Social Presence with Virtual Glass

H. Regenbrecht*, M. Alghamdi, S. Hoermann, T. Langlotz, M. Goodwin, and C. Aldridge
University of Otago, New Zealand

Fig. 1. Performer (left) and Instructor (right) meeting in a Virtual Glass enabled CVE (center)

Abstract

Collaborative Virtual Environments (CVE) with co-located or remote video communication functionality require a continuous experience of social presence. If, at any stage during the experience the communication interrupts presence, then the CVE experience as a whole is affected - spatial presence is then decoupled from social presence. We present a solution to this problem by introducing the concept of a virtualized version of Google Glass™ called Virtual Glass. Virtual Glass is integrated into the CVE as a real-world metaphor for a communication device, one particularly suited for collaborative instructor-performer systems. In a study with 65 participants we demonstrated that the concept of Virtual Glass is effective, that it supports a high level of social presence and that the social presence is rated higher than a standard picture-in-picture videoconferencing approach for certain tasks.

Keywords: Virtual Reality, Collaborative Virtual Environments (CVE), sense of presence, video-mediated, computer-mediated communication (CMC)

Index Terms: H.4.3 Information Interfaces and Presentation (e.g., HCI) - Communications Applications: Computer conferencing, teleconferencing, and videoconferencing

1 Introduction

Collaborative Virtual Environments (CVE) allow people to meet and interact in one shared, virtual space. In an effective CVE users experience a sense of presence in various aspects, most notably spatial presence – i.e. the sense of being part of the virtual environment – and social presence – i.e. the experience of being together. To support the experience of presence we have to design a believable virtual world, where users can interact with the environment and with each other [1,2].

With the Virtual Glass concept introduced here we propose an integrated communication solution combining communication and CVE collaboration in a way that maximizes social presence. This solution is suitable for a range of types of CVEs, in particular for systems which allow an instructor (e.g. therapist) and a performer (e.g. client or patient) to meet in the same virtual space, which we call instructor-performer systems, and fulfilling at least the following requirements: (a) self-controlled, first person view for the performer, (b) WYSIWIS (what you see is what I see) view for the instructor, (c) continuous, two-way audio and video communication between the instructor and the performer, (d) sense of social presence, i.e. a "sense of being there together", and (e) integration and believability of the communication interface.

2 Virtual Glass

In February 2013 Google released the developer version of a head-worn, optical see-through display with an integrated camera, microphone, headphone and a built-in computer controlled by simple gestural commands and voice. Of particular interest for our project are Google Glass's audio and video capabilities, i.e. the combined use of (1) the in-built miniature camera capturing (a portion of) what is in view of the user, (2) the microphone and headphones (bone conduction or otherwise) as communication devices, and (3) the miniature display as a potential video display device. The integrated use of these three components can lead to new forms of videoconferencing: (a) allowing for hands-free, on-the-move remote communication and (b) replacing a "see each other" with a WYSIWIS approach. WISYWIS can also be combined with standard videoconferencing (capturing the face/upper body of the participant) for an asymmetrical setup. Here, one person (e.g. the instructor) sees what the other person (e.g. the performer) sees, while the backchannel communication represents standard videoconferencing (seeing the other's face). This can lead to interesting extensions of the AR Remote Expert scenario [3] with rich real-time communication.

The performer interacts with an environment by navigating around and interacting with virtual objects. The instructor can take different viewpoints of the environment, including – and most relevant – the performer's virtual first person view. Instructor and performer communicate in such a way that the instructor can see the virtual viewpoint of the performer as well as the performer's face (web camera) as a picture-in-picture. The performer sees the instructor's face (web camera) as part of a virtualized version of a Google Glass™ device on a virtual miniature screen.

We have developed the Virtual Glass concept to allow for a continuous experience of social presence for the performer while immersed in the virtual environment and communicating with the instructor at the same time. To test the feasibility of the Virtual Glass concept we implemented a tailored version of our virtual

* holger@infoscience.otago.ac.nz

IEEE Virtual Reality Conference 2015
23 - 27 March, Arles, France
978-1-4799-1727-3/15/$31.00 ©2015 IEEE

reality exposure therapy system (i.e. with virtual stressors removed).

3 User Study

To test whether Virtual Glass is a feasible concept for instructor-performer systems communication and whether social presence can be improved over standard PiP (picture-in-picture)-enhanced CVE systems we built a prototype system implementing the approach.

Method: For evaluation, there were two independent variables focusing on the medium delivering the communication: a standard videoconferencing PiP interface, and the Virtual Glass. The main dependent variables included the users' perceived social presence and their understanding of the Visual Glass metaphor, as well as attention, video quality, audio quality, and the question of whether the communication medium would be seen by the participants as helping them to solve the task. The experiment followed a between-subjects design (Standard (PiP) Videoconferencing x Virtual Glass).

A total of 67 adult participants (aged 16 and above) took part in the experiment in 46 instructor-performer sessions (some participants took part in two session with different roles).

The experiment was conducted in one physical room but with two booths set far enough apart that the instructor and performer could not see or hear each other, except via the system. The instructor and performer were requested to work collaboratively to solve a shared task inside the CVE. The activity required tidying up the clutter in a virtual house in no more than 5 minutes. The instructor tells the performer where to find items and to where to place them. The performer executes the requested tasks (i.e., finding the items, picking up each one, and placing it in the appropriate place).

Fig. 2. Example room with multiple items of clutter (during the experiment only single items would be presented)

In order to have continuous communication and collaboration between the instructor and the performer, the clutter in the house (e.g., rubbish, food, shoes, clothes, tools, toys) was organized in sequence. This meant that there was only one item of clutter in the house at a time, and when the performer dealt with it, another item would appear elsewhere and the performer would need to deal with that one next. The instructor is notified via text instructions on the video screen of the identity and location of each next item, and can use the multi-viewing functionality to observe this and guide the performer.

The performer commences his or her task by "picking up" and "putting on" the Virtual Glass. This action enables the video communication channel. The instructor helps the performer to find the items in the house. This is done by reading the text instructions from the videoconferencing window and providing the performer with appropriate information. The text instructions include the type and location of the item and how the performer should deal with it (e.g., take the rubbish to the rubbish bin). The activity continues until the timer countdown ends. To measure the progress of the joint task, the counter for the instructor shows the number of objects found out of the total. For data collection, a paper questionnaire was used for both participants. Demographic information included gender and age. A seven-point Likert-like scale (from 1 "strongly disagree" to 7 "strongly agree") was used to measure the dependent variables: understanding, attention, video quality, audio quality, and the degree to which the communication medium helped the participant to solve the task. Social presence was measured using an unmodified scale by Short, Williams, & Christie (1976) [4]. This is a series of seven bipolar pairs that include items such as warm-cold, personal-impersonal, sensitive-insensitive, and sociable-unsociable.

In each session, two participants were pseudo-randomly assigned to the two booths/conditions.

Results: Overall, the results show that the Virtual Glass concept was well understood by the majority of the participants. They were clearly comfortable with this mode of communication and it seems to have been no less comprehensible than a standard video conferencing interface. This is quite surprising, as only four of our participants had previously tried a Google Glass or similar interface.

Despite the fact that both tested interfaces provided the same type of communication (video/audio), the interfaces performed significantly differently with respect to perceived social presence. The level of social presence when using the Virtual Glass was rated significantly higher than with the standard videoconferencing interface, thereby supporting our main hypothesis. We argue that the increased social presence results from (a) accepting the concept of the Virtual Glass, as well as from (b) continuously maintaining task performance in the VE, whilst (c) simultaneously communicating with the remote collaborator. Perhaps the Virtual Glass was seen as less "artificial" and therefore affected social presence positively.

While Virtual Glass had a significant effect on social presence, there was no corresponding effect on any of the other dependent variables: attention, video quality, audio quality, and solving the task.

To make an externally relevant comparison possible we chose a realistic setting and tasks. This however might introduce a role and task bias, because the performer and instructor had to act in different roles with different tasks, which included the instructor's ability to switch views, and therefore might have led to different social presence ratings, too. Also, the instructions given to the participants, even if they had been the same, might have contributed to the experience of social presence.

Acknowledgments

The authors wish to thank all of our participants in the experiment and Elgar Storm, George Stamou, Jonny Collins, Patrick Ruprecht, Carl Leichter, and Abdulaziz Alshaer for their help and support during this research. The presentation of this work at the conference is kindly funded by InternetNZ.

References

[1] Biocca, F., Harms, C. and Gregg, J. The Networked Minds Measure of Social Presence: Pilot Test of the Factor Structure and Concurrent Validity. Presented at PRESENCE 2001, 4th Annual International Workshop on Presence, (Philadelphia, USA, May 21-23, 2001).

[2] Paul de Greef, Wijnand A. IJsselsteijn: Social Presence in a Home Tele-Application. Cyberpsy., Behavior, and Soc. Networking 4(2): 307-315 (2001)

[3] Basile M., Ozdirlik B., Terrin JJ. (2009) Urban Projects and Multi-actor collaboration processes using mixed reality technologies. In : International Symposium on revitalising built environments: requalifying old places for new uses.

[4] Short, J., Williams, E., & Christie, B. (1976). The Social Psychology of Telecommunications. London: Wiley.