

Influence of Avatar Appearance on Presence in Social VR

Paul Heidicker*
University of Hamburg

Eike Langbehn†
University of Hamburg

Frank Steinicke‡
University of Hamburg

ABSTRACT

Social virtual reality (VR) has enormous potential to allow several physically separated users to collaborate in an immersive virtual environment (IVE). These users and their actions are represented by avatars in the IVE. In question is how the appearance of those avatars influences communication and interaction. It might make a difference, if the avatar consists of a complete body representation or if only certain body parts are visible. Moreover, a one-to-one mapping of the user's movements to the avatar's movements might have advantages compared to pre-defined avatar animations. To answer these questions, we compared three different types of avatar appearances in a user study. For this, we used estimations of presence, social presence, and cognitive load. The evaluation showed that motion-controlled avatars with full representation of the avatar body lead to an increased sense of presence. Motion-controlled avatars as well as avatars which have only head and hands visible produced an increased feeling of co-presence and behavioral interdependence. This is interesting, since it states that we do not need a complete avatar body in social VR.

Keywords: Collaborative interaction, perception, presence

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality

1 INTRODUCTION

Broad availability of *head-mounted displays* for consumers makes immersive telecommunication systems a viable alternative to audio communication and 2D video conferencing systems. In these systems, participants meet in an *immersive virtual environment* (IVE) that creates a common spatial and social context. IVEs with multiple users are denoted as *collaborative virtual environment* (CVE) or as *social virtual reality* (social VR).

Recently, several commercial applications were announced or released that use the opportunities of social VR. To depict the participants in these CVEs 3-dimensional avatars are used. Because of technical limitations these avatars are often subject to restrictions. We identified three different types of avatars that we wanted to compare in a user study (see Figure 1):

- Avatar with complete body and regular pre-defined idle animations (called *Idle* in the following) that is used in *AltspaceVR* for example¹.
- Avatar with complete body and a one-to-one mapping of the user's movements to the avatar's movements (called *Mapped* in the following) that is used in *High Fidelity* for example².

*e-mail: zheidick@informatik.uni-hamburg.de

†e-mail: langbehn@informatik.uni-hamburg.de

‡e-mail: steinicke@informatik.uni-hamburg.de

¹<https://altvr.com/>

²<https://highfidelity.com/>



Figure 1: Facebook's social VR, AltspaceVR, and High Fidelity are using different approaches of avatar appearance.

- Avatar body that consists only of head and hands and a one-to-one mapping of the user's movements to the avatar's movements (called *SocialVR* in the following) that is used in Facebook's social VR system for example³.

In this context, the question arises what effects do the visual representation of the user and the type of body movements have on the quality of communication in social VR systems. It might be assumed that a different appearance of the avatars has an important influence on the way they are perceived.

One of the most important factors influencing the quality of communication in mediated communication contexts is considered to be *social presence* [3]. It is defined as a measure of the perceived presence of another intelligent being and is determined by a variety of verbal, nonverbal, visual, conscious, and subconscious signals. A more detailed definition of social presence is presented by Biocca et al. [1], the so called *Networked Minds Measure of Social Presence*. They divide social presence into the three dimensions *co-presence*, *psychological involvement*, and *behavioral engagement*.

According to this, it leads to the assumption that systems which are using a one-to-one mapping of the user's movements to the avatar's movements provide a higher degree of social presence than other systems. Furthermore, it is interesting how much difference in social presence there is between a complete body representation and a representation that uses only head and hands. These assumptions were investigated in an experiment.

2 EXPERIMENT

For the experiment, we build a social VR system using Unity3D 5.3 and the Oculus Rift CV1, the Microsoft Kinect v2 as well as the Oculus Remote. For voice communication between the participants the integrated microphone of the Oculus Rift CV1 and the integrated headphones are used. 18 participants (6 female and 12 male, ages 19 – 45, $M = 27.2$) completed the experiment. The total time per participant was 60 minutes. During the experiment, the participants stood in front of the Kinect and wore the head-mounted display while they were able to interact by using a remote (see Figure 2).

³<https://www.facebook.com/mike.booth/posts/10209079557850643>

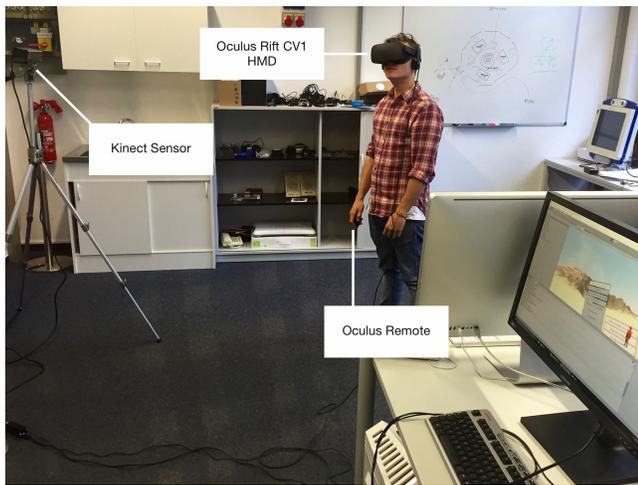


Figure 2: A participant during the experiment with an Oculus Rift CV1, a Microsoft Kinect v2 and an Oculus Remote.

We used a within-subjects experimental design with the three different avatar conditions mentioned in the introduction. During the experiment, two participants were using the social VR system at the same time. Therefore, we had two identical setups in two different rooms. Both participants experienced the same avatar condition at the same time. All participants completed all three conditions. The order of the conditions was randomized. The participants had to solve a collaborative task in the IVE. For that, we used the *Desert Survival Task* [2] which was already used to measure social presence in two- and three-dimensional video conferencing systems [1]. The goal of the task is to sort an unordered list with 15 items according to their importance for surviving in the desert.

After one third and two third of the items were sorted, there was a break to fill out the following questionnaires: The *Networked Minds Social Presence Questionnaire* to measure social presence, the *Slater Usoh Steed Presence Questionnaire* to measure presence, and the *NASA TLX Questionnaire* to measure the cognitive load.

3 RESULTS & DISCUSSION

A significant main effect of avatar appearance on SUS presence has been determined, ($F(2; 34) = 5.45; p = .009; \eta_p^2 = .243$). Post-hoc tests confirmed a significantly increased presence in the Mapped condition in comparison to the Idle condition ($p = .0187$), as well as in the Mapped condition in comparison to the SocialVR condition ($p = .0337$). There is no significant difference between the Idle and SocialVR condition.

The analysis of NASA TLX was not able to gain new insights. We could not find any significant differences of avatar appearance on cognitive load.

A significant main effect of avatar appearance on social presence could not be determined. But we found some effects for the sub-factors of social presence (see Figure 3).

There is a significant main effect within the sub-factor *co-presence*, ($F(2; 34) = 13.134; p = .0004; \eta_p^2 = .436$). Post-hoc tests showed a significant increased co-presence both in the Mapped condition ($p = .001$) and in the SocialVR condition ($p = .01$) in contrast to the Idle condition. There is no significant difference between the Mapped and the SocialVR condition.

A significant main effect has also been found within the sub-factor *behavioral interdependence*, ($F(2; 34) = 3.398; p = .045; \eta_p^2 = .167$). Post-hoc tests showed that compared with the SocialVR ($p = .039$) and the Mapped condition ($p = .048$), the behavioral interdependence in the Idle condition is significantly lower.

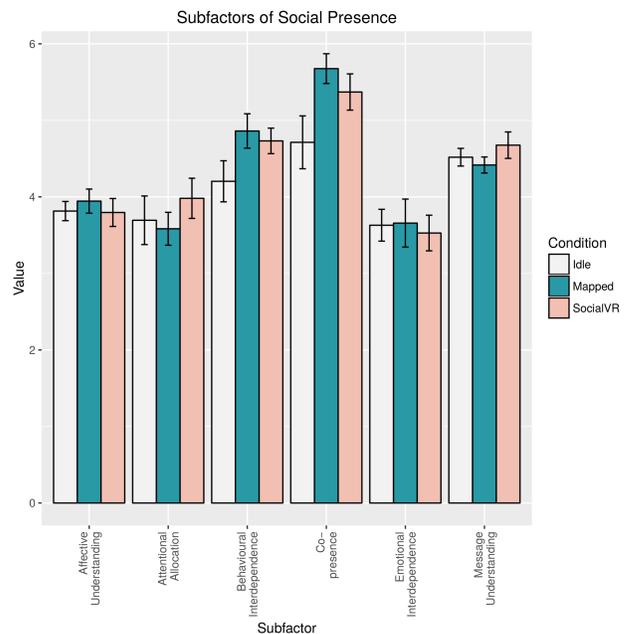


Figure 3: In the Co-presence and behavioral interdependence groups, the results of the Mapped and SocialVR conditions are significantly higher than the Idle condition.

There is no significant difference between the SocialVR and the Mapped condition.

Regarding the sub-factors of social presence it is conspicuous that the results for co-presence, behavioral interdependence, and message understanding are absolutely higher for all three conditions than the results for affective understanding, attentional allocation, and emotional interdependence. An explanation for this is that these sub-factors can be perceived by rough sensory informations, e.g., outlines or inexact gestures. The sub-factors with lower results are connected to emotional and affective states and require detailed gestures and facial expressions. So, it might be that the used system was not able to achieve this degree of details. Especially, the missing representation of facial expressions and posture of fingers belong to these. This might also be the reason why we have no significant effect on the overall factor of social presence.

4 CONCLUSION

It turned out that a complete avatar body with movements which are mapped from the user's movements generates the highest co-presence and behavioral interdependence. But an avatar that consists only of head and hands is not significantly worse, and even better than using a complete avatar body with pre-defined animations. An open question for future research is how many and which body parts have to be visible to reproduce or even surpass this degree of co-presence.

REFERENCES

- [1] J. G. Frank Biocca, Chad Harms. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. *4th annual international workshop on presence*, pages 1–9, 2001.
- [2] J. Lafferty, P. Eady, and J. Elmers. The desert survival problem. *Experimental Learning Methods*, 1974.
- [3] J. Short and e. etc. *The Social Psychology of Telecommunications* -. Wiley, New York, 1976.