Human-Robot Interaction in an Interactive Tour Guide Robot

Helena Wotring
Advisor: Professor Maja Mataric

I. Introduction

Human-robot interaction is a forefront of interactive robotics. As science once again is following science fiction, we strive to emulate the advanced robotics of our fiction. In one fictional story after another, intelligent, interactive robots have been an instrumental part of the stories, assisting humans in many ways. Interactive robotics strives to create such robots that can help people where there are not sufficient human aids to help. Interactive robots are not made to replace humans, only to aid them where time and funding is lacking.

II. Previous Work

Interactive robotics has been a topic of much research. In the city of Lancaster, a non-robotic tour guide system was suggested to give visitors to the city a more complete and interactive view (2). There have also been several instances where robots were developed to give tours in crowded museums. These robots not only sought out people for the tour, but also gave tours while dynamically navigating through the crowded museum (5). While different means of interface were used, such as humanoid features (4) or interaction via the internet (1), it was found in all papers that a robot that people could relate to was much more liked by the tourists than a robot they did not relate to, or seemed more machine like. In creating voices for an interactive robot, or webpage (3), voices that sounded more human were much more effective at being personable than clearly computer generated voices. In addition, when the personality matched that of the web viewer, the credibility of the voice, and thus the author of the content increased.

III. Materials and Methods

I composed my code on a Gentoo Linux machine with a GCC GNU v 3.4.4 compiler. I wrote my code using an iterative approach based upon the rational unified process. The scripts were pre-composed in separate files for easy editing and read in to the speech generator at runtime. Interaction was achieved through script cues from the robot which are responded to by the user, the tour group, via a simple button panel consisting of two buttons representing a yes and a no answer. To help monitor the progress of the tour group, at various parts of the tour, the user is asked to press any button to respond to a general cue.

In previous robots of this lab, a voice actor was used to generate the scripts for the robots. This is less desirable for two reasons: one, the same voice actor is needed each time the script is changed to re-record the script, and two, creating a simple driver to run a

humanoid head based on voice cues from a voice actor is nearly impossible. To get around this problem, I used Linux's Festival v 1.4.3. Using the phonemes it uses to create the spoken utterances, I was able to create a driver that formed those phonemes on the ESRA head in time with the speech.

IV. Results and Conclusions

While the entire robot was not completed by our co-op group by the end of my stay at the university, I was able to come to several conclusions about the humanness of a robotic agent. First, highly computerized TTS systems were found difficult to understand and were less desirable than a more human-like voice. Second, small details like a semi-random blinking pattern while the robot read the scripts helped lend a layer of reality to the robot.

V. Further Work

Instead of continuing using the basic Festival TTS engine, we will use AT&T's Natural Voices engine to create a more natural voice. As our own tests and others (3) have shown, a more human-like voice is perceived as being much more personable and has a more distinguishable personality. A more human-like voice is also easier to understand.

VI. Acknowledgements

I would like to thank my advisors Maja Mataric, and my mentors Emily Mower and Adriana Tapus. I would also like to thank David Feil-Seifer because even though he wasn't really my mentor, he always ended up helping me. I also need to thank Jennie Vongsoasup and Donovan Schafer, my group mates, without whom I would have not been able to work on such a fun project.

VI. References

1. *W. Burgard, A.B. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun.* Experiences with an Interactive Museum Tour-Guide Robot. Technical report CMU-CS-98-139, Computer Science Department, Carnegie Mellon University, 1998.
2. N Davies, K Mitchell, K Cheverst, and G Blair. Developing a Context Sensitive Tourist Guide. *Proceedings of The First Workshop on Human. Computer Interaction for Mobile Devices*, 1998.
3. C Nass and K M Lee. Does Computer-Synthesized Speech Manifest Personality? Experimental Tests of Recognition, Similarity-Attraction, and Consistency-Attraction. *Journal of Experimental Psychology: Applied.* 2001, Vol. 7, No 3, 171-181.
4. *S. Thrun, M. Bennewitz, W. Burgard, A.B. Cremers, F. Dellaert, D. Fox, D. Haehnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz.* MINERVA: A second generation mobile tour-guide robot. *Proc. of the IEEE International Conference on Robotics and Automation (ICRA'99)*, 1999.

5. S. Thrun, J. Schulte, and C. Rosenberg. Interaction with mobile robots in public places. *IEEE Intelligent Systems*, pages 7-11, July/August 2000.