Building Rapport with a 3D Conversational Agent

Whitney L. Cade, Andrew Olney, Patrick Hays, and Julia Lovel

Institute for Intelligent Systems, University of Memphis, 365 Innovation Dr.,
Memphis, TN, USA
`{wlcade,aolney,dphays,jcmnrslv}@memphis.edu`

Abstract. While embodied conversational agents improve a user's experience with a system, systems meant for repeated use may need agents that build a relationship with the user. Anita is a low-cost 3D agent capable of talking, displaying emotions, gesturing, and postural mimicry, all of which may increase the rapport between agent and user. Motion capture and pressure sensors were used to create an agent capable of realistic, responsive motions.

1 Introduction

Embodied conversational agents have been used in task-oriented computer applications to improve task performance and user experience. By talking, gesturing, and displaying affective facial expressions in a humanlike way, embodied agents may increase the rapport between the user and the agent [1]. Rapport improves conversational coordination between speakers and eventually allows disagreements to be an acceptable part of the conversation [2]. Therefore, a rapport-building agent could improve a user's experience with the system.

Rapport is often built nonverbally. For instance, humans naturally mimic the posture of their conversational partner in order to induce liking in the partner [3]. Gestures, another nonverbal mode of communication, create a common ground necessary for communication and understanding [4]. Facial expressions can not only signal the emotions being felt, but can also mark a lack of rapport, e.g. disgust [5].

We have created a low-cost embodied agent capable of speech, affective facial expression, gesture, and postural mirroring. While many agents have a subset of these rapport-building capabilities, we have united all of these aspects in one responsive embodied agent capable of realistic human motion. We believe this may increase attention, motivation, liking, and persistent use of the system over time.

2 Anita, the Reactive Conversational Agent

One of the major objectives in creating a multi-capable agent was to make a flexible 3D model on limited budget. Our conversational agent, Anita, was created using 3D Studio Max, a versatile 3D modeling software offered freely for educational purposes, and operates within the runtime environment Microsoft XNA. This gives the creator the freedom to control the agent in a world space so that any motion is possible, and the motion can be reactive to the user's input.

The agent is capable of gestures and expressions using morphing and skeletal animation. To create human-like movements, we built a motion capture system using four Sony Playstation Eye cameras and a low-cost motion capture program called iPi. The motion capture system was used to map a human actor's gestures and movements into a set of skeletal animations. We selected typical conversational gestures such as head nodding and an arm gesture that suggests that it is the user's turn to speak. These gestures can be dynamically or statically placed into the dialogue of the agent.

The agent also receives posture information from the user so that the agent can mirror the user's posture using Wii Fit Balance Boards to measure posterior body pressure [6]. Center of mass posture information is translated into dynamic, responsive agent leaning. Thus a user leaning heavily to the left would see this posture mirrored by the agent on screen, where the amount of lean is weighted by the posture information received from the Wii Fit.

Visible speech, emotion, and facial expressions are created using morph targets generated using Facegen, which creates high resolution heads. These morphs include visemes (phonemic mouth shapes), which give the agent the ability to lipsync with a text-to-speech voice, facial expressions such as sad, angry, happy, neutral, raised eyebrow, etc. Each morph target may be blended with the others to create composite expressions, such as happy surprise while speaking.

Anita represents both an effort to bring about a high quality agent using low cost equipment, and also the marriage of several avenues shown to create rapport between interlocutors (Demo video: http://www.youtube.com/watch?v=1gS3HLjgB7U).

References

1. Cassell, J., Gill, A.J., Tepper, P.A.: Coordination in Conversation and Rapport. In: Proceedings of the ACL Workshop on Embodied Language Processing, pp. 40–50. Omnipress, Madison (2007)
2. Tickle-Degnan, L., Rosenthal, R.: The Nature of Rapport and its Nonverbal Correlates. Psychol. Inq. 1(4), 285–293 (1990)
3. Chartrand, T.L., Bargh, J.A.: The Chameleon Effects: The Perception Behavior Link and Social Interaction. J. Pers. Soc. Psychol. 76(6), 893–910 (1999)
4. Clark, H., Brennan, S.: Grounding in Communication. In: Resnick, L.B., Levine, J.M., Teasley, J.S.D. (eds.) Perspectives on Socially Shared Cognition, pp. 127–149. American Psychological Association, Washington (1991)
5. Wang, N., Gratch, J.: Rapport and Facial Expression. In: Proceedings of International Conference on Affective Computing and Intelligent Interaction. IEEE Computer Society Press, Los Alamitos (2009)
6. Olney, A., D'Mello, S.K.: A DIY Pressure Sensitive Chair for Intelligent Tutoring Systems. In: Aleven, V., Kay, J., Mostow, J. (eds.) ITS 2010. LNCS, vol. 6095, pp. 456–456. Springer, Heidelberg (2010)