Developing Helper Robot for Senior Care through Sociological Analysis of Human Interaction

Hitoshi NIWA † Naonori AKIYA ‡ Michie KAWASHIMA †

Yoshinori KUNO † Keiichi YAMAZAKI ‡

† : Graduate School of Science and Technology, Saitama University
‡ : Graduate School of Cultural Science, Saitama University

Abstract We have developed a robot that is able to respond appropriately to user's verbal and non-verbal requests. For practical use of helper robots such as in elderly care. However, it is necessary to analyze how elderly persons in naturally occurring settings actually make requests to caretakers in order to enhance the development of user-friendly robots in elderly care. In this paper we report on our observations of interaction between elderly persons and caretaker in a senior care home. We find simultaneous actions coordinated by seniors and helpers. Our analysis shows an importance of projectability in a senior care home. Based on this analysis, we will develop a robot which can understand such projectability in the cooperative work among the senior and the helper.

1 Introduction

In a rapidly aging society, lack of human caretakers is a major social and health issue. Because of this, there will likely be a rising demand for helper robots in the near future. Thus, it is necessary to develop a mechanism and control system to enable robots to perform caretaking tasks. For this purpose, we recognize that one of the most fundamental tasks a robot should be able to do is appropriately respond to mundane requests by the user. In order to respond appropriate requests, naturally a robot will have to be able to take in verbal dialogue.

A previous study has shown the importance of a dialogue interface in helper robots [1]. Dialogue interface has several problems of speech recognition and language understanding. While dialogue interface has room for improvement by working more on speech recognition, we also think that the dialogue interface can be improved by developing robots that can recognize the potential meanings of non-verbal actions. In making requests, human beings often concurrently use both verbal and non-verbal means. For example, a human might point to something while saying “bring that!” to a listener. In this utterance we understand “that” as the thing being pointed to in the immediate scene.

In our research, we have developed a robot that respond to such simplified utterances taking into account both verbal and visual information [2]. In particular, this robot tracks both the human gaze and pointing direction, and detects the object or objects in its direction.

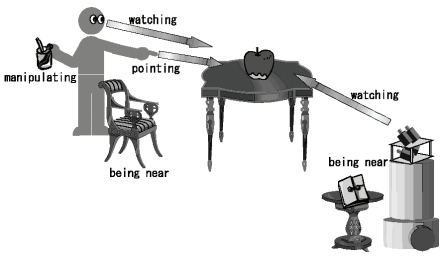

Fig. 1: Behaviors used for speech understanding.

Such requests, however, are not always made so

explicitly. In real time a human helper can interpret subtle bodily movements from elderly persons as requests, even when there is nothing said. Thus, it is highly desirable for a helper robot to be able to interpret and respond to non-verbal movements as potential requests even when there is no verbal information. In order to develop such a robot, we first investigated interaction between caretakers and elderly persons. Our analysis of caregiver-elderly interaction is based on an ethnomethodological approach. Ethnomethodology is a framework for investigating how people achieve mutual understanding in interaction [3]. Through the results of this analysis, we aim at developing a helper robot that is able to respond to verbal and non-verbal requests.

This paper presents the initial stage of developing this type of robot. The outline is as follows. In section 2, we describe a robot that can appropriately respond to utterances that depend upon visual information. In section 3, we describe the relationship between ethnomethodology and developing a helper robot. In section 4, we describe the results of an analysis of everyday interaction in an elderly care home. In section 5, we discuss the development of a helper robot based on the results of the previous analysis.

2 Utterance Understanding Based on Action Recognition

We have developed a robot which understands user's verbal and non-verbal requests [2]. This robot understands "simplified" utterances by taking into account visual information. For example, the robot can understand the words "this", "that" by taking into account the user's point.

Requests can be considered as composed of an "object" (e.g. apple, radio) and "verb" (e.g. bring, turn on). For each request, a robot judges three cases, (1) are the noun and verb are explicit? If either is not explicit (i.e. inexplicit), (2) is the object a pronoun, and is the verb a pro-verb?, (3) are the object and verb omitted? In cases (2) and (3), we assume that inexplicit utterances co-occur with participant bodily movements. We have taken into account four factors, including three types of bodily movements, (a) pointing, (b) manipulating, and (c) gazing, and one type of contextual feature, (d) being near. These are shown in Fig.1. The robot can recognize what is being directed by the user without speech. In cases (2) and (3), the robot can predict the verb from a detected object, and confirm it to a user.

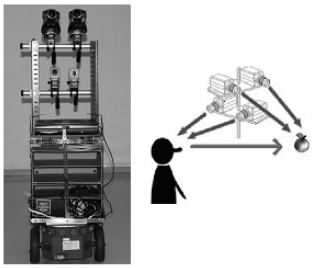

Fig.2: Robot understanding simplified utterances.

The robot is shown in Fig.2. This robot has two sets of stereo-cameras. A lower stereo-camera always watches the user, and detects the direction of the user's gaze and point. An upper stereo-camera detects an object in the gaze and pointing direction provided by the lower stereo-camera. Upon this detection, the robot moves a camera so that two optical axes cross over a half line of gaze and pointing direction, detecting an object by a zero-disparity filter [4]. We conducted an experiment with our robot system, and confirmed that it can understand simplified utterances utilizing visual information. Human beings, however, in interpreting and responding to requests, may use a range of visual information in addition to the four factors listed above. Thus, as a step in developing robots to be used in elderly care, we have been analyzing how elderly persons actually do requests to caretaking in ordinary interaction.

3 Ethnomethodological analysis

This section presents ethnomethodological analysis. Ethnomethodological analysis investigates

how people achieve mutual understanding in interaction [3].

(1) Interaction is organized in relation to sequences that typically have a high degree of temporal ordering. [5][6][7]

(2) Through talk and physical actions, people's intentions and behaviors become accountable to others. [8][9][10][11][12]

(3) For instance, people's physical actions (e.g., head movement, gaze, movement of upper torso, direction of their arm and pointing) indicate orientation towards an object. These actions are furthermore hierarchically organized. [13][14]

Physical action in particular indicates "projectability"; which allows others anticipate the person's desires. For example, when A says, "Please get that" while pointing or reaching toward a book, A typically first gazes towards the book. A's gaze and her verbal message indicates what A desires is the book, and not something else such as the remote control, pen, and the like. Ethnomethodological studies explore how people coordinate their actions smoothly by focusing on the notion of projectability through talk and physical actions. Kuzuoka has shown that the notion of "projectability" is important for the development of a remote robot [14] . In our analysis, we apply the notion "projectability" in interaction between caretakers and elderly persons in a senior care home [14]. In particular, we observe how participants (helper and elderly) respond to other's physical movements, and how they coordinate their actions.

While we observe how people make requests in a senior care home, we noticed in a previous study that there while are a few direct requests, for example ``bring that'', many of the forms of requests are not so direct. We find many instances in which seniors display they want the caretaker to do something for them through physical movements, and the caretaker responds by carrying out the request. That is, even when there may be no talk between a senior and caretaker, they manage to carry out actions. Thus, we think that the role of physical movements are pertinent to making requests and carrying them out.

4 Projectability in the Cooperative Work among the Senior and the Helper

The data we examine here concerns a caretaker assisting an elderly person to move a chair. The elderly is a man, who has no cognitive problems but cannot walk for a far distance, and uses elderly daycare service two to three times a week. In this scene, the elderly are taking part in something called a power rehabilitation program. The elderly man we focus on here has finished his program and is beginning to take a break.

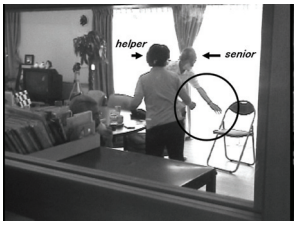

Fig. 3: Chair1

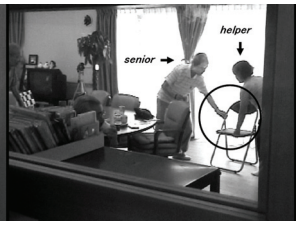

Fig.4: Chair2

In Fig. 3, the caretaker looks towards the man as he is extending his arm towards the pipe chair, and then approaches the senior. We focus on the approaching caretaker's arm as is extends towards the pipe chair. When the elderly man slowly walks towards the chair reaching his hands towards it, the caretaker also moves towards the chair. The senior then stops his action as the helper starts reaching towards the chair. It is clear that the elderly man stop his action because he anticipates the caretaker's action as a sign to assist him.

By stopping his own actions towards the chair, the elderly man indicates two things: (1) an understanding that the helper is attempting to assist him, and (2) an acceptance of this help. In this way, cooperative action between the caretaking and elder man is achieved. Let us now take a closer look at how the caretaker and senior coordinate their

actions. In interactional analysis, we use transcriptions of both the talk and verbal action to show how utterances and physical movements are coordinated in real time[11].

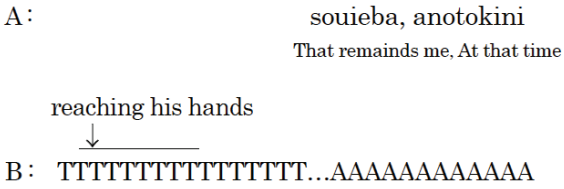

Fig. 5: transcript 1

Fig. 5 shows an example of transcript. In this transcript, upper cases (TTT, AAA) indicate gaze towards an object or persons (T for Tissue box and A for person). "..." indicates transfer of gaze, while underline indicates continuation of physical action. In this example, participant B is gazing towards T, and then transfers his gaze towards participant A. Here, B's gaze and A's utterance are expressed in the same timeline.

In this scene (Fig.7-Fig.8, transcript2 Fig.6), caretaker C helps elderly person B get a small package of tissues that is lying on the center of the table. Elderly person B is a man without any cognitive problems but cannot move the left half of his body freely, and uses daycare service two to three times a week. "T" stands for tissue box. In this scene, the participants are eating lunch.

At first B, an elderly man, extends his right arm towards a tissue box on the center of the table. Since the left side of B's body is partially disabled, it is difficult for him to fully reach out towards the box with his right hand, which is the only hand to support his body while sitting down. It is important to take note that B does not make a verbal request to the helper to assist him, even though he cannot reach the tissue box. Through his gaze towards the tissue box, direction of his upper torso, reach towards the tissue box and withholding of this action, B makes it possible for others to notice that he is attempting to achieve an intended action. Caretaker C, who is sitting next to B, not only notices B's actions but his proximity to B allows him to pull the tissue box towards B. B's withholding action indicates his understanding that C is attempting to help him. As mentioned previously, an elderly man displayed his understanding of a caretaker's movement towards the chair as an initiation of assistance by ceasing his own action. Similarly, in this example, B displays his understanding that the helper will assist by withdrawing his own action towards the tissue box.

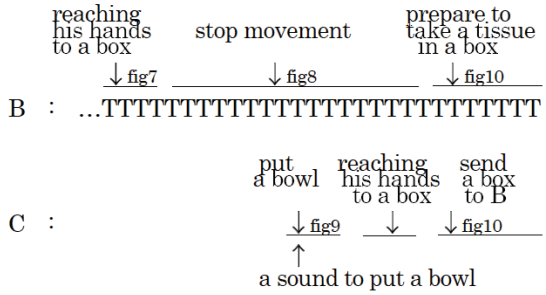

Fig.6: transcript2

Fig. 7: Box1

Fig. 8: Box2

Fig. 9. Box3

Fig. 10: Box4

Now let us examine how a person might understand the projected action of another person's physical action. Again, let us consider the above excerpt. Here we will first focus on the caretaker's action of putting down the rice bowl. When B reaches towards the tissue box, C puts down the rice bowl, reaches for the box and then pulls it towards B. At this time, B does not look towards C. This may mean that B does not notice C's action. However, when C puts down the rice bowl, it makes a noticeable sound. This sound can be an important resource for B to realize that C has momentarily stopped and is available to do an alternative action.

When B hears this sound he may recognize it as a sign that C is now available to help B with gaining access to the box of tissues. Therefore, by withholding the hand movement toward the tissue box immediately following hearing the noise of the rice bowl, B indicates that he is in need of assistance from C.

As shown above, we can see how assistance is achieved through two important resources: physical movement and sound. The above analysis has suggested that assistance in elderly care homes, are not achieved soley through verbal "request-response" sequences. These observations lead to the following two points. First, elderly persons and caretakers jointly coordinate assistance. Second, this coordinated assistance is made possible through the "projectability" of elderly and caretaker actions. Our analysis suggests the importance of actions coordinated by elderly persons and caregivers. For example, a caregiver can assist the elder by understanding what has been anticipated through his or her reach towards an object. In elderly care situations, caretakers coordinate their activities along with the elderly's embodied actions in order to carry out caretaking tasks smoothly.

5 Knowledge for Robot Development

In order to develop helper robots for elderly care, it is important to study the actual interactions that occur within senior care facilities. Now, we would like to clarify how our analyses can be used to develop helper robots. We have considered "projectability" in a study of a remote robot [14].

(1) A helper robot should be able to recognize the projected action of an elderly person when the person is attempting to accomplish some task but not being able to do so on his or her own.

(2) A helper robot should be able to display that it is attempting to help without actually helping the senior directly

In this way it is necessary for a helper robot to detect an action of a elderly person through projectaiblity, and to display the action to a senior by coordinating its action with this projectability. It becomes possible that cooperative tasks are achieved between a helper robot and a senior by such projectability on both sides.

We have to distinguish meaningful information from visual information to detect projectability through coordinated action of the elderly. In particular, it is important that a robot recognizes a "withholding movement" as projecting a certain action. In addition, we described the relations of physical actions and utterances in section 2. We have shown that it is necessary to recognize how assistance is achieved without speech in section 4. It is necessary that robots recognize such requests for assistance at high speed since such recognition is a chance to show projectability through coordinated actions from a robot to elderly.

It is necessary to consider what kinds of movements are more effective in showing projectability for the elderly. While observing human-human interaction, we also recognize that it is imperative to examine how people interact with robots, especially the movement of head, upper torso, arms as well as utterances, as this type of interaction is likely qualitatively different.

Furthermore, in the case of humans, we discovered that the elderly continues their own action when they recognize projected actions of the caretaker's movement as an initiation of assistance. Therefore, in the case of a robot, a similar situation might be anticipated. The robot should be able to recognize what is anticipated in the elderly person's action as well as display the robot's attempt to assist. We think it is necessary to develop a system able to detect the projectability of human actions and to display its intentions to humans through physical movements for natural communication between a helper robot and a human being.

In our analyses, we considered communication between humans. However in order to develop a robot able to respond to human requests similar to the way that humans respond to these requests,

more research is needed. In particular it would be beneficial to conduct experiments that compare situations of robot-human and human-human interaction in the context of caretaker-elderly requests and assistance. We plan to carry out such experiments, and analyze the interaction through ethnomethodological analysis. We will examine a robot that does not carry a elderly person's request immediately and observe the how the elderly person responds to such situations. The robot should be able to display projectability through coordinated actions by displaying what the next robot's action is and achieving its tasks cooperatively.

6 Conclusions

In an elderly care home, our analysis of interaction suggests that even when elderly persons do not express his/her request through speech, the caretakers ascertain their intentions through their bodily actions. In many cases, the elderly gains assistance without the use of any words. Elderly persons and caretakers achieve such action cooperatively. During such instances projectability of action is crucial. This projectability is made possible through bodily behavior and material resources in the environment. Based on our analysis, we plan to develop a helper robot that can provide assistance to elder persons in the future. We will develop a robot that responds to projectability in order to achieve cooperative actions with the elderly persons.

Acknowledgment This work was supported in part by the Ministry of Internal Affairs and Communications under the Strategic Information and Communications R&D Promotion Program and by the Ministry of Education, Culture, Sports, Science and Technology under the Grants-in-Aid for Science Research (KAKENHI 18049010, 16330095).

References

[1] K. Dautenhahn, H. Bond Alan, and L. Canamero, ``Socially Intelligent Agents : Creating Relationships with Computers and Robots'' Springer, 2002.

[2] Z.M. Hanafiah, C. Yamazaki, A. Nakamura, and Y. Kuno, “Human-robot Speech Interface Understanding Inexplicit Utterances Using Vision,” CHI Extended Abstracts, pp.1321-1324, 2004.

[3] K. Yamazaki, “An Introduction to the Practice of Ethnomethodology,” Yuhikaku, 2004. (in Japanese)

[4] D. Coombs, and C. Brown, “Real-time Binocular Smooth Pursuit” Int. J. Comput. Vis. vol.11, no.2, pp.147-164, 1993.

[5] H. Sacks, and G. Jefferson, “A Simplest Systematics for the Organization of Turn-taking for Conversation,” Language, 50, pp.696-735, 1974.

[6] E. Schegloff, G. Jefferson, and H. Sacks, “The Preference for Self-correction in the Organization of Repair in Conversation,” Language, 53, 2, pp.361-382, 1977.

[7] E. Schegloff, and H. Sacks, “Opening up Closings,” Semiotica, VIII, 4, pp.289-327, 1973.

[8] C. Goodwin, “Action and Embodiment within Situated Human Interaction,” Journal of Pragmatics 32, pp.1489-522, 2000.

[9] C. Goodwin, “Conversational Organization: Interaction between Speakers and Hearers,” New York: Academic Press, 1981.

[10] C. Goodwin, “Professional Vision,” American Anthropologist 96(3), pp.606-33, 1994.

[11] C. Goodwin, “Transparent Vision,” In Interaction and Grammar, ed, E. Ochs, E. Schegloff, and S. Thompson, Cambridge University Press, pp. 370-404, 1996.

[12] C. Heath, “Talk and Recipiency: Sequential Organization in Speech and Body Movement,” In: M. Atkinson and J. Heritage, ed: Structures of Social Action. Studies in Conversation Analysis, pp.247- 265, 1984.

[13] A. Kendon, “Conducting Interaction: Patterns of Behavior in Focused Encounters,” Cambridge: Cambridge University Press, 1990.

[14] H. Kuzuoka, J. Kosaka, K. Yamazaki A. Yamazaki, and Y. Suga, “Dual Ecologies of Robot as Communication Media: Thoughts on Coordinating Orientations and Projectability,” in Proc. of CHI2004, pp.183-190, 2004.