Museum Guide Robot
Based on Sociological Interaction Analysis

Yoshinori Kuno, Kazuhisa Sadazuka, Michie Kawashima, Keiichi Yamazaki,
Akiko Yamazaki*, Hideaki Kuzuoka**

Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama, Saitama 338-8570, Japan
{kuno, sadazuka, mkawashi}@cv.ics.saitama-u.ac.jp, yamakei@post.saitama-u.ac.jp

ABSTRACT

We are currently working on a museum guide robot with an emphasis on "friendly" human-robot interaction displayed through nonverbal behaviors. In this paper, we focus on head gestures during explanations of exhibits. The outline of our research is as follows. We first examined human head gestures through an experimental, sociological approach. From this research, we have discovered how human guides coordinate their head movement along with their talk when explaining exhibits. Second, we developed a robot system based on these findings. Third, we evaluated human-robot interaction, again using an experimental, sociological approach, and then modified the robot based on the results. Our experimental results suggest that robot head turning may lead to heightened engagement of museum visitors with the robot. Based on our preliminary findings, we will describe a museum guide robot that first works autonomously and, if necessary, can turn into remote-control mode operated by a human to engage in more complex interaction with visitors.

Author Keywords

Nonverbal behavior, ethnomethodology, guide robot, computer vision.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): User Interfaces – *interaction styles*.

INTRODUCTION

We are currently working on a museum guide robot with emphasis on "friendly" human-robot interaction displayed through nonverbal behaviors. There were several museum guide robot projects, e.g., [7], which mainly focused on the autonomy of the robots and did not emphasize interaction with humans. In interacting with humans, it may be important for the purposes of presenting a "friendly" robot to utilize nonverbal behavior (e.g. head turning, arm

gestures). The research by Sidner et al. [10] has been conducted in a similar situation utilizing a guide robot, designed to explain some innovative items. Bennewitz et al. [1] have recently presented a humanoid guide robot that interacts with multiple persons. This robot can direct the attention of its listeners towards objects of interest through pointing gaze. Shiomi et al. [9] have done a longitudinal study on human-robot interaction at Science Museum. The aforementioned studies have not examined how gestures and other body movements can potentially be coordinated with talk in human-robot communication.

Recent research has attempted to develop gestures such as head movement in human-robot communication by studying human communication with a focus on gesture, head movement and eye gaze. In particular, Sidner et al. [10] developed a penguin robot and examined how users reacted towards the robot under two conditions in the context of the robot explaining an exhibit. In the first, the robot continuously gazed towards the user, while in the second the robot moved its head and arms occasionally during the explanation. Under the second condition, user attention apparently increased, as users responded to the robot's head movement and gaze direction by changing their own gaze and head directions. In another study, Breazeal [2] focused on emotion. Her work on nonverbal behavior of robots suggests the importance of nonverbal interaction between humans and robots. In summary, research on human-robot communication has suggested a link between head movements and gesture of robots to listener attention and response.

This paper describes an on-going collaborative research project by researchers in engineering and sociology. In a previous paper [4], we presented results on the behavior of human guides and visitors by using experimental sociological methods that included conversation analysis. We also showed preliminary experimental results regarding the effects of turning the robot's head with a simple prototype robot. In this paper, we report on further experiments using a humanoid robot. Then, we describe a museum guide robot that moves its head communicatively. The robot also observes the visitor's head direction to help determine whether or not s/he may have questions. If the

  * Future University-Hakodate, akikoy@fun.ac.jp
  ** University of Tsukuba, kuzuoka@iit.tsukuba.ac.jp

robot sees that s/he is gazing towards the robot for a continuous period of time, the robot presumes that s/he has a question, and turns itself into remote-control mode, which is then taken over by a human operator. The operator then responds to the visitor's questions.

PREVIOUS WORK

This section briefly describes our previous work. Details can be found in [4]. As mentioned in the introduction, we began by observing human behaviors. We performed two experiments where guides explained exhibits to visitors. We carried out a total of 11 sessions each from 15 to 30 minutes. We examined at what points in the talk the guides turned their heads towards the visitors. We extracted 136 cases, and made detailed transcripts for these cases. We then classified the cases as shown in Table 1. In the table, transition relevance places (TRPs) are points at which a speaker is likely to hand over the turn to a listener [8] such as upon finishing a sentential unit. We found that TRPs are the most frequent point at which guides turn their gaze towards the visitor.

Based upon the results of our experiment using human guides, we developed a prototype museum guide robot that moves its head while explaining exhibits. The robot has two pan-tilt-zoom cameras (EVI-D100, Sony). We attached a plastic head on the upper camera, and used the pan-tilt mechanism of the camera to move the head. We did not use the images of the upper camera. The robot uses the images of the lower camera to make eye contact and to observe the visitor's face [6].

Sixteen visitors participated in our experiments (14 females, 2 males, ages from 20 to 28, students and office clerks). In the experiment, when a visitor stood near an exhibit and made eye contact with the robot, the robot came close to the visitor and then began explaining the work. The robot explained the work in one of two modes: the *proposed* mode where it turns its head towards the visitor at predetermined points, and the *fixed* mode where it continuously gazes towards the exhibit (without turning its gaze towards the visitor). Half the participants underwent the fixed mode, then the proposed mode. The other half did so in reverse order. We did not tell participants the differences between the two trials.

	Number of occurrences
TRP (transition relevance place)	61
When saying keywords with emphasis	14
When saying unfamiliar words or citing figures	6
When using deictic words such as 'this'	26
With hand gestures	41
When the visitors asked questions	12

Table 1. Number of cases that the guides turned their heads in the two experiments (Total 136 times. Counted multiple if multiple conditions were satisfied).

As a quantitative evaluation, we examined when and how often participants turned their heads. In the proposed mode, the robot turned its head 7 times for each trial at predetermined points. Participants turned their heads 1.6 times on average for the fixed mode and 4.1 times for the proposed mode. The number of head turns of participants increased significantly in the proposed mode ($p < 0.01$, paired t-test).

EXPERIMENTS WITH HUMANOID ROBOT

Our previous experiment with the prototype robot shows that participants move their heads more frequently when the robot turns its head than when the robot does not turn its head. However, from this experiment we cannot say for certain whether participants' head movements are driven by the robot's head movements in appropriate timings or are simple reflections of the robot's motions. Sidner et al. [10] reports that robot head movements increase participants' engagement. Their result also comes from comparing two cases: one with head movements and another without.

In subsequent experiments, we used a humanoid robot Robovie-R ver. 2 [3], which can move its head more smoothly than the prototype robot. We conducted two types of experiments. The robot was programmed to explain two posters. We prepared two modes for the robot motion: *random* mode in which the robot turns its head at unsystematic points and the *proposed* mode in which the robot turns its head at interactionally significant points.

As shown in Table 1, the human guide often turns its head towards the visitor at points where his turn is about to come to completion (TRP). By turning his head, the guide may be able to check the visitor's understanding or non-understanding, as well as confirm that the visitor is listening. In addition to TRPs, the guide turns his head when saying key terms. These head movements again allow the guide to check the visitor's visible displays of understanding. By using these two interactionally significant points for robot head movements, we examined how visitors respond to such head movements. It should be noted that we did not implement head movements at deictic terms since the robot is programmed to point at the posters during this experiment and deictic terms are often accompanied with pointing in the experiments on human interactions. We focus only on head turns in this paper. Research on the concurrence of head and arm motions will be addressed in the future.

We had twelve participants, who experienced both random and proposed modes. Half of the participants underwent the random mode first, and the other half the proposed mode first. In addition, half of the participants were guided in poster 1 first and then poster 2. The other half underwent the opposite order. We videotaped both cases as shown in Figure 1.

We counted participants' movements that started within one second after the robot finished its head turn. We found two

typical movements: nodding and mutual gaze. When the robot turned towards the participant, the participant often made vertical head movements towards the robot (i.e. nodded). Participant nodding may display an attempt to show understanding of the explanation. Furthermore, when nodding occurs at TRPs, it may function as a "continuer", or request to keep going on with the explanation [8]. Second, participants often gazed towards either the robot or the poster by moving their vision in either direction as the robot turned towards the participant or the poster during the explanation. We call this kind of behavior mutual gaze since it seems to reveal participant attempts to engage along with the robot. In summary, the results suggest that participants' head nods and gaze may be displays of understandings and engagement with the robot and the poster.

We found an increase in participant nodding and mutual gaze in the proposed mode (i.e. robot turns its head at interactionally significant points). Figure 2 shows a significant difference especially in poster 1. In poster 2, the difference between these two modes is not as apparent, which may be because the content of poster 2 (gaze interface) is slightly more difficult to understand than that of the poster 1 (motion analysis of dances). Figure 3 shows a statistically significant difference in the ratio of nodding in the proposed mode (p< 0.01, paired t-test). This suggests the participants may be actively displaying their understanding of the robot's explanation when the robot turns its head towards the participants at interactionally significant places. Even though an increase of mutual gaze in the proposed mode is not apparent, it may be due to the fact that the participants are simply affected by the robot's head movements as observed in our previous experiment and that by Sidner et al.[10].

When nodding occurs together with mutual gaze, participants display an even greater heightened orientation towards the robot's explanation. Such co-occurrence increases significantly in the proposed mode (p<0.01, paired t-test) shown in Figure 4. In excerpt 1, the participant displays that she registers the explanation by looking at the robot and nodding toward the robot. This is the beginning part of the robot's explanation of poster 1.

Excerpt 1 (line 1: Robot's utterance in Japanese, line 2:
    Word by word translation, line 3: Equivalent English.
    R: Robot, **P**: Participant. Gaze direction and gesture:
    X= gaze toward the poster, P=gaze toward the participant,
    R=gaze toward the robot, N=nodding)

R: Buyou no keiko shisutemu ni tsuite setsumei shimasu.
    Dance training   system about   explain will
       'I will explain a training system for dancing.'
R:XXXXXXXXXXXXXXXXXX…PPPPPPPPPPPPPPPPPPP…XX
P: XXXXXXXXXXXXXXXXXXX…RRRRRRRRNNRR…XX

As this excerpt illustrates, slightly after the robot turns its head towards the participant near the end of the sentence at line 1, the participant starts looking at the robot. As the

robot is about to finish the sentence, the participant nods several times while looking at the robot. This suggests that the participant is engaging with the robot and may even be a display of understanding about what the robot has just said (through it also may be a display of simulated understanding due to the experimental setting). In addition, this turn is placed at the very beginning of the robot's explanation projecting to the listener further expansion of the robot's talk. By nodding at the end of this turn, the participant also shows her acceptance of the further expansion of the robot's explanation. Moreover, such reaction may indicate that the participant is treating the robot as if the robot is a ratified speaker, able to convey information the way a human guide would. This in turn shows that appropriate timing of the robot's head movement increases participant reactions that consider the robot as an interactive agent.
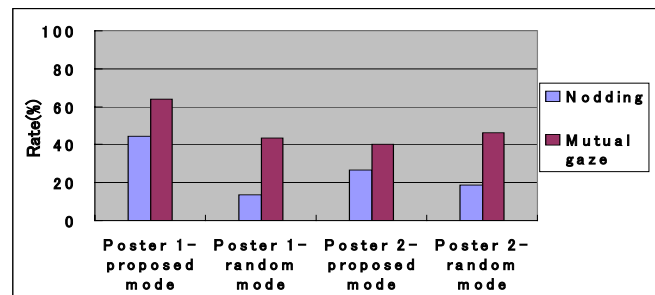

Figure 1. Experimental scene.

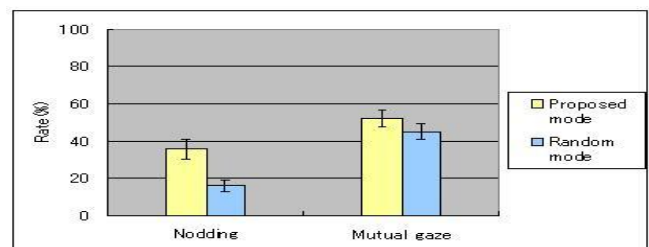**Figure 2. Ratio of participants' reactions toward each poster in the two modes.**

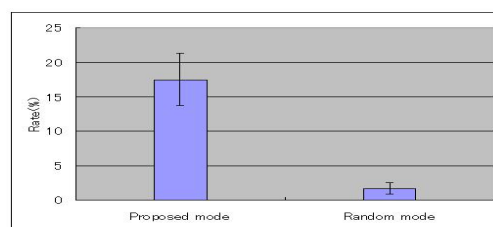**Figure 3. Ratio of participants' movements in each mode.**

Figure 4. Ratio of co-occurrence of nodding and mutual gaze.

DEVELOPMENT OF MUSEUM GUIDE ROBOT

Based on findings from the above experiments, we have developed a museum guide robot that turns its head towards visitors at particular points during the explanation (at similar points that human guides do). In addition, the robot continuously monitors the head direction of the visitor, and if it finds that the visitor keeps looking at the robot, the robot turns toward the visitor and asks, "Do you have any questions?" As shown in the bottom row of Table 1, human guides turn towards the visitor when the visitor asks them questions. During the questioning, the visitor often continues looking at the guides. Natural language understanding technology, however, is not yet enough to process and respond to questions from visitors. Because of this we have combined the autonomous operation mode described above and the remote-control operation mode, which allows for human response to visitors' questions [5].

Figure 5 shows a picture of the robot. The robot base is Robovie-R ver.2 used in our last experiment. We attached three cameras on its chest. If the robot locates a visitor from these three camera images, it turns its body in order to capture the visitor's face with the middle camera. We used the same face image processing method presented in [6]. Images from the three cameras are sent to the remote operator site. At the remote site, we obtain the head motion of the operator by using the three cameras attached on the displays. In the remote-control mode, the robot's head moves in the same manner as the operator's head.

As in the prototype robot, this robot makes eye contact with the visitor, and approaches him/her. The robot takes its body position as mentioned above, then starts explaining the exhibit. If the robot finds that the visitor keeps looking at the robot during the explanation, the robot turns its head toward the visitor and asks, "Do you have any questions?" The operation mode then changes to the remote-control mode. The operator watches the three displays. The head direction of the operator with respect to the three displays is sent to the robot to move its head in that direction. The robot shows which direction the operator is paying attention to through its head motion. Such head motion, which is similar to the autonomous mode, helps ensure smooth communication between human and robot [5], which is considered a substitute agent of the operator.

The combination of autonomous and remote-control modes allows a single human guide to operate multiple robots.

Figure 5. Guide robot with three cameras on its chest and the remote site.

CONCLUSION

Head movements play an important role in communication. We have analyzed human head gestures through an experimental, sociological method and have developed a prototype robot based on the results. We have also analyzed communication between this robot and humans. Based on these results, we have developed a museum guide robot where the autonomous mode changes automatically to the remote-control mode depending on the situation. While the original purpose of this project was to create a robot that is "friendly" (i.e. more human-like), we have not dealt with the issue of comprehension. That is, does the incorporation of head movements lead to increased visitor comprehension of contexts explained by the robot? This is a question for future research. We are also planning to observe this robot in the natural context of an exhibit.

REFERENCES

1. Bennewitz, M., Faber, F., Joho, D., Schreiber, M., and Behnke, S. Towards a humanoid museum guide robot that interacts with multiple persons. In *Proc. 5th IEEE-RAS Int. Conf. on Humanoid Robots* (2005), 418-423.

2. Breazeal, C. Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59 (2003), 119-155.

3. Intelligent Robotics and Communication Laboratories, http://www.irc.atr.jp/index.html.

4. Kuno, Y., Hiroyuki, S, Tsubota, T, Moriyama, S, Yamazaki, K., and Yamazaki, A. Museum guide robot with communicative head motion. In *Proc. RO-MAN 06*, (2006), 33-38.

5. Kuzuoka, H., Yamazaki, K., Yamazaki, A., Kosaka, J., Suga, Y., Heath, C. Dual ecologies of robot as communication media: thoughts on coordinating orientations and projectability. In *Proc. CHI 2004* (2004), 183-190.

6. Miyauchi, D., Nakamura, A., and Kuno, Y. Bidirec-tional eye contact for human-robot communication. *IEICE Trans. Inf. &. Syst. E88-D* (2005), 2509-2516.

7. Nourbakhsh, I., Kunz, C., and Willeke, T. The Mobot museum robot installations: a five year experiment. In *Proc. IROS 2003* (2003), 3636-3641.

8. Sacks, H., Schegloff, E., and Jefferson, G. A simplest systematics for the organization of turn-taking in conversation, *Language, 50* (1974), 696-735.

9. Shiomi, M., Kanda, T., Ishiguro, H., and Hagita, N. Interactive humanoid robots for a science museum. In *Proc. HRI 2006* (2006), 305-312.

10. Sidner, C.L., Lee, C., Kidd, C.D., and Rich, C. Explorations in engagement for humans and robots. *Artificial Intelligence, 166* (2005), 140-164.