User Modeling and Reception Interaction by Service Robots

Rintaro Manabe
Graduate School of Science and Technology, Seikei University
Musashino-shi, Tokyo
Japan
dm156212@cc.seikei.ac.jp

Yutaka Takase
Dept. of Computer and Information Science, Seikei University
Musashino-shi, Tokyo
Japan
yutaka-takase@st.seikei.ac.jp

Yukiko I. Nakano
Dept. of Computer and Information Science, Seikei University
Musashino-shi, Tokyo
Japan
y.nakano@st.seikei.ac.jp

***Abstract*— There are many situations that social robots can serve, such as shopping malls, care facilities, and museums. As one of such situations, this paper focuses on a restaurant/café situation, and discusses the importance of recognizing the user's characteristics as a technical challenge in developing a social robot for the wild. First, we will discuss the user characteristics that the robots should recognize and explain how to recognize them. Then, we describe our prototype system in which such user modeling functionality is implemented.**

Keywords—Café situation; User modeling; Social robot

I. INTRODUCTION

There are many situations that social robots can serve, such as shopping malls, care facilities, and museums [2, 5, 6]. The required skills and behaviors of the robots are different depending on the situation and the task of the robots. But, more importantly, social robots should choose their behaviors according to the users with whom the robot interacts. This paper focuses on a restaurant/café situation, where robots should show the customers to their table, recommend the speciality, and take an order, and discusses the importance of user adaptive behaviors of social robots. For example, when a group of people are entering the café, they may be a family, business men, or university students. A clerk may choose different table, recommend different foods or services, and select different small talk topics depending on the group. Moreover, it is highly possible that the clerk may predict the leader of the group, and talk to him/her when asking to make a decision in the group.

Based on the discussion above, this paper discusses the importance of recognizing the user's characteristics as a technical challenge in developing a social robot for the wild; a café situation. In the next section, we will discuss the user characteristics that the robots should recognize and explain how to recognize them. Then, in section III, we will describe our prototype system in which such user modeling functionality is implemented. Section IV proposes discussion topics, and finally, in section V, we will discuss the future work.

II. SENSING USERS' CHARACTERISTICS

In a café situation, the following information may be useful to create a user model based on which the robot decides its behaviors;

- Gender: If the system can estimate the gender of the customers, this may be used in deciding what type of food the robot should recommend and what type of small talk topics should talk to the customer.
- Age: If the system can estimate the age of the customers, the system may change the food recommendation and small talk topics. The system may also infer whether the group is a family based on the age configuration of the group.
- Combination of gender and age information: If all the group members are middle-aged men, the system can assume that they are visiting the café for a business meeting.
- Proxemics: Recognizing the seating behavior contributes to inferring the intimacy and leadership. If customer A and B sit side by side and the other person C sits in front of them, we can assume that A and B are more intimate or have a good relationship with each other compared to with C [3]. Moreover, it is also assumed that the leader of the group may be at the head and other members may follow the leader.
- Eye gaze: By measuring eye gaze during the conversations at the table, it may be possible to estimate the intimacy and the dominance. Mutual gaze may be used as the measure of intimacy [1]. A person who more looks at others and is more looked at by others may be dominant in the conversation [4].

In order to recognize these user profiles and infer the social relationship between the customers, the system is required to employ various human sensing technologies. For example, computer vision technologies are necessary to recognize person, age, and gender. But, to estimate higher level information, such as intimacy and dominance, it is necessary to combine multimodal information and extract social signals in the conversations.

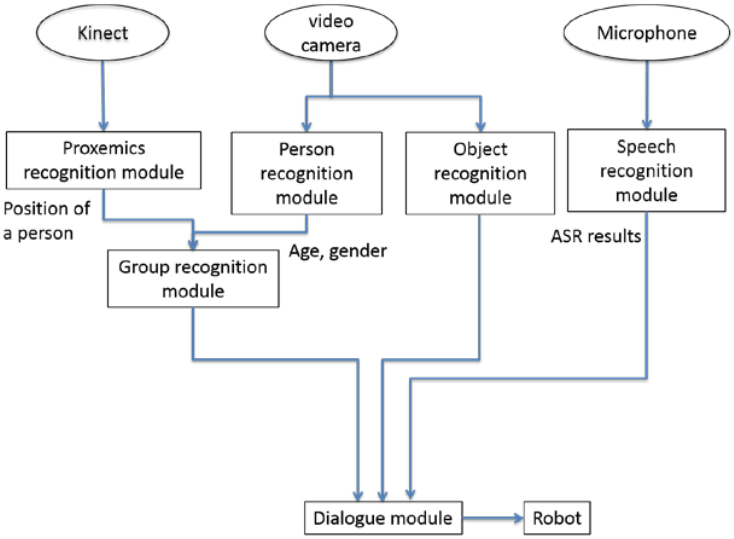

Fig. 1. Sysetm architecture

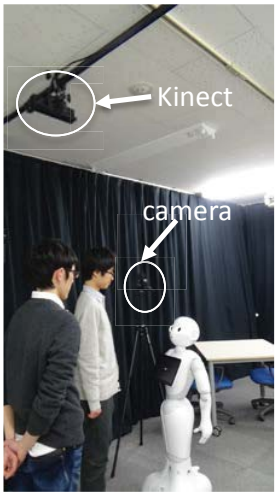

Fig. 2. Equipments and environemnt

III. Prototype System

A. System Architecure

Fig. 1 is the system architecture of our prototype system and our experimental environment is shown in Fig. 2. We use a humanoid robot, Softbank Pepper, to interact with the customers. We will explain each module as follows;

Proxemics recognition module: We set a Microsoft Kinect sensor on the ceiling to measure the depth information. When this module receives the depth information from a Kinect sensor, it detects the region of a person. By calculating the center coordinates, the system can estimate the position of the customer.

Person recognition module: This module recognizes a person and estimates the age and gender using visual data obtained from a video camera (embedded in the environment or installed in the robot). We use the Microsoft oxford computer vision API for age and gender recognition.

Group recognition module: This module combines the information from the Proxemics recognition module and the Person recognition module, and recognizes who is standing or sitting at where. For example, the module may recognize two middle aged men followed by a younger man.

Dialogue module: This module receives visual information from the Group recognition module and the speech recognition results from the Speech recognition module, and controls the conversation with the group of customers. In the current implementation, only greeting scenario (conversation at the entrance of the café) is implemented.

B. Expected Communication

We have implemented a conversation expected at the entrance of the café. First, when a group of customers are entering the café, the Proxemics recognition module detects the customers. This information is sent to the conversation module and the robot greets the customers: “Hello, how many in your party?” Then, the Group recognition module starts processing to recognize the group members. If one of the customer answers, the system refers the age and the gender of this customer, and chooses a topic for small talk, such as weather or business news.

IV. Discussion

In order to estimate user profiles and communication attitudes, human behavior sensing is indispensable. The sensors may be installed in the environment as we did in our experimental room. Otherwise, cameras and sensors installed in the robot are used. Thus, the following issues need to be discussed in order to build socially skillful service robots that can work.in the wild.

- Should the sensors be installed in the environment or in the robot?
- In case of being installed in the environment;
  - What are the useful devices to be installed in the environment?
  - How to build sensor rooms suitable for a café situation?
  - What is the ideal architecture that integrates sensor information and robot manipulation?
- In case of using a robot as a multi-sensor device;
  - How to measure both the customers to whom the robot interacts with and the environmental information?
  - Are the state-of-the-art robot technologies capable of measuring and processing all the information necessary to enable the robots to work in the wild?

V. Conclusion and Future Work

This paper focused on a café situation and discussed what functionalities and technologies are necessary to implement a socially skillful robot that can estimate user models and change the communication based on the user model. We are working on implementing the functionality for intimacy and dominance estimation by analyzing gaze and speech data during the conversation at the table. Our work is still preliminary and conceptual, but we believe that a café situation is a good practice to study HRI in natural interaction settings. We will investigate whether sensing users' personal and group characteristics may contribute to social interaction in such situations.

References

[1] Argyle, M. and Dean, J. Eye-Contact, Distance and A liation. Sociometry. 28(3): p. 289-304, 1965.

[2] Kanda, T., Shiomi, M., Miyashita, Z., Ishiguro, H., and Hagita, N. An Affective Guide Robot in a Shopping Mall. in Proceedings of HRI' 2009. p. pp.173-180, 2009.

[3] Knapp, M.L. and Hall, J.A. Nonverbal Communication in Human Interaction. 2010: Wadsworth, 2010.

[4] Nakano, Y.I. and Fukuhara, Y. Estimating Conversational Dominance in Multiparty Interaction. in Proceedings of 14th ACM International Conference on Multimodal Interaction (ICMI2012). p. 77-84, 2012.

[5] Sabelli, A.M., Kanda, T., and Hagita, N. A Conversational Robot in an Elderly Care Center: an Ethnographic Study. in Proceedings of HRI2011. 2011.

[6] Yamazaki, A., Yamazaki, K., Kuno, Y., Burdelski, M., Kawashima, M., and Kuzuoka, H. Precision Timing in Human-Robot Interaction: Coordination of Head Movement and Utterance. in Proceedings of CHI2008. p. 131-140, 2008.