

Behavioral Expression Design onto Manufactured Figures

Yoshihisa Ishihara

Shinshu University

4-17-1 Wakasato, Nagano, Japan

16w2003f@shinshu-u.ac.jp

Kazuki Kobayashi

Shinshu University

4-17-1 Wakasato, Nagano, Japan

kby@shinshu-u.ac.jp

Seiji Yamada

NII / SOKENDAI / Tokyo

Institute of Technology

2-1-2 Hitotsubashi, Chiyoda,

Tokyo, Japan

seiji@nii.ac.jp

ABSTRACT

Natural language user interfaces, such as Apple Siri and Google Voice Search have been embedded in consumer devices; however, speaking to objects can feel awkward. Use of these interfaces should feel natural, like speaking to a real listener. This paper proposes a method for manufactured objects such as anime figures to exhibit highly realistic behavioral expressions to improve speech interaction between a user and an object. Using a projection mapping technique, an anime figure provides back-channel feedback to a user by appearing to nod or shake its head.

Author Keywords

Human-agent interaction; speech interaction; projection mapping; human-like agent; interface agent

ACM Classification Keywords

H.5.1 Multimedia Information Systems: Artificial, augmented, and virtual realities; H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

INTRODUCTION

In recent years, natural language user interfaces, such as Apple Siri and Google Voice Search, have been embedded in consumer devices. These interfaces allow users to talk to devices to do everything from a phone call to Web searching. However, people sometimes hesitate to use the interfaces in public. Interestingly, most people speak to someone on their smartphones in public without hesitation, but feel awkward using voice interfaces in public. Although there is a privacy concern, the difference in the acceptance between voice calls and voice interfaces is attributed to the human-likeness or intelligence of such devices.

There have been several studies on increasing the human-likeness of machines and speech interaction between users and machines. Especially, in the context of using nonverbal information, Breazeal et al. [3] showed that human-robot

teamwork is enhanced using nonverbal information. Goetz et al. [4] reported that people's acceptance of a robot performing tasks that were more social in nature improved when the robot appeared more human-like. Additionally, Powers et al. [7] compared a screen agent and a robot in an interaction experiment and found that users took a more positive attitude toward the robot than the screen agent.

In this study, we attempted to improve an object's human-likeness by using nonverbal information to reduce the awkwardness users might feel when talking to the object. Nonverbal information in communication is an important factor because it affects consensus formation, speech content, and duration of the conversation [8]. Our approach is a way for already-existing objects to exhibit additional behavior in speech interaction rather than creating an entirely new object. We developed a system for an anime figure as a manufactured object to provide back-channel feedback to a user via a projection mapping technique [1] that gives the figure the appearance of nodding and shaking its head.

The proposed method has no shape limitations and provides natural expressions without mechanical vibration from motors because it does not use actuators. Although there are human-like software agents [2], such as MMDAgent [6], they are less realistic than embodied agents, such as robots [5].

BEHAVIOR-EXPRESSING AGENT

We developed a behavior-expressing agent using the projection mapping technique. It includes a small projector in a 310 × 240 × 110 mm case and the anime figure. The projector displays an animated image on the face of the figure.

Agent Behavior

The agent provides back-channel feedback during speech interaction by exhibiting nodding behavior as a positive reaction and head-shaking behavior as a negative reaction to a user. It also says “un” or “hai” in Japanese which mean “uh-huh” in English when it is nodding, and says “uunn” in Japanese which means “well...” in English when shaking its head. The responses of the agent were recordings of a female voice, with minor modifications in pitch. Figure 1 shows the nodding behavior using the projection mapping technique. The agent continuously monitors the voice volume of the user, and provides back-channel feedback

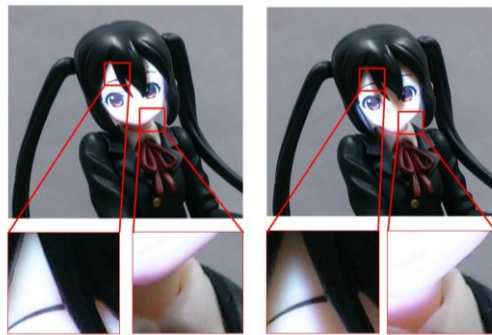
Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/author(s).

HAI '16, October 04-07, 2016, Biopolis, Singapore

ACM 978-1-4503-4508-8/16/10.

<http://dx.doi.org/10.1145/2974804.2980484>



(a) beginning of nodding (b) middle of nodding



(c) projected base image

Figure 1. Projected nodding behavior and base image

when the voice volume of the user exceeds a preset threshold after a 200-ms period of the volume being below the threshold.

Behavior Implementation

The nodding behavior of the agent is achieved by projecting a solid-white, face-shaped image onto the face of the anime figure (Figure 1c). It moves down 20 pixels in 150 ms and moves up 20 pixels in 150 ms in a single nod, as shown in Figure 1. The head-shaking behavior of the agent is achieved by projecting a face-shaped image that moves left and right 10 pixels from the center in 150 ms in one head-shaking, respectively. We empirically designed the behavior by projection mapping so that it clearly shows back-channel feedback behavior.

The image projection system for the agent works in a Web browser. The system includes a control window and an animation projection window, which is displayed on a projector screen that is connected to a personal computer. The system was developed in JavaScript and PHP, and the projection window controls the animation according to instructions in a text file shared between it and the control window. The system moves a face-shaped image that has been manually modified using a geometric transformation to fit the face of the anime figure.

We have a plan to perform experiments with participants to assess the participants' feelings toward the agent and the interaction using a questionnaire, and the number of utterances, the utterance duration, the wordless duration, and the number of fillers during the speech interaction between the agent and the participant.

CONCLUSION AND FUTURE WORK

A method for manufactured objects such as anime figures to exhibit highly realistic behavioral expressions to improve speech interaction between a user and an artifact was

proposed. We developed a listener agent based on an anime figure to provide back-channel feedback in the form of nodding and head-shaking behavior using a projection mapping technique.

The proposed type of projection mapping-based feedback could be applied to a voice interface to decrease the uncomfortable feeling when talking to objects. Our next steps are to develop various feedback expressions using the projection mapping technology and perform experiments in realistic situations.

ACKNOWLEDGEMENTS

This study was partially supported by JSPS KAKENHI "Cognitive Interaction Design" (No. JP26118005).

REFERENCES

1. Daniel G. Aliaga, Yu Hong Yeung, Alvin Law, Behzad Sajadi, and Aditi Majumder. 2012. Fast high-resolution appearance editing using superimposed projections. *ACM Transactions on Graphics* 31, 2 (2012), 13:1–13:13.
2. Timothy Bickmore and Justine Cassell. 2001. Relational agents: a model and implementation of building user trust. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI'01)*, 396–403.
3. Cynthia Breazeal, Cory D Kidd, Andrea Lockerd Thomaz, Guy Hoffman, and Matt Berlin. 2005. Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *Proceedings of the International Conference on Intelligent Robots and Systems (IROS '05)*, 708–713.
4. Jennifer Goetz, Sara Kiesler, and Aaron Powers. 2003. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *Proceedings of the International Workshop on Robot and Human Interactive Communication (ROMAN'03)*, 55–60.
5. Sara Kiesler, Aaron Powers, Susan R. Fussell, and Cristen Torrey. 2008. Anthropomorphic interactions with a robot and robot-like agent. *Social Cognition* 26, 2 (2008), 169–181.
6. Albert Lee, Keiichiro Oura, and Keiichi Tokuda. 2013. MMDagent – a fully open-source toolkit for voice interaction systems. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP'13)*, 8382–8385.
7. Aaron Powers, Sara Kiesler, Susan Fussell, and Cristen Torrey. 2007. Comparing a computer agent with a humanoid robot. In *Proceedings of the International Conference on Human-Robot Interaction (HRI'07)*, 145–152.
8. Bjorn Schuller, Dino Seppi, Anton Batliner, Andreas Maier, and Stefan Steidl. 2007. Towards more reality in the recognition of emotional speech. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP'07)* 4, 941–944.