Spatial Presence in Real and Remote Immersive Environments

Nawel Khenak, Jean-Marc Vezien, David Thery, Patrick Bourdot

Spatial Presence in Real and Remote Immersive Environments

Nawel Khenak[1,3] Jean-Marc Vézien[1,3] David Thery[2,3] Patrick Bourdot[1,3]

(1) VENISE Team (2) CPU Team (3) LIMSI, CNRS, Paris-Saclay University, Orsay, France

ABSTRACT

This paper presents an experiment assessing the feeling of spatial presence in both real and remote environments (respectively the so-called "natural presence" and "telepresence"). Twenty-eight (28) participants performed a 3D-pointing task while being located in a real office and the same office remotely rendered over HMD. The spatial presence was evaluated by means of the ITC-SOPI questionnaire and users' behaviour analysis (trajectories of head during the task). The analysis also included the effect of different levels of immersion of the system – visual-only versus visual and audio – rendering in such environments. The results show a higher sense of spatial presence for the remote condition, regardless of the degree of immersion, and for the "visual and audio" condition regardless of the environment. Additionally, trajectory analysis of users' heads reveals that participants behaved similarly in both environments.

Keywords: Telepresence, Spatial Presence, Immersion, User Evaluation, Reality-Test.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information System—Virtual reality, Telepresence System; [Information Interfaces and Presentation]: Multimedia Information System—Evaluation

1 INTRODUCTION AND RATIONALE

Spatial presence defined as the sense of "being there" [1] encompasses the user's ability to experience a feeling of presence [7]. If the environment is real and non-mediated, the user experiences a *natural* presence; if the environment is real and mediated, the user experiences a *remote* presence; and if the environment is computer generated, the user experiences a *virtual* presence [6]. While studies have focused on the evaluation of the sense of presence across real and virtual environments, no study was conducted to compare between real and *remote* environments within the specific context of spatial presence. This was mainly due to the low immersive quality of technologies at that time. Indeed, many scholars emphasized the importance of high-quality natural sensory channels in the emergence of remote presence [6]. Recent technological breakthroughs, including HMDs visual quality, sound spatialization and overall system latency reduction, allow for more sophisticated user experiments in terms of realistic sensation and higher level of immersion, permitting better comparison of the sense of spatial presence between real and remote environments. The main goal of the conducted study is thus to evaluate the user's sense of spatial presence within real and remote immersive environments in almost similar visual and auditory conditions.

(3) firstname.lastname@limsi.fr

Figure 1: (Top) General setting of participants (Left) In the operating room. (Right) In the teleoperating room. (Bottom) First person view of participants.

So far, assessing presence was mostly performed through subjective questionnaires. Lessiter et al.'s ITC Sense of Presence Inventory (ITC-SOPI [3]) was chosen to measure the sense of spatial presence. This popular questionnaire has the advantage to be quite easy to administer and score. In parallel, an approach based on behavioural indicators sought to establish reliable and validated measures for presence [4]. Then, objective metrics related to user's behaviour were logged to study their relation with the feeling of spatial presence and whether they are conceivable tools for its objective assessment.

2 EXPERIMENTAL DESIGN AND METHOD

Two rectangular rooms with a similar office setting were used for the experiment. Subjects were either located in an "operating room", where they performed the task directly, or in a "teleoperating room", from which they performed the task through a teleoperation system. The task consisted in pointing a sequence of images that were displayed sequentially on 12 different tablets in a random but fixed order, one at a time (as soon as a tablet was pointed at, the image disappeared, and another one appeared in another location). The tablets were attached to the 4 walls of the operating room using a uniform sampling procedure. Sounds were played from the tablets. Thus, the subjects were seated in the middle of the operating or the teleoperating room and had to perform a sequence of pointing actions for a fixed period of time (3 minutes).

2.1 Hadware and Software set-up

Remote visual capture was obtained with a Ricoh Theta-V 360° panoramic camera placed at the centre of the operating room. Images were streamed at a resolution of 3840 x 1920@30fps to a real-time rendering engine visualized in a HTC Vive VR headset, with a 90Hz refresh rate. Because the field of view of the headset was limited to about 110°, subjects in the operating room wore a headset mock-up with an identical field of view.

An infrared tracking system was set up in the operating room consisting of a six-camera ARTrack5 tracking network. Subjects wore infrared markers on both their pointing hand and their head. In the teleoperation setting, head tracking was achieved using the Vive lighthouse system. Remote action was made possible by tracking user gesture with a Leap Motion sensor attached to the headset.

In the operating room, the subjects naturally experienced 3D sound perception coming from the tablets around them. To reproduce a similar auditory perception in the remote setting, a Tetramic First-Order ambisonic microphone was used to capture the sound, which was rendered binaurally over headphones (Sennheiser HD 280) in the teleoperating room. A patch designed in the Max MSP software managed all the audio processing pipeline. A reverberation delay of 1.2 second was added to simulate the acoustics of the operating room while the source aperture was fixed to 90°. Finally, the RMS level was set equal in both real and remote conditions to 75 dBA.

3 User Evaluation

A mixed-design analysis was run with two independent variables: the type of environment with 2 modalities labelled “real” and “remote”, representing respectively the operating room and the teleoperating room, and the level of immersion with 2 modalities labelled “with sound” and “no sound”, representing respectively the multisensory condition – visual with audio rendering – and the visual-only, silent condition. The order of the factors was counterbalanced across the 28 participants of the study (19 males, 9 females; Mean = 27,4 years, Sd = 4,4 years), using a Latin Square.

Once the pointing task was completed, participants were administered the ITC-SOPI questionnaire [3]. The participants performed the task and filled the questionnaire twice, for the “with sound” and “no-sound” condition, but only in one of the two environments. Finally, personal demographic variables (age, gender, experience with VR systems) were also collected.

In total, 54 trials were registered. The study comprised 2 different dependent variables, namely the perceived spatial presence deducted from responses to the questionnaire, and the user behaviour described by the head-related hesitation movement, i.e. the number of subject’s turnarounds to locate the targets during the evaluation.

4 Results and Discussion

The analyses were divided into a between-subject study between the environments and a within-subject study between the levels of immersion. All the analyses were performed using RStudio.

The perceived spatial presence was assessed using a two-way repeated measures ANOVA between both environments and levels of immersion. Statistically significant results were found between “real” and “remote” conditions $[F(52) = 16.7, p < .001, r = 0.2]$ and between “no-sound” and “with sound” conditions $[F(52) = 11.5, p < .01, r = 0.2]$. Thus, a Wilcoxon signed rank dependent and independent tests were performed respectively between levels of immersion and environments. A statistically significant higher sense of presence was reported in the “with sound” condition in both real and remote environments ($[V = 1.5, p < .01, r = 0.8]$; $[V = 7, p < .01, r = 0.9]$). A higher sense of presence was also found in the “remote” condition compared to the “real” condition for both “with sound” and “no-sound” conditions ($[W = 153.5, p < .01, r = 1.1]$; $[W = 158.5, p < .01, r = 1.0]$).

The higher sense of presence with audio shows the usefulness of a high degree of immersion to improve the feeling of presence [6]. More surprising, the higher sense of presence reported in the remote environment show that subjects experienced an illusion of more “real” presence in the remote environment. This hyper-presence [1] could be resulting from the lack of familiarity of participants with VR headsets (66,7% were beginners), especially in such a telepresence configuration.

Trajectory analysis was performed with the help of an ad-hoc hesitation indicator, which measured how often subjects would “change their mind” about the location of the pointing target.

A two-way repeated measures ANOVA was conducted to compare between the environments and levels of immersion. The results showed that for both real and remote environments, the hesitation of participants was reduced significantly when audio renderings were provided. Thus, participants had similar reaction whereas in the real and remote condition. However, results of comparison between the environments showed that participants hesitated more in the “remote” condition, especially in the “with sound” condition. A possible explanation is that localization was probably degraded by the low spatial resolution of the First-order Ambisonic microphone used during the evaluation. This highlights the importance of accurate reproduction of the auditory component to reach renderings comparable to real situations [5].

Finally, these results contradict those from the spatial presence scale of the questionnaire. Thus, it would be interesting not only to analyze the trajectory in its totality (research phase and pointing phase) but analyze each phase separately to provide more reliable tools based on behavioral indicators.

5 Conclusion

This paper presented a user evaluation of the sense of spatial presence between a real and a remote environment captured through telepresence technologies, with different levels of immersion. Such comparative analyses between real-world and telepresence conditions will benefit the presence research community as well as the designers of future teleoperation systems.

Acknowledgements

This research was supported by EquipEx DIGISCOPE (ANR-10-EQPX-26-01) operated by the French National Research Agency (ANR), as part of the program Investissements d'Avenir.

References

[1] F. Biocca. The cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of Computer-Mediated Communication*, *3*(2), 0-0, 1997.

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

[3] J. Lessiter, J. Freeman, E. Keogh, & J. Davidoff. A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and virtual environments*, *10*(3), 282-297, 2001.

[4] K. Mania, A. & Chalmers, A. The effects of levels of immersion on memory and presence in virtual environments: A reality centered approach. *CyberPsychology & Behavior*, *4*(2), 247-264, 2001.

[5] B. N. Postma, & B. F. Katz. Perceptive and objective evaluation of calibrated room acoustic simulation auralizations. *The Journal of the Acoustical Society of America*, *140*(6), 4326-4337, 2016.

[6] T. B. Sheridan. Musings on telepresence and virtual presence. *Presence: Teleoperators & Virtual Environments*, *1*(1), 120-126, 1992.

[7] S. Zhao. Reconceptualizing presence: Differentiating between mode of presence and sense of presence. In *Proceedings of PRESENCE 2002* (pp. 260-273), 2002.