Embodied Interaction in Immersive Virtual Environments with Real Time Self-animated Avatars

Trevor J. Dodds
Max Planck Institute for Biological Cybernetics, Spemannstrasse 38, 72076 Tübingen, Germany
trevor.dodds@tuebingen.mpg.de

Betty J. Mohler
Max Planck Institute for Biological Cybernetics, Spemannstrasse 38, 72076 Tübingen, Germany
betty.mohler@tuebingen.mpg.de

Stephan de la Rosa
Max Planck Institute for Biological Cybernetics, Spemannstrasse 38, 72076 Tübingen, Germany
delarosa@kyb.tuebingen.mpg.de

Stephan Streuber
Max Planck Institute for Biological Cybernetics, Spemannstrasse 38, 72076 Tübingen, Germany
stephan.streuber@tuebingen.mpg.de

Heinrich H. Bülthoff
Max Planck Institute for Biological Cybernetics Spemannstrasse 38, 72076 Tübingen, Germany.
Department of Brain and Cognitive Engineering, Korea University
heinrich.buelthoff@tuebingen.mpg.de

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.
ACM 978-1-4503-0268-5/11/05.

Abstract
This paper outlines our recent research that is providing users with a 3D avatar representation, and in particular focuses on studies in which the avatar is self-animated in real time. We use full body motion tracking, so when participants move their hands and feet, these movements are mapped onto the avatar. In a recent study (Dodds et al., CASA 2010), we found that a self-animated avatar aided participants in a communication task in a head-mounted display immersive virtual environment (VE). From the perspective of communication, we discovered it was not only important for the person speaking to be self-animated, but also for the person listening to us. Further, we show the potential of immersive VEs for investigating embodied interaction, and highlight possibilities for future research.

Keywords
Animation, Avatars, Communication, Embodiment, Interaction

ACM Classification Keywords
H.5.1 Multimedia Information Systems–*Animations; Artificial, augmented, and virtual realities*.

General Terms

Experimentation, Human Factors

Introduction

Multiple users interact together using computer technology as part of their daily lives, e.g. using Twitter, Wikipedia, Second Life, World of Warcraft [7].

In the former examples, interaction and communication can be carried out using text based interfaces (e.g. discussing an article on Wikipedia). In the latter, people have an embodied representation (e.g. avatars talking to each other in a Massively Multiplayer Online Game). We ask the question, what is important about having an embodied representation in VE interaction. Specifically, an embodied representation in our work means a self-animated avatar: we use full body motion tracking to allow people to animate their own avatar in real time (Vicon optical tracking, and Xsens MVN suits). Our participants view their avatar using a head-mounted display (HMD). The view can be first-person (when the user holds up their hands in front of their eyes, they see their avatars hands) or third-person (over-the-shoulder perspective). See Figure 1.

This ability to control an avatar using motion tracking creates new and exciting opportunities for naturalistic interaction. While controlling an avatar's movements is not new (e.g. moving using a control pad, using the keyboard to make one's avatar 'wave', 'nod', 'shrug'), there is evidence to suggest a naturalistic interface (i.e.

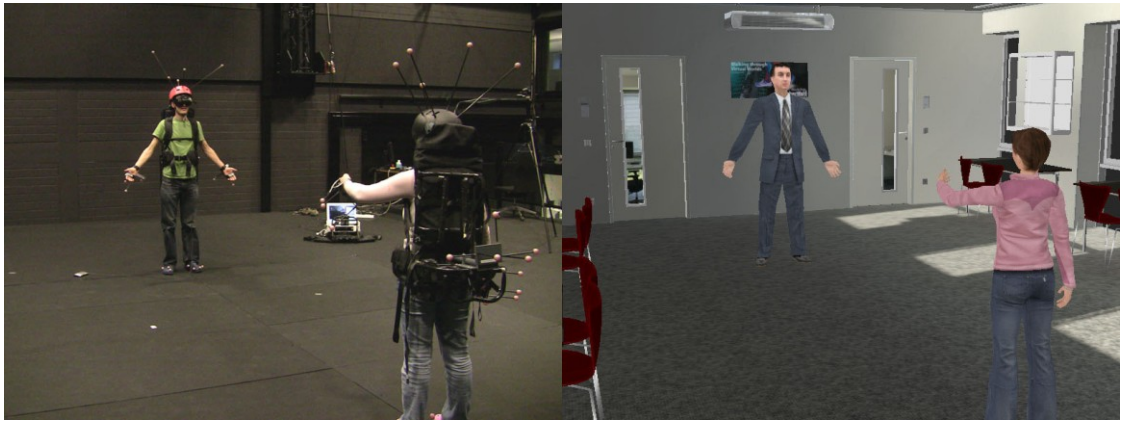

Figure 1: Two users (left) wearing rigid body objects on their hands, feet, helmet and backpack, tracked by Vicon cameras, and mapped onto their avatars (right).

using motion tracking) is actually increasing our sense of avatar ownership [5]. And recent advances in tracking technology, e.g. Microsoft's Kinect, make this type of interaction much more affordable.

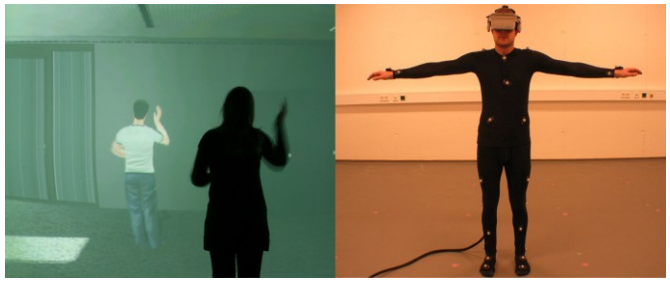

Figure 2: Examples of technology for self-animated avatars used by our research: Wearing an Xsens MVN suit underneath normal clothes (left); and Vicon body tracking wearing an NVis head-mounted display (right).

Our group's research has investigated the importance of a self-animated avatar from the perspectives of perception and interaction. In perception, giving people an avatar in a HMD environment helped alleviate the effects of distance compression that are well known to occur in VEs. Participants' distance judgements improved further when their avatar was self-animated. First or third-person perspectives did not have an impact on the magnitude of the effect [4].

In another study, [6], we investigated the effects of knowledge about one's avatar on task performance on three behavioral tasks: locomotion, object interaction and social interaction. We did not find effects of pre-exposure to a self-avatar on these tasks. We did however find effects of testing environment (VE / real world) and testing order (VE first, real world first) on participants behavior. We will further investigate if presence of the self-avatar during the task (rather than prior to task execution) will cause the performance on these three tasks to be closer to real world performance.

A study in [2] focused on the importance of embodiment from the perspective of multiuser interaction. We consider communication an essential subtask of any collaborative interaction, and therefore this experiment investigated how two users communicated in HMD virtual environments.

In HMD VEs, we can systematically manipulate different aspects of our environment, the functionality, and our appearance. In this study, we focused on the manipulation of the availability of body language and gestures in our self-animated avatar. We compared conditions with static avatars, with conditions where users had full-body motion tracking as a naturalistic interface for controlling a self-animated avatar.

The media richness theory claims that more interaction cues such as body language and gestures would improve communication [1]. In addition, work from psychology and psycholinguistics tells us that the gestures that naturally occur with speech carry additional meaning [3]. Finally, work in virtual reality shows evidence that we can have a sense of ownership over our avatars, and we would predict people to use gestures and body language in the virtual world in a similar way that participants would manipulate their own bodies in the real world (i.e. we would see participants and their self-avatars produce some of the gestures that naturally occur with speech).

To evaluate the importance of gestures participants played a communication game where they had to describe the meanings of words in rounds of three minutes, without saying the word itself. This gave us a measure of performance (mean number of words described per round).

The results show participants moved more and performed better in the communication game when both avatars were self-animated, compared to both static (we see the importance of nonverbal feedback, in addition to the usage of gestures by the speaker). This effect occurred in the third-person perspective, and we note the limitations of seeing one's own body in first-person perspective HMD environments (small field of view, meaning one's arms, legs and body are not visible when one is looking at someone in front of them).

Future work will investigate embodied interaction using large screen immersive projection technology, with two users wearing inertial motion tracking suits (Figure 2) and communicating over a network. In this setup we will investigate: What is important about seeing one's own body in the real world, with a naturalistic field of view, while communicating and interacting in VEs? Overall, immersive VE technology, where you can manipulate the visual body of the user in various ways, identity, movement, selective behavior (gaze, gestures), is a powerful tool for investigating embodied interaction between two people.

Acknowledgements

This work was supported by a post doctoral grant from the Max Planck Society, and by the World Class University program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (R31-10008). We thank Ivelina Alexandrova, Tobias Meilinger and Paulina Teneva.

References

[1] Daft, R.L. and Lengel, R.H. *Organizational Information Requirements, Media Richness and Structural Design*. Management Science, 32, 5 (1986), 554-571.

[2] Dodds, T.J. and Mohler, B.J. and Bülthoff, H.H. *A Communication Task in HMD Virtual Environments: Speaker and Listener Movement Improves Communication*. CASA: Proceedings of the 23rd Annual Conference on Computer Animation and Social Agents (2010).

[3] McNeill, D. *Gesture & Thought*. The University of Chicago Press (2007).

[4] Mohler, B.J., Creem-Regehr, S.H., Thompson, W.B. and Bülthoff, H.H. *The Effect of Viewing a Self-Avatar on Distance Judgments in an HMD-Based Virtual Environment.* Presence: Teleoperators and Virtual Environments, 19, 3 (2010), 230-242.

[5] Sanchez-Vives, M.V., Spanlang, B., Frisoli, A., Bergamasco, M. and Slater M. *Virtual Hand Illusion Induced by Visuomotor Correlations*. PLoS ONE, 5, 4 (2010), doi:10.1371/journal.pone.0010381.

[6] Streuber, S., de la Rosa, S., Trutoiu, L.-C.C., Bülthoff, H.H. and Mohler, B.J. *Does Brief Exposure to a Self-avatar Affect Common Human Behaviors in Immersive Virtual Environments?* Eurographics: Proceedings of the 30th Annual Conference of the European Association for Computer Graphics (2009).

[7] Yee, N. *The Demographics, Motivations, and Derived Experiences of Users of Massively Multi-User Online Graphical Environments.* Presence: Teleoperators and Virtual Environments, 15, 3 (2006), 309-329.