper robot behaviors for interacting with humans.

The relevance theory [3] proposes a communication model for recognizing situations and the humans' experiences. It employed a new term named mutual manifestness that represents mental states where two or more humans recognizes the same situation or recalls similar experiences. The relevance theory regards human's communications as a process of gaining mutual manifestations by passing messages to others. This concept of mutual manifestness is same as that of focus of attention; and it is called "joint attention" in social psychology [4] in the case where people frequently focus on the same object while communicating each other.

We have verified the effect of the joint attention. First of all, we have prepared two groups: one was given Robovie with eye contact, and the other was given Robovie without eye contact. Robovie performed the attention expression for both groups. The target of the attention expression was a poster on a wall. We recorded the number of subjects who looked at the poster according to the attention expression. As the results, we could verified that eye contact is significantly effective for achieving joint attention; and the robot behaviors using the eye contact are proper for establishing a communicative relationship with a human.

From these psychological experiments, we have obtained four ideas as follows:

1. Rich robot's behaviors induce various human communicative gestures that help utterance understanding.

2. Attention expression by the robot guides the human's focus to the robot attention.
3. Eye contact by the robot indicates robot's intention of communication to the human.
4. Sharing of a joint viewing point and a proper positional relation establish a situation where the human can easily understand robot's utterance.

4. A Robot Architecture for Generating Interactive behaviors

Based on these ideas, we have designed a new architecture of the robot and implemented to the developed robot "Robovie". The basic structure of the architecture is a network of situated behavior modules. Fig. 3 shows the meta-structure of Robovie's software. All of the behaviors are classified into four categories; and Robovie performs behaviors belonging to one of them. A unique point is that the category "Play with humans" has two sub-categories of greeting to say "Hello" or "Bye" when switching the category.

The behavior models belonging to the category include elemental behaviors for communications as shown in Fig. 4. The elemental behaviors that implement the above-mentioned ideas are the most important point in this architecture. The robot behaviors developed so far do not have the function to entrain humans into the communication. By combining the elemental behaviors and other task-oriented behaviors, we can realize various interactive behaviors. Fig. 5 shows an example of the interactive behavior that the robot asks a human to look at a poster.

Fig. 6 shows the all-over software architecture. Basically, this is an extension of the architecture based on situated modules [5]. The architecture proposed in our previous work has two merits: easy development of behavior modules and robust execution by dynamic switching of the behavior network.

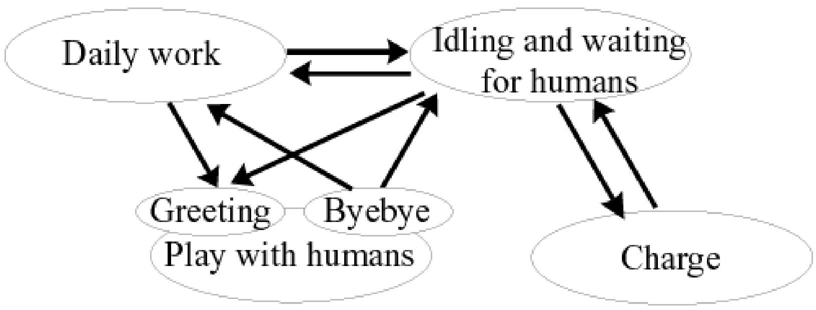

Fig. 3: The meta-structure of the architecture

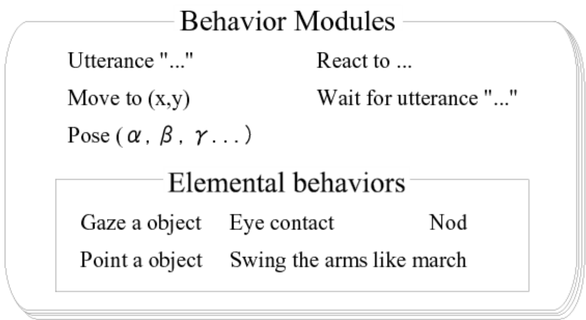

Fig. 4: Behavior modules

These two ideas are also important to implement sensory data processing. In previous robotics, robots needed to perform perfect sensor data processing to execute particular tasks. The robot in this paper, however, entrains humans into the interaction loops by the interactive behaviors and it does not require perfect sensory data processing. Humans, rather, adapt to the robot's ability. With the architecture, the robot can continuously generate rich behaviors for communication even if the sensory data processing is not perfect.

Finally, Fig. 7 shows all of developed modules and their relationships. The robot behaviors generated from the various interactive behaviors and the complicated network has given human-like impressions. The typical behaviors of Robovie are: "greeting", "hand-shake", "playing the game of 'paper, stone and scissors' ", "hugging", "kiss", "short conversation", "exercise", "pointing the poster", and "saying good bye". Robovie also takes idling behaviors such as "scratching the head", "folding the arms", and so on.

Fig. 8 shows a scene where two Robovies interact with a chilled. In the beginning, the child was concentrated on drawing something on the notebook. But she focused on the interaction when her mother started interacting with the robot (the 1st row of the left column). Then she began playing (from the 1st row of the center column to the 2nd row of the right column). While talking each other, she imitated ro-

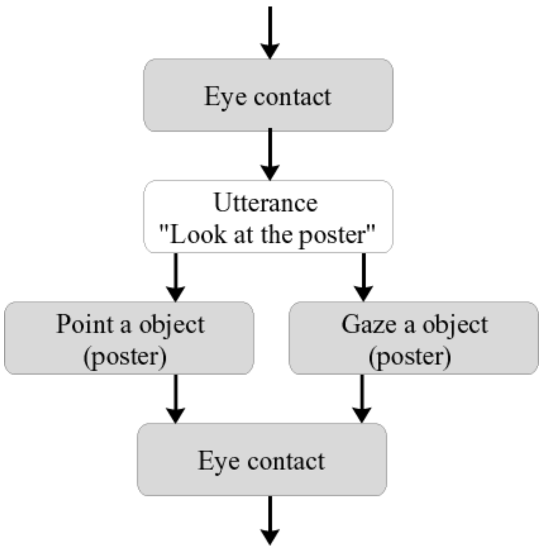

Fig. 5: An example of behavior module

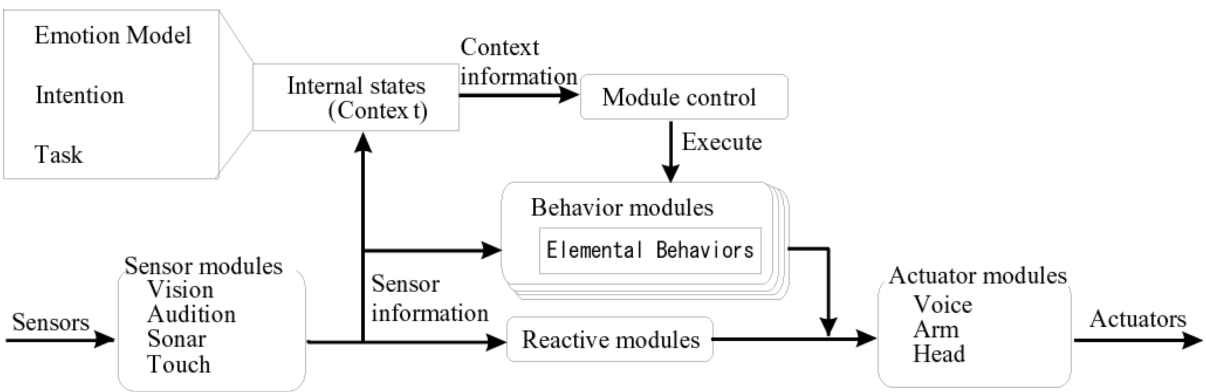

Fig. 6: Software architecture based on behavior modules

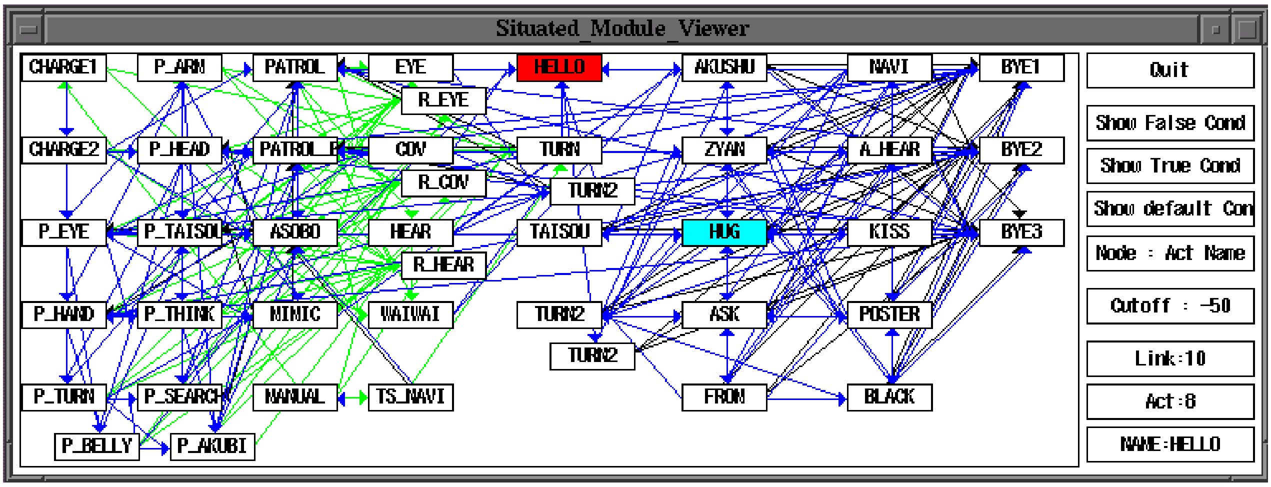

Fig. 7: All behavior modules and their relationships

tated robot's behaviors (the 3rd row of the left and center columns) and hook robot's hands (until the 4th row of the right columns). When the robot pointed a poster, she watched the poster. This interaction never happens without the sophisticated eye motions of the robot (three images in the 5th row). As shown in the figure, the child naturally communicated and played with Robovie. It is difficult to explain how impressive this scene is, but we have felt strong possibilities of robotic creatures.

5. Conclusion

This paper has reported on a new humanoid robot called "Robovie". The unique aspect of Robovie is the mechanism designed for communication. Robovie can generate human-like behaviors with the actuators and sensors. In the design, we have performed two psychological experiments and developed the behaviors obtained from them. Our next step is to implement more interactive behaviors to Robovie and try to establish more sophisticated relationships between the robot and humans.

We have started this project on August 1999. After the development of Robovie on July 2000, Robovie has appeared in many robot exhibitions and been reported by almost all major newspapers and several TV programs in Japan. These are not only advertisements but also valuable chances to gather comments from ordinary people. For developing a robot work in our daily life, these activities bring much information in addition to the cognitive experiments. For more detail of this project, please refer to the following WebPages. It includes many videos where Robovie is interacting with humans.

http://www.mic.atr.co.jp/~michita/everyday-e/

References

[1]David McNeill, "Psycholinguistics: A New Approach," Harper & Row, 1987.

[2]T. Ono, M. Imai, and R. Nakatsu, "Reading a robot's mind: a model of utterance understanding based on the theory of mind mechanism," Advanced Robotics, Vol.14, No.4, pp.311-326, 2000.

[3]D. Sperber and D. Wilson, "Relevance: Communication and Cognition," Oxford: Basil Blackwell, 1986.

[4]C. Moore and P. J. Dunham, "Joint Attention: Its Origins and Role in Development," Lawrence Erlbaum Associatcs, Inc., 1985.

[5]H. Ishiguro, T. Kanda, K. Kimoto, and T. Ishida, "A Robot Architecture Based on Situated Modules," Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 1617-1623, 1999.

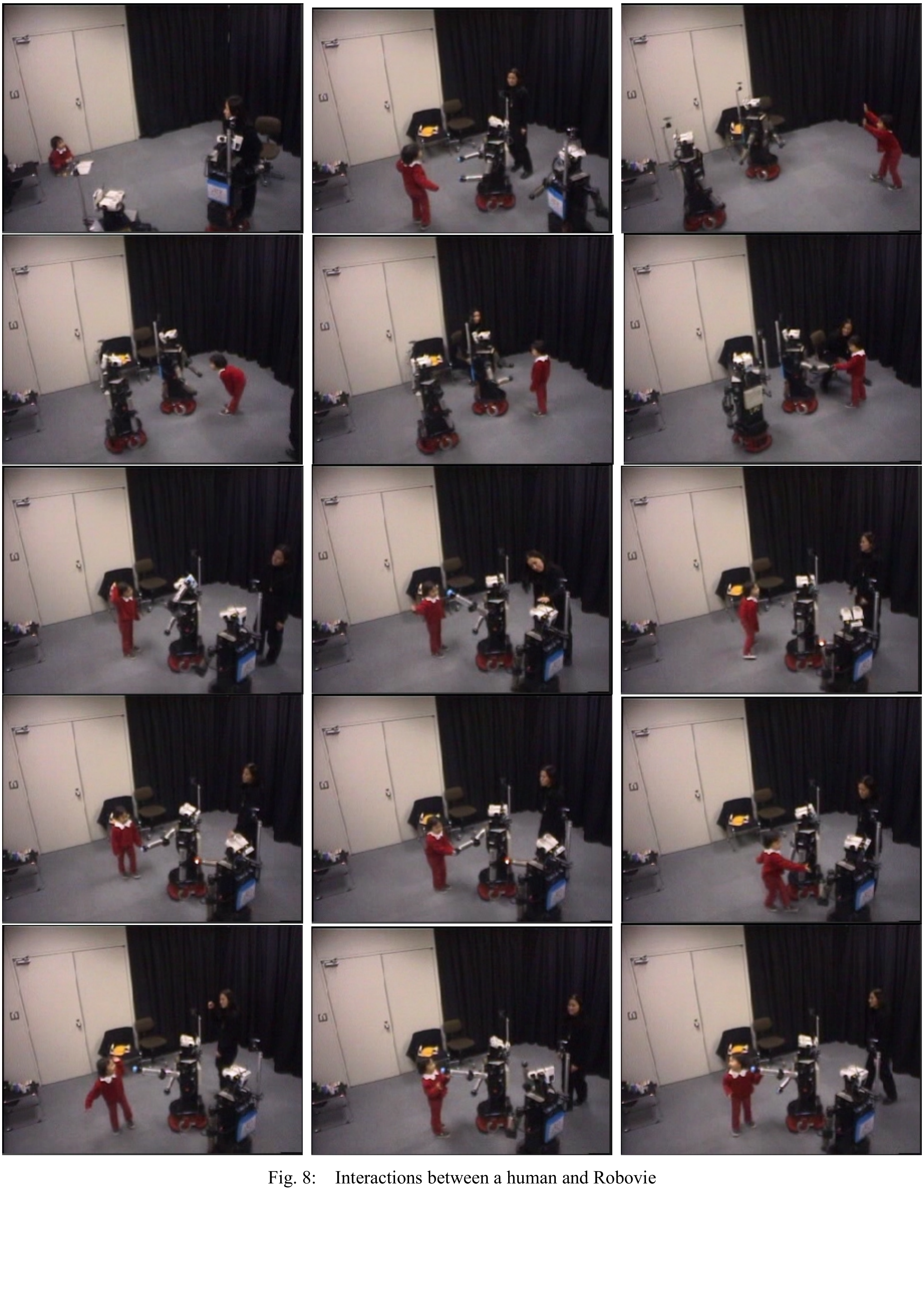

Fig. 8: Interactions between a human and Robovie