The Study of Interpersonal Communication Using Virtual Environments and Digital Animation: Approaches and Methodologies

Daniel Roth[1,2]

[1] University of Cologne, Germany

[2] University of Würzburg, Germany

Flashlight Paper

Presentation on the 66th Annual Conference of the International Communication Association (ICA), June 9-13 2016, Fukuoka, Japan.

Citation:

Roth, D. (2016). The Study of Interpersonal Communication Using Virtual Environments and Digital Animation: Approaches and Methodologies. *Presentation at the 66th Annual Conference of the International Communication Association (ICA), June 9-13 2016, Fukuoka, Japan*.

Abstract

Virtual technologies inherit great potential as methodology to study interpersonal communication. Current Virtual Reality (VR) technology is becoming increasingly available and accessible for research fields not limited to computer science. This flashlight will explain technological approaches and their use for the study of communication phenomena, presenting perspectives and potentials for communication research. It emphasizes technological requirements and workflows to take advantage out of the many possibilities VR and Virtual Environments (VE) offer for objective measurements that can be applied to interpersonal communication research. The present paper therefore focuses on approaches to motion tracking, physical appearances of avatars as stimuli, and exemplary objective measurements. Results of own studies are presented in order to give examples for technological adoptions to answer related research questions.

Keywords: Virtual Environments, Virtual Reality, Interpersonal Communication, Empirical Methods, Digital Animation

The Study of Interpersonal Communication using Digital Animation and Virtual Reality: Approaches and Methodologies

HIGHLIGHTS

Shared Virtual Environments (SVEs) inherit great potential as methodology to study interpersonal communication. Current Virtual Reality (VR) and Tracking Technology is available and accessible for research fields not limited to computer science. This flashlight will compare current approaches and their use for the study of communication phenomena, presenting perspectives and potentials as well as exemplary studies for communication research.

Introduction

Driven by the gaming industry, technological developments and market behavior have led to an increased attention to VR, Mixed Reality (MR) and Augmented Reality (AR).

“We’re making a long-term bet that immersive, virtual and augmented reality will become a part of people’s daily life.” (Marc Zuckerberg)

These developments led to the creation of low-cost solutions for head-mounted displays (e.g. Oculus Rift) and motion tracking sensors (e.g. Kinect) and accessories as well as the accessibility of development environments and game engines (e.g. Unity 3D) that may be used in future communication technology and are able to create interactive virtual worlds with decreasing effort. For research, this results in an increased methodological accessibility to study social interaction, communication and gaming not only from a reality- and perceptional view, but also to perform embodiment experiments that enable anthropomorphism and body ownership illusions (Lugrin, Latt, & Latoschik, 2015) relevant to induce a perceptional real-world similarity in mediated communication.

Studies in computer-mediated communication (CMC) often topic the social perception and processing of information through media channels. Amongst others, the social information processing theory by Walther (for a current review see (Walther, Van Der Heide, Ramirez Jr, Burgoon, & Peña, 2015) follows a functional approach to communication, concluding that any communicative message can be conveyed by a combination of alternative cues. Today's technology for virtuality can invert this rationale to allow for the experimental control and modification of physical appearances of the interactors (see Figure 1), available behavioral channels, and their transformation (Bailenson, Beall, Loomis, Blascovich, & Turk, 2005), overcoming traditional methodological problems (Blascovich et al., 2002). Besides "online" asynchronous or synchronous real-time communication and interaction with virtual environments, "offline" perceptual studies using collected behavioral data represent a powerful tool for psychological research and therapy and necessary cross-validation from a third-person perspective (Bohil, Alicea, & Biocca, 2011; Georgescu, Kuzmanovic, Roth, Bente, & Vogeley, 2014).

Virtual Environments

Virtual Environments (VEs) and VR technology have been defined as "a computer-generated display that allows or compels the user to have a feeling of being present in an environment other than the one they are actually in and to interact with that environment" (Schroeder, 1996, p. 2). It is to note that this definition explicitly excludes the narrowed "replication" of reality, which is identified to be impossible as to limits imposed by the laws of physics (Slater, 2014). VEs as a broad term is often used in the context of 2D Desktop environments and multiplayer games as well as social virtual communities giving the user possibilities to "interact" with the environment, no matter in which form. Immersive (embodied) VEs, often analogue to the term VR emphasize the sensorimotor coupling of head and body motion and the representative visual feedback for the user (e.g. “I move the mouse and the image changes” – VE, in contrast to “I turn my head and the image changes” - IVE).

Technological foundations

Working principles of VE and VR systems are in line with classical theoretical views on communication such as the approach by Shannon (2001). The information source in the case of VEs and VR is represented by the human user. The user’s actions are detected by sensors (e.g. gamepad, motion tracking, microphone). Messages from these sensors are discrete or continuously transmitted to the VR module (engine), which triggers calculations of , physics, interdependencies between the user and the environment as well as the final rendering of the current user’s viewport accordingly to drive the actuator components (e.g. display, headphones/speakers, haptic feedback) as depicted in Figure 3. Current systems often allow the user to interact via multiple modalities, which can therefore be described as (real-time) multimodal interaction.

A wide range of components can be selected to build an “embodiment lab” (Spanlang et al., 2014). Building a lab is seldom an “out of the box” experience, as many protocols exist and support between two interfacing components is required. Therefore researchers may adapt requirements analysis from engineering to find a fitting solution with respect to budget and cross-compatibility. While there are various ways for visual feedback depending on the task (head mounted displays, 3D screens/projections), motion input (motion tracking/capturing) is crucial for its reproduction or immersive VE/VR embodiment experiences.

Motion Tracking

Motion data that is typically mapped to avatars in real-time or post-processing can be distinguished along three major tracking fields: body motion, facial expression, and gaze.

Whereas methods for gaze-tracking are well established, measurements of body movement and facial expression are continuously advancing. Tracking of body movement can be achieved using i) active marker systems, ii) passive marker systems, iii) marker free optical (infrared pattern) based systems, iv) time of flight based systems, v) inertial sensor based systems. One crucial point is the overall system latency (end-to-end latency), as high latency may increase simulator sickness (Spanlang et al., 2014). Most of these systems provide access to a data stream that can be interfaced with proprietary recording software or a real-time based (game) engines to animate the avatar's body movement, gaze, and facial expression. The tracking of facial expressions recently advanced with the development of low cost methods using unobtrusive RGB-Depth sensors. Algorithms today allow for more robust feature tracking and expression estimation by combining texture and geometry registration. Some sensors and software provide an additional support to include gaze tracking.

Including the users' gaze data into avatar based communication systems allows for i) the transmission of gaze behavior via the users avatar and ii) recording of gaze behavior for later analysis of focus of attention, both leading to different requirements. Whereas an exact analysis requires precise measurement, the real-time transmission and mapping of gaze behavior requires low latency and the fasted possible approximation to quickly reproduce human behavior.

Physical Appearances and Behavior of Avatars

There is evidence that perception of avatars is similar to the perception of humans (Bente, Krämer, Petersen, & de Ruiter, 2001; de Borst & de Gelder, 2015; McDonnell, Jörg, McHugh, Newell, & O'Sullivan, 2008). Virtual characters that are used as avatars usually consist of a so called mesh of polygons that is rigged (constrained) simplified adaptation of the human skeleton with a number of joints, that follow either direct data input or are driven by a kinematic model. Similarly to the body, the face mesh is rigged to a skeleton model of the head. In addition to classical facial rigging today, so called blend shapes (morphologic deformations), represent a widely used option for facial animation (Orvalho, Bastos, Parke, Oliveira, & Alvarez, 2012). The avatar's visual appearance is displayed by textures that are applied to its mesh. Considering this structure (see Figure 2), a variety of modification possibilities arise:

i) Modifying the appearance of the avatar to study group effects such as stereotypes and racial biases (Peck, Seinfeld, Aglioti, & Slater, 2013)

ii) Transforming or dampening behaviors, decoupling the human behaviors from avatar behaviors (Bailenson et al., 2005; Boker et al., 2009)

iii) Activation, deactivation of cues and behavioral channels.

Whereas the latter point may comprise the highest potential to study interpersonal communication (and CMC), it also represents the biggest challenge to current VR systems as the multimodal coupling of body motion, facial behavior and gaze with the parallel use of head mounted displays still requires expansive consumer adaptations or custom beta-stage solutions.

Stimulus Creation for Cross Validating Perception Studies

Analog to the mapping process in online VE's/VR, researchers may use motion data of recorded single subjects or dyadic interactions to reproduce stimuli that can be controlled to a large degree using software for (professional) digital animation. Typical processes in animating behavioral data include *data management*, *data cleaning* (filling data gaps, correcting marker swaps), *filtering* (correct jitter artifacts), *actor/character mapping* as well as *movement modification* for the final *rendering* process. A broad range of virtual characters are available, reaching from low cost self-prepared characters to commercial high resolution characters with extensive skeleton construction and high definition textures. For the researcher it is important to understand the character's features and quality for the selection process. For example, it is important to know whether predefined blend shapes or skeletal movements for facial expression are included, or need to be added manually, which can lead to a time consuming process.

Once the material is prepared and the character is mapped accordingly, the final images for video stimuli are rendered. As with the creative direction in cinematography, this process allows modifications in lighting, camera angle and field of view and therefore gives the researcher control over the stimulus appearance.

Objective Measures of Communication Phenomena and Interpersonal Coupling

There are a number of objective measures to study communication using digital animation and VEs/VR. First, *eye-tracking* can be used when presenting dynamic stimuli in order to investigate focus of attention in perceiving social signals. Second, *video based as well as motion data* based mathematical post experimental analysis can result in estimating correlations in behaviors and social signals such as head nods that can be linked to interpersonal coupling or rapport and cross validated by perceptual studies with avatar animated or video material. For example legged cross correlations can be used to identify aspects of mimicry, synchrony and rapport using parameters such as motion energy, movement complexity, and head or chest angular deviations. Third, communication outcomes and performance times can give insights into the impact of absence/presence of communicative channels and change of physical appearance in mediated vs. real encounters.

Exemplary studies and results

Study 1: Affective displays of body motions were recorded using a motion capture system. Out of the recorded set, we selected five motions per group that were rated as happy, angry or neutral in a pretest. Onto these motions we mapped an avatar and modified the facial expression to a happy, angry or neutral, resulting in a total of 45 congruent/incongruent stimuli viewed by observers. Both, subjective rating and eye-tracking data suggests that in this study, facial expressions were the dominant channel in perceiving affect (in prep.).

Study 2: Dyadic interactions were recorded from two cultural groups (Culture 1, Culture 2) in female-female, male-male settings. We used neutral avatars and in twenty one minute slices (10 x Culture 1, 10 x Culture 2, five male and five female takes each) rapport was rated by observers using a continuous response measurement. Results indicate that in both cultures, females show higher rapport (in prep.).

Study 3: To study VR mediated communication, subjects were asked to do a performance task (ball game) as well as two permuted negotiation roleplays in both, real-life face to face interaction as well as in immersive embodied VR in which only the cues that could be tracked (body movement) were presented, whereas avatars did not have any face or eyes and were non-human mannequins. Results show significant difference in the performance task, which we interpreted was mostly due to a latency breaking sensorimotor coupling. However, no significant difference was found in the number of consensuses, time to consensus or interactions with the environment. Questionnaire analyses indicate a perceptual switch of attention to bodily cues in the VR scenario (in prep.).

Conclusion

The use of VE/VR and digital animation which is becoming more accessible allows for new methodological perspectives and control to the study of interpersonal communication and future CMC. The article identified common approaches, potential pitfalls and limitations that researchers should be aware of. The use of technology for virtuality requires both, financial resources, and interdisciplinary human resources. Motion Data and VR can give objective insights in research areas such as nonverbal communication, CMC, social interaction and perception. Exemplary study results for VR mediated communication are in line with traditional theory of CMC (e.g Walther et al., 2015).

References

Bailenson, J. N., Beall, A. C., Loomis, J., Blascovich, J., & Turk, M. (2005). Transformed social interaction, augmented gaze, and social influence in immersive virtual environments. *Human Communication Research, 31*(4), 511.

Bente, G., Krämer, N. C., Petersen, A., & de Ruiter, J. P. (2001). Computer animated movement and person perception: Methodological advances in nonverbal behavior research. *Journal of Nonverbal Behavior*, *25*(3), 151–166.

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, *13*(2), 103–124.

Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews Neuroscience*. http://doi.org/10.1038/nrn3122

Boker, S. M., Cohn, J. F., Theobald, B.-J., Matthews, I., Brick, T. R., & Spies, J. R. (2009). Effects of damping head movement and facial expression in dyadic conversation using real–time facial expression tracking and synthesized avatars. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3485–3495.

de Borst, A. W., & de Gelder, B. (2015). Is it the real deal? Perception of virtual characters versus humans: an affective cognitive neuroscience perspective. *Frontiers in Psychology*, 6: 576.

Georgescu, A. L., Kuzmanovic, B., Roth, D., Bente, G., & Vogeley, K. (2014). The Use of Virtual Characters to Assess and Train Non-Verbal Communication in High-Functioning Autism. *Frontiers in Human Neuroscience, 8*. http://doi.org/10.3389/fnhum.2014.00807

Lugrin, J.-L., Latt, J., & Latoschik, M. E. (2015). Anthropomorphism and Illusion of Virtual Body Ownership. In *International Conference on Artificial Reality and Telexistence Eurographics Symposium on Virtual Environments*.

McDonnell, R., Jörg, S., McHugh, J., Newell, F., & O'Sullivan, C. (2008). Evaluating the emotional content of human motions on real and virtual characters. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization* (pp. 67–74). ACM.

Orvalho, V., Bastos, P., Parke, F., Oliveira, B., & Alvarez, X. (2012). A facial rigging survey. In *in Proc. of the 33rd Annual Conference of the European Association for Computer Graphics-Eurographics* (pp. 10–32).

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779–787.

Schroeder, R. (1996). *Possible worlds: the social dynamic of virtual reality technology*. Westview Press, Inc.

Slater, M. (2014). Grand Challenges in Virtual Environments. *Frontiers in Robotics and AI*, *1*(3). http://doi.org/10.3389/frobt.2014.00003

Spanlang, B., Normand, J.-M., Borland, D., Kilteni, K., Giannopoulos, E., Pomés, A., … Slater, M. (2014). How to Build an Embodiment Lab: Achieving Body Representation Illusions in Virtual Reality. *Frontiers in Robotics and AI*, 1. http://doi.org/10.3389/frobt.2014.00009

Walther, J. B., Van Der Heide, B., Ramirez Jr, A., Burgoon, J. K., & Peña, J. (2015). Interpersonal and Hyperpersonal Dimensions of Computer-Mediated Communication. In S. S. Sundar, *Handbook of the Psychology of Communication Technology*. John Wiley & Sons, Inc.

Figures

Figure 1. *Examples for possible control options using avatars in perceptional studies and virtual environments. From top to bottom: clothes and appearance, body shape and facial features, gender and culture, presence/absence of nonverbal cues (face).*

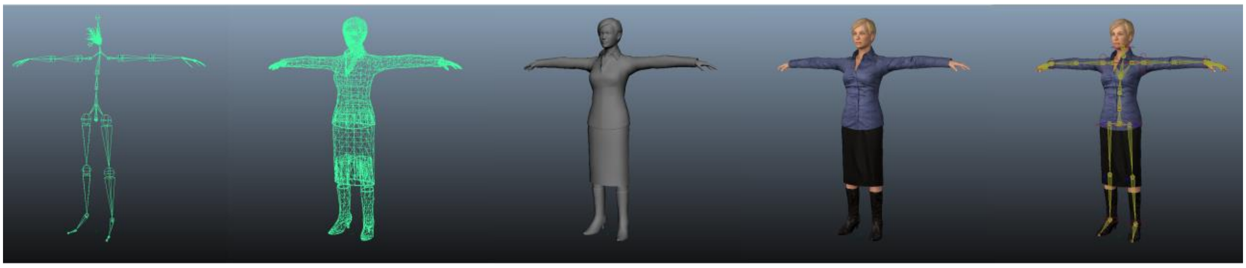

Figure 2. *Avatar components. From left to right: skeleton, mesh, surface, texture, control rig.*

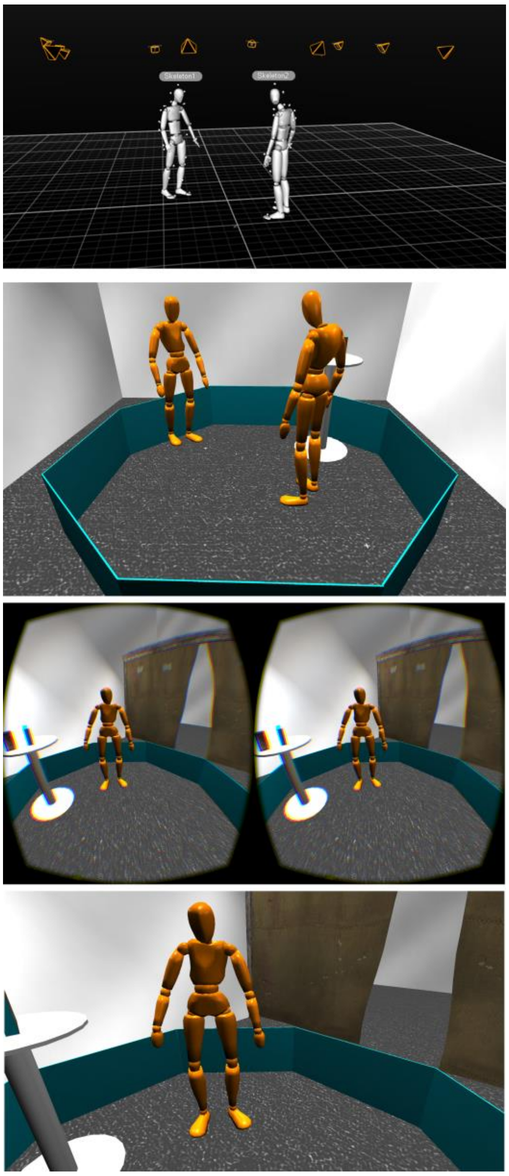

Figure 3. *Users in VR. From top to bottom: Motion Capturing Software, 3rd person perspective on VR scene, first person rendering for the head mounted display, first person camera viewport. Adapted from (Authors, in prep.).*