Dynamic Communication of Humanoid Robot with multiple People Based on Interaction Distance

Tsuyoshi Tasaki, Shohei Matsumoto, Hayato Ohba, Mitsuhiko Toda,
Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno
Graduate School of Informatics, Kyoto University
Yoshida-honmachi, Sakyo-ku
Kyoto, 606-8501, Japan
E-mail {tasaki, shohei_m, hayato, mtoda, komatani, ogata, okuno}@kuis.kyoto-u.ac.jp

Abstract

Research on human-robot interaction is getting an increasing amount of attention. Since almost all the research has dealt with only communication between one robot and one person, there have been quite few discussions about communication between a robot and multiple people. This paper proposes a method which enables robots to communicate with multiple people using the 'selection priority of the interactive partner' based on the concept of 'Proxemics'. In this method, a robot changes active sensory-motor modalities based on the 'interaction distance' information. The proposed method is implemented into a humanoid robot SIG2 using subsumption architecture. SIG2 has various sensory-motor modalities to interact with humans. A demonstration of SIG2 showed that the proposed method works well during interaction with multiple people.

1 Introduction

Studies concerning human-robot interaction are represented by Sony-Aibo, Robovie [1], Kismet [2], and Wamoeba [3] have gotten a lot of attention. Since almost of them have dealt with only one-on-one communication between one robot and one person, there have been quite few discussions on the methodology for the communication between a robot and multiple people. Considering the yet-to-be developed human-support robot, it would be expected that a robot be able to interact effectively with multiple people at the same time. This paper proposes a design method for such humans-robot communication.

Basically, robots cannot communicate with multiple people at the same time, except when these people can be regarded as one unit as with an audience in the classroom in a lecture, and so on. People select the 'interactive partner' dynamically, based on various criteria such as "intimacy". People also change their communication strategy of sensing and their behavior according to the situation. For example, if the distance between people is small, they can identify the other individual easily, speech recognition and facial expressions are effective for communication. If the distance between them is great, they would use body gestures etc. Thus people's personal space, or 'interaction distance' is an important criterion for selecting appropriate sensory-motor function.

In this paper, we propose a method of humans-robot dynamic communication in which the robot selects an interactive partner from multiple people by using the 'priority' based on the interaction distance. In this method, the robot refines recognition and behavior by selecting an appropriate sensor and motion based on interaction distance.

In Section 2, 'Proxemics', a social psychology theory is introduced as a basic concept of our method and the details of our method are described. In Section 3, a humanoid robot used in this study and actual implementation used subsumption architecture (SA) are described. In Section 4, some demonstrations of the robot behavior for communicating with multiple people are described. In Section 5, the effects of the proposed method are discussed. Section 6 concludes this paper.

2 Communication based on Interaction Distance

We focused on "Proxemics [4]" and attempted to propose a novel concept for a robot to respond appropriately to each person in a group of people based on the distance between the robot and each person. We propose our concept of "Intimacy" which reflects the relationship between the robot and people in terms of behavior.

2.1 Categorization of Robot Functions based on Proxemics

We divided the various functions of the humanoid robot into four groups based on the distances listed in Table 1. For example, if a target person is standing far from the robot, the robot cannot use both speech recognition and face recognition because these functions require highly reliable sensory information.

Proxemics is a social psychology theory which posits that two humans interact at an appropriate physical distance from one another based on their relationship. In this theory, an interaction distance is classified into roughly four groups: intimate distance, personal distance, social distance, and public distance.

At an 'Intimate distance (approx. 50 [cm])', people can communicate via physical interaction and express strong emotion. At a 'Personal distance (approx. 50-120 [cm])', people can talk intimately. People maintain 'Social distance (approx. 120-360 [cm])', when they are talking and don't know each other. At a 'Public distance (approx. 360 [cm] -)', people who have no personal relationship with each other can comfortably coexist. The distance values shown in parentheses are just typical examples. These depend on the person's personality and cultural background.

2.2 Robot Intimacy based on Proxemics

Proxemics suggests that the more intimate the communication, the nearer the target person stands. The parameter of intimacy is introduced to reflect the relationship between a robot and humans. The robot uses this parameter to determine communication priority among multiple people who are in the same situation, and then behaves appropriately based on its relationship with each person.

The I is a parameter of 'Intimacy' of which the range is 0 to 1. It represents how intimate each partner is. Since the I changes dynamically during the communication, its level is modified by using the following equation:

$$I(0) = P \quad (1)$$

$$\frac{dI}{dt} = \left(\frac{I+P}{2}\right) \cdot D - I \cdot \left(\frac{P \cdot I + 1}{2}\right) + S_k \quad (2)$$

The P is the constant parameter defined a priori as the robot personality. The first term of Eq. (2) shows the influence of the distance. The D is the parameter of the distance designed as follows:

$$D = \begin{cases} 0.04 & \text{(if Intimate distance)} \\ 0.02 & \text{(if Personal distance)} \\ 0.0 & \text{(otherwise)} \end{cases} \quad (3)$$

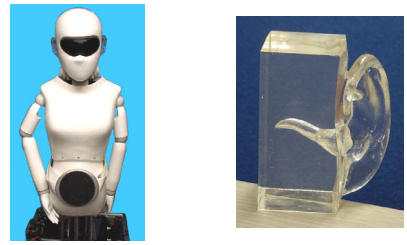


Figure 1: SIG2(left) and its ear(right)

The I is defined as the summation of friendliness of the robot and intimacy of the person. If the robot recognizes the person it is intimate with, the I increases and vice versa. The second term of Eq. (2) is a damping factor. If the robot has no communication with people for a while, the I converges to 0. The S_k is a parameter of the influence of stimulus. It changes I based on the human behavior. The proposed method influencing stimulus includes greeting (increase), brushing (increase), and hitting (decrease).

3 Humanoid SIG2 and Implementation

We implemented the proposed method into the humanoid robot SIG2 shown in Figure 1. Our method dynamically determines the priority of various modalities of sensory system and motor system, based on the interaction distance. The Subsumption Architecture (SA) [5] is used for the implementation, taking into consideration the hierarchy of the distance listed in Table 1.

3.1 Humanoid SIG2

SIG2 has two microphones on both sides of the head. Each microphone is covered with silicon in the shape of the human external ear (Figure 1). The head and the upper body are covered with soft skin containing 19 touch sensors. And an ultra-directional parametric speaker is equipped around its waist.

The touch sensor, which consists of piezo elements covered by silicon, can detect the pressure velocity. It can recognize three kinds of contact: touch, brush, and hit. SIG2 can measure the distance to the partner by stereovision which is composed of two cameras equipped in the head. Based on the interaction distance, SIG2 selects the sensory functions as listed below.

- Public distance, Social distance

  At these distances, SIG2 can locate humans using skin color information of the vision system and locate sound sources by integrating Interaural

Table 1: Relation between Distance and Function

	Intimate distance	Personal distance	Social distance	Public distance
Input	sense of touch face detection speech recognition face localization sound localization	face detection speech recognition face localization sound localization	face localization sound localization	face localization sound localization
Output	utterance	utterance action	utterance action	utterance action moving

Intensity Difference (IID) and Interaural Phase Difference (IPD) using the Dempster-Shafer theory [6].

- Personal distance

  At this distance, besides the functions mentioned above, SIG2 carries out speech recognition and face recognition (Figure 2) based on Mahalanobis distance between the detected face and the face data which was registered beforehand.

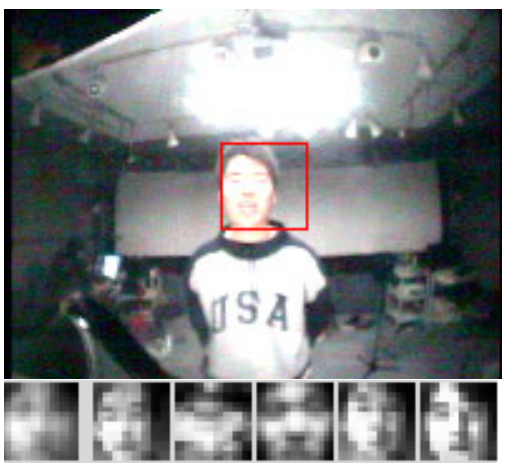


Figure 2: Face Recognition and Localization

- Intimate distance

  At this distance, besides the functions used in calculating personal distance, SIG2 recognizes three kinds of contact: touch, brush, and hit, (Figure 3) using absolute values and the number of zero-crossing points of output of touch sensor.

SIG2 has a four-degree-of-freedom rotation such as nod, incline, rotation of its neck and rotation of its

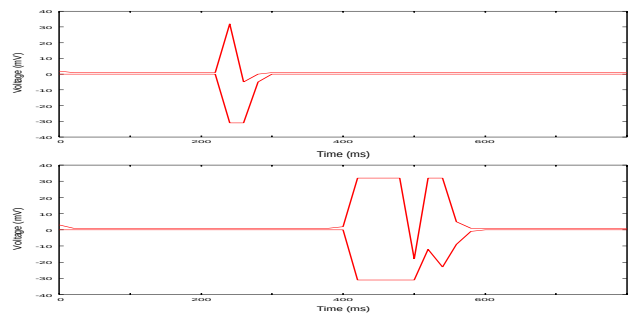


Figure 3: Touch sensor, hit(above) and brush(below)

body as movement functions. SIG2 also has movement by the cart, utterance enabled by the two kind of speakers (directional, omnidirectional), as movement functions.

Based on the distance to the person, SIG2 selects the movement functions as listed below.

- Intimate distance
  At this distance, SIG2 uses the omni-directional speaker for speaking.

- Personal distance
  At this distance, besides the omni-directional utterance, the gesture facilitated by four motors is effective.

- Social distance
  At this distance, besides the functions used in personal distance, it is effective to use the directional speaker to speak to a person standing far away from the SIG2.

- Public distance
  At this distance, besides the functions used in social distance, SIG2 can move via the cart to get close to the target person or people.

3.2 Implementation by Subsumption Architecture

The SA is effective for the implementation of the proposed method considering the hierarchal structure listed in Table 1. This enables SIG2 to process sensor information efficiently.

All sensory information is sent to all action modules. Each action module processes input information in parallel to output the result. The output of upper modules suppresses or inhibits that of lower modules to subsume the output of action modules. By implementation of the top module which inhibits outputs of lower modules based on the interaction distance, the dynamic modality-selection can be achieved (see Figure 4).

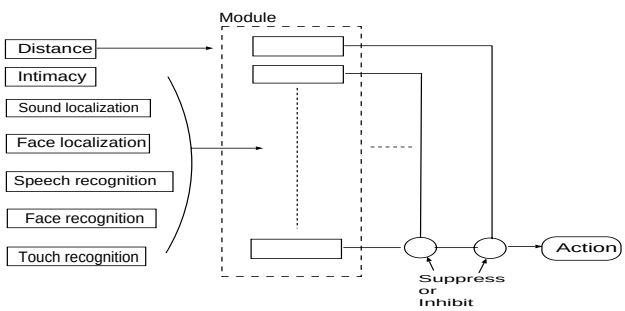


Figure 4: System Outline

4 Experiments

We conducted two experiments to ascertain the functions of the proposed method. The outlines of these experiments follow below.

4.1 Scene 1 : Selecting sensory modalities by distance

In this experiment, SIG2 interacts with two people who talk from different distances, far and near, by changing input modalities. The SIG2 urges the farthest person from the robot to approach. Figure 5 shows the structure of SA used in this section.

Step 1 Person A said "Hello SIG2." at a social distance.

As a result of sound localization, SIG2 turns to Person A (turnFaceToSound) and as a result of face localization, it continues looking at him (lookAtFace). It detects he is positioned at a social distance by the stereovision. Consequently it inhibits the act of calling his name (greetWithName) and replying with a greeting (replyGreet).

Step 2 Person B approached with in an intimate distance and said "Hello SIG2".

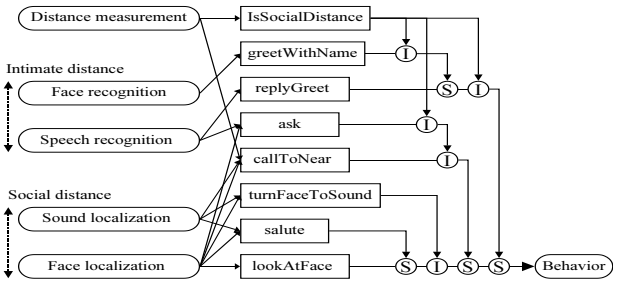


Figure 5: Implemented Modules

As a result of sound localization and face localization, SIG2 turns to Person B and continues looking at him. Because Person B is positioned at an intimate distance, SIG2 bows slightly (item salute), calls Person B's name, and greets Person B.

Step 3 Person A called to SIG2.

As a result of sound localization and face localization, SIG2 turns to Person A and continues looking at him. Greeting Person A is inhibited by the history of Person A's behavior, and requesting Person A to approach (callToNear) and asking about Person A's business (ask) become active. However asking about Person A's personal business is inhibited because of social distance.

Step 4 Person A followed instructions and approached SIG2.

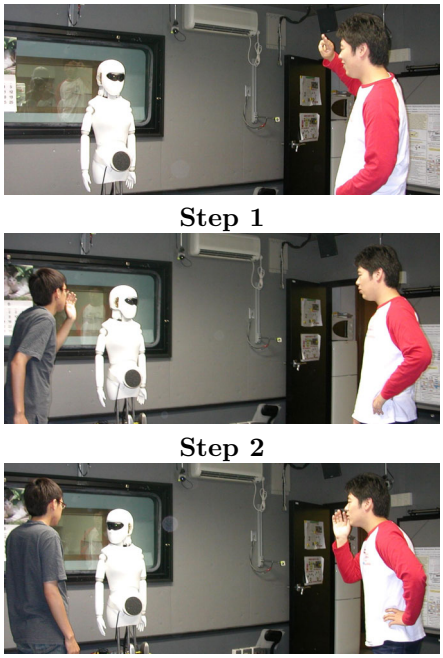
As a result of face localization, SIG2 continues looking at Person A. SIG2 detects Person A is positioned at an intimate distance. Then it asks Person A's business and inhibits requesting Person A to approach.

4.2 Scene 2 : Changing behavior by the intimacy

In this experiment, SIG2, between two people, changes its conversation partner based on intimacy. Figure 7 shows the SA used in this section.

Step 1 Person A greeted within an intimate distance.

As a result of sound localization and face localization, SIG2 turns to Person A and continues looking at him. SIG2 bows slightly, calls Person A's name and replies to Person A because of their intimate distance. The intimacy with Person A increases.



Step 1

Step 2

Step 3

Figure 6: Scene 1

Step 2 Person B greeted from a social distance.

As a result of sound localization and face localization, SIG2 turns to Person B and continues looking at him. Calling his name and offering a greeting is inhibited because of social distance. SIG2 compares the intimacy it experienced with Persons A and B, and then returns to Person A, with whom it has higher intimacy (turnToIntimatePerson).

Step 3 Person A brushed SIG2.

The intimacy with Person A increases. (updateIntimacy)

Step 4 Person B called SIG2 from a social distance.

As a result of sound localization, SIG2 turns to Person B with increasing frequency. However SIG2 continues looking at Person A since the intimacy with Person A is over the threshold value (turnToIntimatePerson). SIG2 does not reply to Person B.

Step 5 Person A banged SIG2.

The intimacy with Person A decreases, and SIG2 avoids him. (avoidHostilePerson)

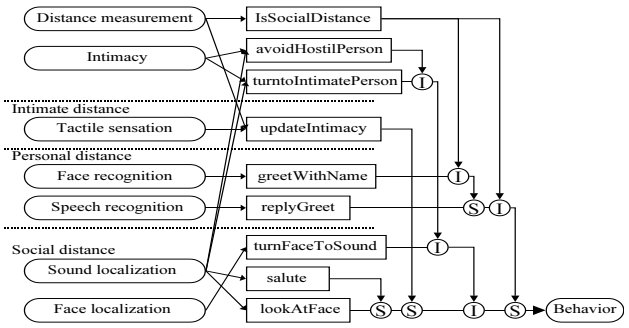


Figure 7: Implemented Modules

5 Discussion

In this section, we discuss the effects of the proposed method.

5.1 Efficiency of design and operation

Selecting sensory modalities using distance information, we can design robot behaviors without considering all combinations of all modalities. Our method is also very smart at computing costs compared with the way of selecting behavior in consideration of all possible modalities.

5.2 Efficiency of communication

We make robots avoid incorrect recognition by selecting sensors and behavior using distance information, therefore we expect the communication of exact information. As a result, we reduce the number of interactions and enable efficient communication.

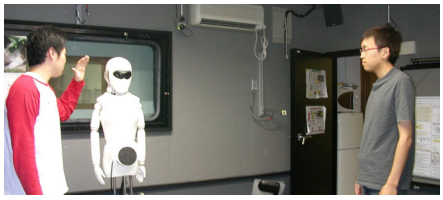
5.3 Priority based on interaction distance

The experiments of the humanoid robot showed that the proposed method enables the robot to select an appropriate target-person in communication by dynamically changing the intimacy. The kinds of demonstrations described in this paper would be a natural style of communication of a human-supported robot in the future.

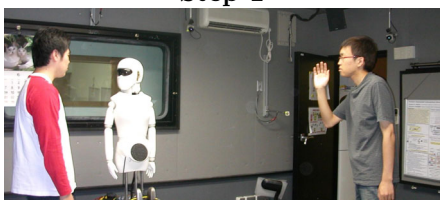
6 Conclusion and Further Work

This paper proposed a model of robot intimacy based on the interaction distance as a criterion to determine the communication priority between multiple people. This method was implemented in a humanoid robot called SIG2 using SA. We discussed the effectiveness of a proposed method by referring to the demonstrations of SIG2.

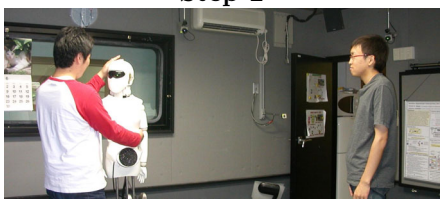
There are three main areas of focus for the future. The first area is the introduction of the 'reliability of sensory information'. Since the robot just selects a sensor modality using the proposed method, the unselected sensors are not used at all. It is thought



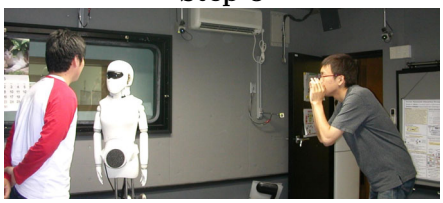
Step 1



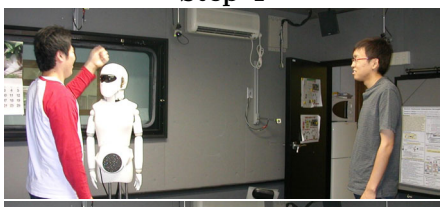
Step 2

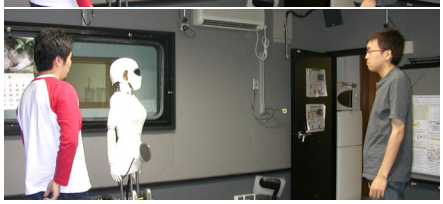


Step 3



Step 4





Step 5

Figure 8: Scene 2

that the heterogeneous sensors should be integrated according to their reliability to generate appropriate robot behavior. We attempt to design the reliability of the sensor modality using not only the interaction distance but also other environmental information, such as sound noise, brightness, and so on.

The second area is the definition of 'intimacy'. The intimacy in the proposed method completely depends on interaction distance. However, it is evident that the intimacy should be influenced by variations in communication. For example, though 'hitting' occurs at an 'Intimate distance', it does not mean intimacy. We should design a more appropriate model which incorporates the machine learning technique.

The third area requiring attention is the definition of an evaluation method for this kind of humans-robot communication. Almost all conventional studies concerning robot communication with 'a person' have used a subjective impression derived from the questionnaire as evaluation criteria. It is quite difficult, however, to deal with such subjective impressions of 'multiple people' in a complete evaluation. We should consider the way of analysis of the dynamic transition of communication between a robot and multiple people.

Acknowledgements This research was supported by the JPSP 21st Century COE program on informatics research for development of knowledge society infrastructure.

References

[1] H. Ishiguro, T. Miyashita, T. Kanda, T. Ono and M. Imai, "Robovie: An interactive humanoid robot", *In Video Proc. of 2002 IEEE International Conference on Robotics and Automation*, 2002.

[2] Breazeal. C. L "Designing Sociable Robots Intelligent Robots and Autonomous Agents", *A Bradford Book*, 2001, ISBN 0262025108.

[3] Tetsuya Ogata, Shigeki Sugano "Emotional Communication between Humans and the Autonomous Robot WAMOEBA-2 Which has the Emotional Model", *JSME International Jounal, Series C, Mechanical Systems Machine Elements and Manufacturing*, Vol.43, No.3, pp. 568-574, Sept. 2000.

[4] Edward T. Hall "Hidden Dimension", *Doubleday Publishing*, 1966.

[5] Rodney A. Brooks "A Robust Layered Control System For A Mobile Robot", *IEEE Journal of Robotics and Automation*,pp.14-23,Vol.2,No.1,1986.

[6] Hiroshi G. Okuno, Kazuhiro Nakadai,Ken-ichi Hidai, Hiroshi Mizoguchi, Hiroaki Kitano "Human-Robot Interaction Through Real-Time Auditory and Visual Multiple-Talker Tracking", *IROS*, 2004.