Socially Distributed Perception

Marek P. Michalowski, Carl DiSalvo,
Didac Busquets, Laura M. Hiatt,
Nik A. Melchior, Reid Simmons
Carnegie Mellon University
Pittsburgh, PA 15213

Selma Sabanovic
Rensselaer Polytechnic Institute
Troy, NY 12180

ABSTRACT

This paper presents a robot search task (*social tag*) that uses social interaction, in the form of asking for help, as an integral component of task completion. We define *socially distributed perception* as a robot's ability to augment its limited sensory capacities through social interaction.

Categories and Subject Descriptors: I.2.9 [Artificial Intelligence]: Robotics–*Operator interfaces, Sensors*; K.4.0 [Computers and Society]: General

General Terms: Design, Human factors

Keywords: Human-robot interaction, social robotics, mixed initiative

1. INTRODUCTION

At the AAAI 2005 Robot Exhibition the robot GRACE (Graduate Robot Attending a ConferencE) [3] played a game of *social tag* in which the task was to locate and rendezvous with a team member who was wearing a pink hat [2]. In this game, our purpose was not to create an object localization task (such as a scavenger hunt); rather, we wanted to create a task that would require a "socially situated" [4] robot that could enlist the help of humans through frequent social interactions. We designed the game of social tag so that the robot's information about the whereabouts of the team member came primarily through social interactions with strangers in the environment.

Most machine perception systems are still severely limited compared to the sensory abilities of humans. Our project was motivated by the idea that robots built to operate in human environments (which are usually well-suited for human perception) would benefit from an ability to off-load sensory demands to human partners. It is up to a robot's designers to ensure that a robot is capable of *requesting* this help and *accepting* it when given. We introduce the term *socially distributed perception* to describe a robot's ability to augment its own perception through social interactions with people. Such a robot is perceptually embedded within a social as well as a physical environment and can treat as affordances both the social and the physical characteristics of that environment [1]. In order for the robot to take advantage of these social affordances, it must be able to engage in social interaction with people.

GRACE is an RWI B21 mobile robot. For this task, it was equipped with a laser scanner, a camera, an LCD monitor with an animated face, and an LCD touchscreen. The robot has two computers on board: one for controlling mobility, sensing with the laser, avoiding obstacles, and handling the touchscreen interface, and the other for vision, control of the face and voice, and general task control.

Copyright is held by the author/owner.
HRI'06, March 2–4, 2006, Salt Lake City, Utah, USA.
ACM 1-59593-294-1/06/0003.

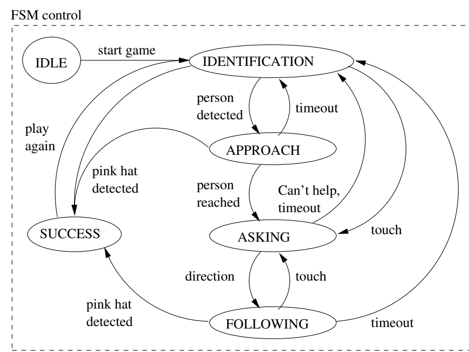

Figure 1: The finite state machine that comprises GRACE's control task for Social Tag.

2. SOCIAL TAG

Many tasks in robotics research use distinctively colored objects, such as pink hats, to simplify the machine perception problem. Despite this simplification, in large crowded rooms the task of finding a hat is still extremely difficult. GRACE relies on the assistance of humans with fully developed senses of sight and hearing and the ability to communicate. In this scenario, the pink hat was intended to be as much for the benefit of other people as for GRACE herself. Accordingly, while GRACE depends on her own sense of vision to achieve the goal (i.e. recognizing the pink hat), her primary mode of gathering information is asking people for help in an intuitive and socially acceptable manner. The task itself has five main phases: *identification* of approachable humans; *approach* toward a human; *asking* for directions; *following* those directions until a pink hat is found visually or more help is required; and demonstrating *success* when the hat has been found (Fig. 1).

3. OBSERVATIONS AND RESULTS

We evaluated GRACE's performance by observing and analyzing three aspects of interaction: the influence of the spatial and social nature of the environment on the human-robot interaction; the nature of social interactions between

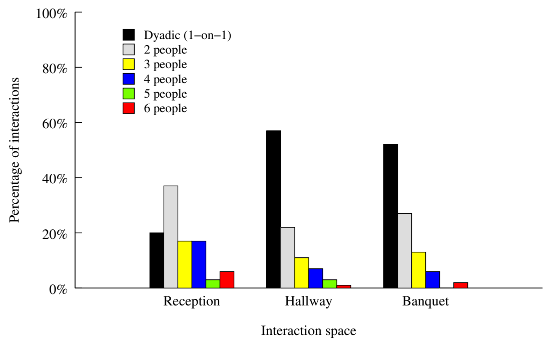

Figure 2: The size of groups with which GRACE interacted.

GRACE and conference participants and participants with each other; and the robot's success in using social interaction to obtain perceptual assistance.

GRACE operated for approximately 15 hours over the course of three days at the conference and played over 100 games. Video recordings and live observations of GRACE's performance were made during this time. Using Noldus Observer software, we coded and analyzed 3.6 hours of videotaped behaviors such as speech, spatial movement, gesture, and gaze as performed by GRACE and conference participants.

For analytical purposes, and in accordance with our emphasis on socially and physically situated interaction, we categorized the data according to the social spaces in which the interactions between GRACE and participants occurred. There are three categories corresponding to areas in the conference venue that varied in their spatial configuration and social use: the *reception*, a social event held in a large hall in which people were contained and crowded (31.25 minutes recorded); the *hallway*, a place through which people walked on their way to the various conference presentations and where they examined poster displays (104 min.); and the *banquet*, a social event during which the hallway was furnished with food tables for the occasion (82.4 min.). A comparison of interactions within the three categories shows the effect of both social and spatial factors on human-robot, as well as interpersonal, interaction. Observation and analysis of the video showed that there were salient differences in the way people were affected by and interacted with GRACE in these three spatial and social environments.

A total of 171 touchscreen interaction sequences were observed in the video. Contrary to expectation, 53% of interactions involved more than one person gathering around GRACE and participating in the task by either taking turns giving her directions, helping each other understand the task, or locating the person in the pink hat before pointing GRACE in the right direction. The incidences of dyadic and multiple-person interactions in the various socio-spatial locations (hallway, reception, and banquet) were quantitatively different (Fig. 2). While the banquet and hallway had very similar distributions of interactions, in the reception hall there were actually more interactions with GRACE involving two people (37%) than with one person (20%).

Next, we looked at the effect of GRACE's actions on how conference participants interacted with the robot. A lag sequential analysis (a technique which measures the number of times certain events are preceded or followed by other events) was performed to see how participants reacted to GRACE's engaging, disengaging, and "random" wandering movements within 5 seconds of their occurence. There were a number of interesting differences between the interaction spaces with respect to participants' responses to GRACE's actions. For example, the reception had the highest overall rate of participants' engagement with GRACE. Furthermore, when GRACE made an engaging movement (turning toward a participant), people were more likely to make an engaging movement than a disengaging movement in the reception and banquet. On the other hand, in the hallway, people were equally likely to engage or disengage in response to an engaging movement by GRACE. This is possibly due to the transitionary nature of the hallway; the other two situations, although they were in different spaces, had a similar social purpose. Another notable phenomenon was that in the reception, after GRACE made a *dis*engaging action, people were more likely to re-engage the robot, either through movement, gaze, or using the touchscreen, than in the other two spaces.

It is apparent that GRACE's design made her more effective in instigating interaction in certain environments than in others. This suggests that the social and spatial situation in which the robot will be placed during the interaction should be seriously considered when designing social robots.

Finally, we looked at whether the robot was able to use interaction effectively in order to seek and obtain human help in achieving its goals. We logged the robot's internal state for 7.2 hours of her operation. The robot initiated interactions 57% of the time. When the robot initiated a logged interaction, humans provided help 83% of the time. When humans initiated the interaction, they provided help 91% of the time.

4. CONCLUSION

We have introduced *social tag* as a way of developing and demonstrating *socially distributed perception*, or a robot's ability to augment its sensory capabilities through social interaction. Our observations suggest that the robot was able to successfully use social interaction to request and accept assistance from conference participants with the aim of finding the person in the pink hat, and that the specific social situation had a significant effect on the nature of this interaction.

5. ACKNOWLEDGMENTS

This project was partially supported by National Science Foundation ITR projects #IIS-032914, #IIS-0121426, and #SES-0522630.

6. REFERENCES

[1] J. J. Gibson. *The Ecological Approach to Visual Perception.* Lawrence Erlbaum Associates, NJ, 1979.

[2] M. Michalowski, D. Busquets, C. DiSalvo, L. Hiatt, N. Melchior, R. Simmons, and S. Sabanovic. Social tag: Finding the person with the pink hat (technical report). In *Proceedings of the National Conference on Artificial Intelligence (AAAI '05) Mobile Robot Competition Workshop*. AAAI Press, July 2005.

[3] R. Simmons et al. GRACE: An autonomous robot for the AAAI Robot Challenge. *AAAI Magazine*, 24(2):51–72, 2003.

[4] L. Suchman. *Plans and Situated Actions: The Problem of Human/Machine Communication.* Cambridge University Press, Cambridge, UK, 1988.