[PRE-PRINT]
Part of the workshop "Robots in public spaces. Towards multi-party, short-term, dynamic human-robot interaction" at ICSR 2013.

A museum guide robot: Dealing with multiple participants in the real-world

Karola Pitsch, Raphaela Gehle, Sebastian Wrede

Bielefeld University, CITEC & CoR-Lab

Abstract. Using video-recordings from a real-world field trial of a museum guide robot, we show how a robot's gaze influences the visitors' state of participation in group constellations. Then, we compare the robot's conduct to a human tour guide's gaze strategies. We argue that a robot system, to deal with real-world everyday situations, needs to be equipped with knowledge about interactional coordination, incremental processing and strategies for pro-actively shaping the users' conduct.

1 Introduction

When exploring robot systems in real-world situations, the scenario of a museum guide robot ([5] [3]) reveals a set of challenges involved in multi-party situations: The robot is faced with visitor groups of varying size who could join an encounter or become disengaged and walk away whenever they like to. Thus, the robot needs to (i) detect multiple visitors and their shifting states of participation and (ii) have at its disposal a set of strategies suitable for dealing with them and sustaining their engagement. Controlled laboratory studies suggest that a robot using gaze behavior has increased persuasive power [1] and that - in multi-party settings - a robot's gaze being directed at one participant leads to positive user ratings [4]. Investigating the interactional implications of a robot's gaze conduct in multi-party situations, [2] show that the robot's short glances (8% of the interaction time) to the user invite him to interpret his interactional role as 'bystander', and glances for more than 71% as 'addressee'. Yet, little is known about the users' interpretation of a robot's gaze conduct in multi-party real-world situations where users are mobile, the system has to deal with insecurity of perception and adjusts its gaze dynamically to the users' (changing) positions.

In this paper, we explore how the gaze behavior of a robotic museum guide - which adjusts its head orientation to the nearest visitor at the end of each utterance [3] - is interpreted by the visitors and how it influences their state of participation in a real-world museum field-trial. In a second step, we compare this HRI-situation to a human guide's gaze strategies in semi-circle formations.

2 Museum guide robot: Addressing the nearest visitor

While previous studies show the impact of a robot's gaze on the users' participation in controlled laboratory conditions [2], we are interested in exploring how users interpret the gaze conduct of a semi-autonomous robot system under real-world conditions in a multi-party situation. Here, additional factors come into play: the dynamics of visitors moving in space, and the insecurities of the system's perception of users on which its observation-production loops are based.

To investigate these aspects as a basis for future modeling, we set up a NAO robot (3+ version) to give explanations of exhibits in a museum. The robot was equipped with a person-tracking module, predefined dialog patterns (talk, gestures, head orientation), and a coordination module following a dual dynamics inspired arbitration scheme. The robot detected/tracked up to three visitors using a Vicon motion capture system (visitors wore hats with markers) providing information about their position in space and head orientation (to robot vs. elsewhere). The robot's gaze conduct was designed using interactionally relevant, yet simple gaze behavior: (i) During the utterances explaining a painting, the robot shifted its head between painting and visitors at relevant moments. (ii) At the end of each utterance, the system's person hypothesis was updated, and the robot adjusted its head orientation dynamically to face the nearest visitor. Given [2], this design (ii) was chosen as we expected visitors to react to a robot favoring *one* visitor. We conducted a field trial at the Bielefeld arts museum, in which the robot was set up in the corner of an exhibition room and gave explanations (2:30 min.) to uninformed visitors who passed by during their regular visit. The resulting video/Vicon corpus comprises of 260 runs with varying numbers of visitors, from which 51 are suitable for analysis of visitor participation.

In [3], we present two case analyses from this corpus and investigate how visitors interpret the robot's gaze conduct. Using micro-analysis inspired by multimodal Conversation Analysis, we reveal the ways in which the robot's shifting head orientation (gaze) engenders shifts in the visitors' state of participation.

Case (1) *From a 1:1 situation to engaging a group*: While the robot is giving its explanation to one visitor (V1), two further visitors (V2, V3) approach. Although they stay behind V1, the robot turns its head to face V2 (perceptual 'insecurity'). V1 reacts by following the robot's head orientation and takes a step back-and-sidewards adjusting the spatial arrangement to include V2/V3. Then, the robot's head shifts back to V1, remaining oriented in between V1 and V2. Now, both visitors respond to the robot's questions, which indicates that also V2 has become an active co-participant. Thus, through the robot's gaze shifts, the initial 1:1-situation emerges to a group constellation.

Case (2) *Disengaging an individual from a group*: When two visitors enter the room, the robot directs its head to the participant who arrives first (V1, Fig. 1). The second visitor (V2) positions himself a step behind V1 to his left and observes him interacting with the robot. At moments, the robot's head shifts slightly to the right (i.e. away from V2). V1 follows the robot's orientation and repositions himself in the line of sight; however V2 disengages and walks away.

These examples show that while the robot used a simple gaze conduct for dealing with multiple users, the situational dynamics of visitors moving in space and the insecurities of real-time perception lead to not-preprogrammed, dynamic shifts in the robot's head orientation. These head orientations are interpreted

Fig. 1. Case (2) HRI: Interactional micro dynamics of disengaging

by the visitors as meaningful gaze strategies, to which they attempt to sequentially adjust their own conduct on a fine-grained level. These observations extend the analyses in [2] in that they reveal aspects of the users' spatial rearrangement engendered by the robot's gaze conduct. They propose an interactional view on the coarse-grained evaluation of the robot's behavior suggested in [4] and the need for interactional coordination in HRI.

3 Human guide: Shifting gaze in semi-circle configuration

In parallel to investigating HRI, we also wanted to know how human tour guides organize their gaze when giving explanations to groups of visitors. We conducted a study with professional human tour guides in the same arts museum. Four guides were equipped with SMI Eyetracking Glasses during a 1 hour exhibition tour with a group of three to five visitors. Here, we focus on those parts of the tour during which the group is positioned in front of an exhibit (and guide) forming a semi-circle constellation, and the guide offers an explanation of the exhibit (see also [5]). The spatial configuration of the participants is comparable to the HRI situation, yet the sequences considered are part of a longer event.

Initial analysis reveals that in such explanatory situations, human tour guides use a 'shifting gaze'-strategy during which their gaze alternates between the exhibit and the participants. Fig. 2 shows the target positions of the guide's gaze with the connecting lines indicating the order of gaze shifts. The guide's gaze was considered as 'focusing' a person/object (vs. 'passing' during gaze shifts), when it fixed the position for at least 250 milliseconds. During an explanation, each visitor is addressed at least once, and the position in the middle of the semi-circle appears as privileged for being gazed at. The guide's gaze often moves from one person to the next, but also shifts occur to the second-next participant or the exhibits. Considering the relationship

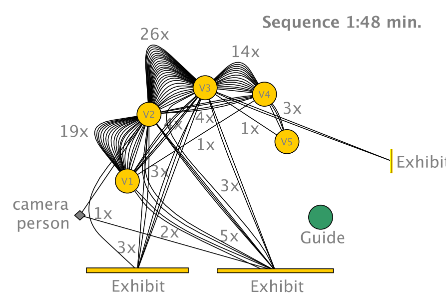

Fig. 2. Human guide: Gaze

between the guide's gaze and the wording of his talk, we wondered whether gaze shifts would co-occur with specific syntactical units/boundaries. However, our analysis so far does not provide evidence for such local talk-gaze patterns. Instead of being predictable, the guide's 'shifting gaze' strategy offers an important amount of variability.

Different from the HRI situation, visitors in our data generally don't (visibly) disengage from the guide's explanations. The rare cases, in which a visitor momentarily walks away from the group, occur at transitions when the group moves from one exhibit to the next one. On the one hand, these differences in the visitors' (dis-)engagement may be due to the different strategies deployed by the robot and the human guide. On the other hand, this also points to the different circumstances under which the human and the robotic guide act. This raises the question to which extent strategies from HHI can be used as inspiration for modeling robot conduct or when the interactional tasks might be too different.

4 Future work

The HRI study presented in this paper uses a basic robot setup aiming at an initial understanding of the practical tasks, interactional organisation and requirements for a robotic museum guide. Our future work will consist of linking the ideas of interactional coordination and incrementality with more performant (internal) observational capabilities of the robot system. We believe this to be an important basis for enabling robot systems to deal with everyday situations.

5 Acknowlegment

The authors acknowlegde the financial support from the Volkswagen Foundation (Dilthey Fellowship 'Interaction & Space', K. Pitsch), from CITEC (project 'Interactional Coordination and Incrementality in HRI') and from the CoR-Lab.

References

1. J. Ham, R. Bokhorst, and J. Cabibihan. The influence of gazing and gesture of a storytelling robot on its persuasive power. In *ICSR 2011*, 2011.
2. B. Mutlu, T. Shiwa, T. Kanda, H. Ishiguro, and N. Hagita. Footing in human-robot conversations: how robots might shape participant roles using gaze cues. In *HRI '09*, pages 61–68, New York, NY, USA, 2009.
3. K. Pitsch, R. Gehle, and S. Wrede. Addressing multiple participants: How a museum guide robot's gaze shapes the micro-dynamics of visitor participation. In *ICSR 2013*, pages 587–588, 2013.
4. Q. Xu, L. Li, and G. Wang. Designing engagement-aware agents for multiparty conversations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2233–2242, New York, NY, USA, 2013. ACM.
5. K. Yamazaki, A. Yamazaki, M. Okada, Y. Kuno, Y. Kobayashi, Y. Hoshi, K. Pitsch, P. Luff, D. vom Lehn, and C. Heath. Revealing gauguin: engaging visitors in robot guide's explanation in an art museum. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, New York, NY, USA, 2009. ACM.