Intelligent Human Systems Integration (IHSI 2023), Vol. 69, 2023, 428–434
https://doi.org/10.54941/ahfe1002862

Visual Cues Improve Spatial Orientation in Telepresence as in VR

Jennifer Brade[1], Tobias Hoppe[2], Sven Winkler[1], Philipp Klimant[1], and Georg Jahn[3]

[1]Institute for Machine Tools and Production Processes, Chemnitz University of Technology, Germany
[2]University of Twente in Enschede, The Netherlands
[3]Professorship for Applied Geropsychology and Cognition, Chemnitz University of Technology, Germany

ABSTRACT

This article addresses the question of how the number of visual cues in a telepresence scenario affects spatial orientation and whether the results are comparable with results in virtual reality studies. 30 participants completed a standard spatial orientation task, the triangle completion task, while using a remote-controlled telepresence robot. In a within-subject experiment, three cue conditions were examined that differed in terms of the availability of visual cues. The results of earlier research in VR were confirmed in telepresence by a reliably better orientation performance with rich compared to sparse visual cues.

Keywords: Telepresence, User studies, Triangle completion task, Spatial orientation

INTRODUCTION

When moving in reality, successful spatial orientation is enabled through continuous updating of egocentric spatial relations to the surrounding environment (Riecke et al., 2002). But in Virtual Reality (VR) or telepresence, cues of one's own movement are rarely provided, which typically impairs spatial orientation. While a large number of studies investigated spatial orientation in virtual environments, spatial updating in telepresence remains largely unexplored. Virtual and telepresence environments share the common feature that the user is not physically located in the mediated environment and thus interacts in an environment that does not correspond to the body-based cues generated by posture and self-motion in the real environment. We are interested in investigating whether findings from virtual reality research can be confirmed in a telepresence setting. Telepresence robots that are increasingly employed also due to the COVID-19 pandemic (Murphy et al., 2020) are typically operated by minimal real movements of the user via PC-based controls, which entail a lack of real translations and rotations and thus can disrupt spatial orientation (Cherep et al., 2020). This could become problematic if the tasks to be performed require a certain degree of spatial orientation in the mediated environment and could lead to detrimental effects on situation awareness and thus on task completion (Chen et al., 2007). Studies in VR show that a certain degree of spatial updating is possible without body-based cues to self-motion (vestibular, proprioceptive, motor efference) solely through continuous visual information about the change in orientation (Wraga et al., 2004). Various studies examined the type and number of visual cues and their influence on spatial orientation in virtual environments: Cherep et al. (2020) conducted five VR-experiments where participants performed a triangle completion task under different conditions concerning the availability of body-based, visual self-motion cues (optic flow), and other visual cues such as landmarks and boundaries. They showed that in conditions where there is a lack of body-based cues, additional landmarks improve spatial updating performance (Cherep et al., 2020). Kelly et al. (2009) compared different types – geometric and featural – of environmental cues and the cue quantity as well as their ambiguity in a virtual environment, displayed with a Head-Mounted-Display (HMD). Their results showed that with two environmental cues, participants stayed oriented compared to conditions with zero environmental cues. To the best of our knowledge, there have not yet been any studies that tested spatial orientation in telepresence applications with triangle completion. In addition, common, commercially available telepresence systems can usually only display the environment on a 2D monitor. The 2D monitor impairs the operator's depth perception compared with 3D presentation in VR (Forster et al., 2015, Boustila et al., 2017). Studies comparing stereoscopic and monoscopic vision show that monoscopic visualisation leads to worse distance estimation performance (Boustila et al., 2017) and worse navigation performance (Luo et al., 2021). Thus, it cannot be assumed without verification that the spatial orientation in 2D telepresence systems can be compared with that in VR systems. Therefore, we employed this standard spatial orientation task with a telepresence robot to evaluate if results concerning the number of visual cues turn out similar to findings in VR-studies. For that purpose we used a typical task for investigating spatial updating - the triangle completion task (TCT). During this task, the participants are guided or translated along two sides of a triangle and are then asked to move on the shortest way back to their starting point (Riecke et al., 2002).

METHODS

Based on previous studies and the associated research gap, we decided to examine the performance in a TCT using the remote, via 2D screen and mouse, controlled telepresence robot Double 3 (Robotics) (see Figure 1). To evaluate the influence of the number of visual cues on the performance in the TCT, three conditions that varied in the amount of visual information provided for navigating the third leg were presented in a within-subjects design. Each participant went through the same three orientation cue conditions in a TCT: sparse, medium, and many visual cues. The three different cue conditions were numerically evenly distributed and presented in random order. A total of 12 different triangles (three conditions x four different triangles) were used that were completed twice by each participant in pseudo-randomized orders, resulted in a total of 24 trials. The trial types differed by the combination

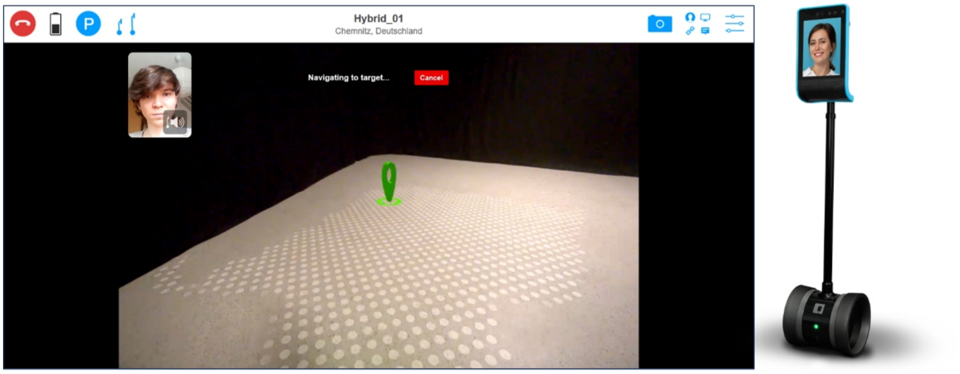

Figure 1: Left: web interface for controlling the telepresence robot; right: Double 3 telepresence robot.

of cue condition, TCT start position, and distance to be covered to the start position (in the following referred as homing distance).

All procedures were determined by the applicably body (Ethics committee of the Faculty of Behavioural and Social Sciences of Chemnitz University of Technology) not to require in-depth ethics evaluation (V-332-15-GJ-Telepresence-13052019).

Participants: Overall 35 people participated in the experiment, but due to technical problems (tracking, stability of the internet connection), data of only 30 participants (17 female, 13 male) could be included in the analysis. They had an average age of 29.17 years (SD = 11.39). The full experimental session took approximately 90 minutes, and participants were compensated by course credits. All participants had normal or corrected-to-normal vision, participated in the study voluntarily, and were informed that they were free to abort the experiment at any time.

Material: The experiment was carried out using a video conferencing tool for communication between the participant and the experimenter, online questionnaires, and remote control of a telepresence robot from the participant's home, in an environment unknown to the participant. A webcam, a microphone and a PC were required for participation. The participants used screens with a diagonal of 20.00 to 68.58 cm (M = 35.95). The participants controlled the telepresence robot remotely using a computer and a mouse and they communicated with the experimenter via microphone and chat.

The Double 3 was used as a telepresence robot, into which the participants dialled in via a web link (Figure 1). The Double 3 is a robot with a self-balancing wheel, stereo vision depth sensors, ultrasonic range finders, wheel encoders, and an inertial measurement unit. It has two 13 megapixel cameras as well as speakers and microphones for communication. The robot is height-adjustable between two positions min-height: 119 cm and max-height: 150 cm. In this study, only a height of 119cm was used, as this enables a faster movement speed.

TCT trials were performed with the Double 3 in a room of the "WohnX-perium" located in Chemnitz. The room size of the experimental area was

Figure 2: Presentation of the three conditions based on the richness of their visual cues: sparse visual cues (left), medium visual cues (middle) and many visual cues (right).

set to 5 × 9 meters by adjustable walls. The ceiling height of the room was 3.50m with a grid at 3m height. Spotlights were distributed on the ceiling of the room in order to sequentially indicate the vertices of the triangle on the floor as movement instructions for the participants. To record the movement of the robot, a mobile VIVE controller was attached to the robot, which recorded together with the STEAM VR base station 2.0 the movement of the Double 3 robot. Using two additional Garmin Virb Ultra 30 cameras placed diagonally to one another on the ceiling, the experimenter in the adjoining room monitored the experiment.

All three cue conditions were realized in a single preparation of the test room. The walls of the room were covered with black curtains and the following visual conditions were set up (Figure 2): two corners of the room were covered (sparse visual cues), one corner of the room was left visible (medium visual cues), and the other corner of the room was filled with furniture (many visual cues).

To examine all three conditions in the same test room, each of the four different triangles was rotated in such a way that a different number of visual stimuli could be seen when driving back to the start position. Which condition was examined was determined by the field of view of the robot camera during the last rotation, (the rotation in the direction of the starting point) and view of the travel path (third side of the triangle). This can be seen in Figure 3 by, for example, the orange triangle: on the last turn (top left), the robot turns to the left and during the return path to the starting point only the curtain is visible in the field of view of the camera.

Procedure: For each participant the following procedure was employed: First, the participants were welcomed by the experimenter in the online conference system and informed about the experiment. The participants answered demographic questions in an online survey and then dialled into the telepresence robot via a web link. After following an explanation on how to control the telepresence robot, the participants went through two example trials, after which the experimenter placed the robot on one of the 12 starting points with the view to the center of the room. The experimenter switched on a first spotlight via a mixer, which indicated the first vertex of the triangle for the participant. Then the participant moved (via click(s) on the floor of the mediated environment) the robot to the point of light on the floor.

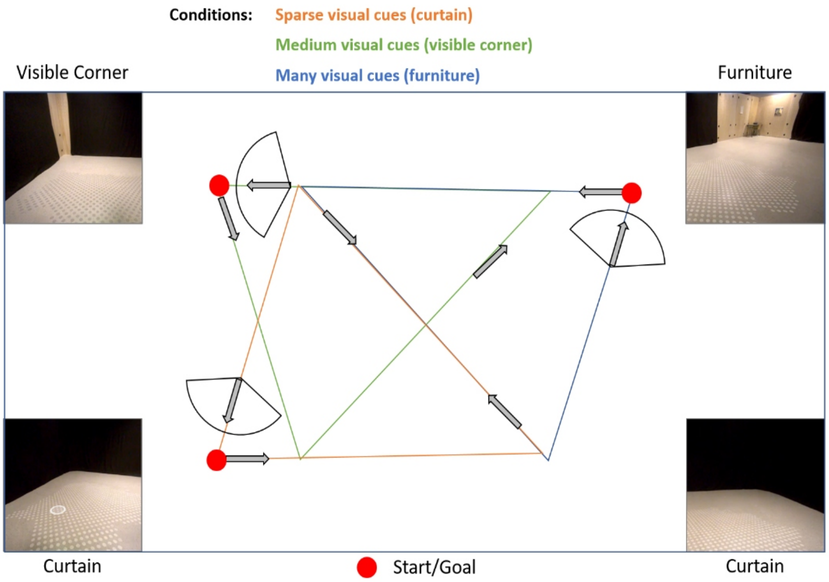

Figure 3: One of the four triangles and how it was placed in the room in each of the three conditions, e.g., orange triangle path in the Curtain condition starting on the bottom left to the right, heading to the top left on the second leg, and oriented towards the bottom left corner with a curtain on the third returning leg.

Upon reaching the first point, the spotlight was turned off and the second vertex of the triangle was visually indicated by a second spotlight. The participants turned the robot (via arrows on the screen) to orient towards the second specified point, moved there and were then asked to return the robot to its starting point. After they had completed one trial, the robot was set to the next starting position.

RESULTS

Our analysis focussed on the distance between the assumed and the actual starting position (in the following called "deviation") in the TCT. The mean deviation was lower in conditions with more visual cues (Figure 3), with the numerically best performance in the many visual cue condition ($M = 0.87$ m, $SD = 0.35$ m), followed by the medium visual cue condition ($M = 0.97$ m, $SD = 0.50$ m) and the sparse visual cue condition ($M = 1.06$ m, $SD = 0.38$ m).

A repeated-measures ANOVA with Greenhouse-Geisser-correction indicated a small and unreliable effect of the amount of visual cues ($F\ (1.63, 47.27) = 2.57$ with $p = .097$; partial $\eta^2 = .081$). To examine the trend visible in Figure 3 in more detail, paired t-tests were carried out. The difference between the sparse and the many condition ($t(29) = 3.129$, $p = .004$, $d = 0.57$, $[0.066, 0.314]$ 95% CI) was confirmed, while the differences between the sparse and medium ($t(29) = -0.956$, $p = .347$, $d = 0.17$, $[-0.279, 0.101]$ 95% CI) and the medium and the many ($t(29) = 1.078$, $p = .290$, $d = 0.19$, $[-0.091, 0.292]$ 95% CI) conditions were small and unreliable.

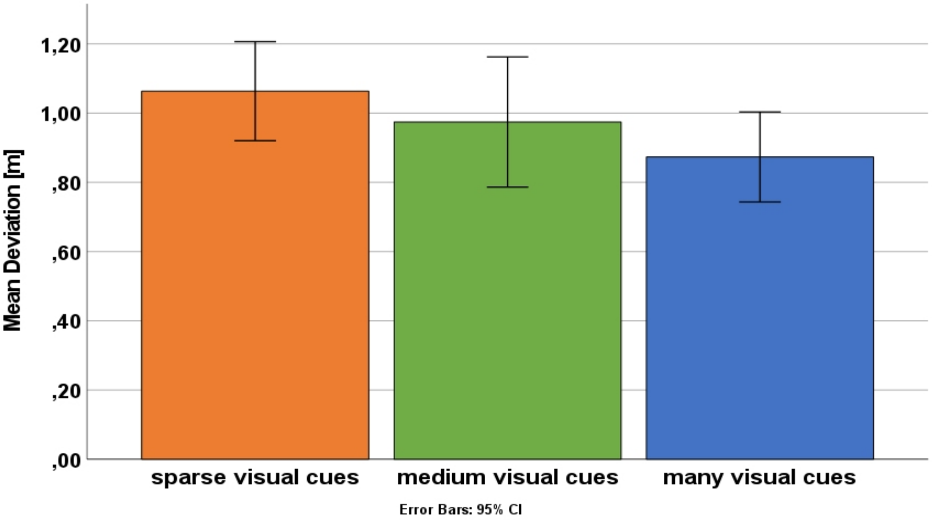

Figure 4: Bar charts for the distance error from the starting point for each condition (Error bars: 95% CI).

DISCUSSION AND CONCLUSION

Similar to studies that showed support of spatial orientation in triangle completion by visual cues in VR (Cherep et al., 2020, Kelly et al., 2009), the number of visual cues available while navigating the third leg supported triangle completion with a telepresence robot. This was confirmed by the trend of reduced error with more visual cues and a reliable difference between the conditions with sparse and many visual cues. Connecting results obtained in VR with telepresence and teleoperation scenarios is valuable to inform designing telepresence and teleoperation interfaces. We demonstrated that a standard task for studying spatial orientation performance is applicable with telepresence robots. In order to get a better understanding of how spatial orientation can be improved when interacting using telepresence systems, we are planning further studies varying the field of view, the type of interaction device, and investigating the differences between 2D and 3D displays. Another topic for future studies is orientation support (e.g., landmarks and pointers, auditory cues, maps, additional perspectives) and its effects on performance in this and additional tasks (e.g., distance estimation).

ACKNOWLEDGMENT

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 416228727 – CRC 1410.

REFERENCES

Boustila, S., Bechmann, D. & Capobianco, A. (2017). Effects of adding visual cues on distance estimation, presence and simulator sickness during virtual visits using wall screen. Proceedings of the computer graphics international conference. 1–6.

Chen, J. Y., Haas, E. C. & Barnes, M. J. (2007). Human performance issues and user interface design for teleoperated robots. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 37, 1231–1245.

Cherep, L. A., Lim, A. F., Kelly, J. W., Acharya, D., Velasco, A., Bustamante, E., Ostrander, A. G. & Gilbert, S. B. (2020). Spatial cognitive implications of teleporting through virtual environments. Journal of Experimental Psychology: Applied, 26, pp. 480–492.

Forster, Y., Paradies, S., Bee, N., Bülthoff, H., Kemeny, A. & Pretto, P. (2015). The third dimension: Stereoscopic displaying in a fully immersive driving simulator. Proceedings of DSC 2015 Europe Driving Simulation Conference & Exhibition. Max Planck Institute for Biological Cybernetics Tübingen, Germany, 25–32.

Kelly, J. W., Mcnamara, T. P., Bodenheimer, B., Carr, T. H. & Rieser, J. J. (2009). Individual differences in using geometric and featural cues to maintain spatial orientation: Cue quantity and cue ambiguity are more important than cue type. Psychonomic bulletin & review, 16, 176-181.

Luo, Y., Wang, J., Liang, H.-N., Luo, S. & Lim, E. G. (2021). Monoscopic vs. stereoscopic views and display types in the teleoperation of unmanned ground vehicles for object avoidance. 2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN). IEEE, 418–425.

Murphy, R. R., Gandudi, V. B. M. & Adams, J. (2020). Applications of robots for COVID-19 response. arXiv preprint arXiv:2008.06976.

Riecke, B. E., Veen, H. a. H. C. V. & Bülthoff, H. H. (2002). Visual homing is possible without landmarks: A path integration study in virtual reality. Presence: Teleoperators & Virtual Environments, 11, pp. 443–473.

Robotics, D. Available at: https://www.doublerobotics.com/double3.html (Accessed: 05 January 2023).

Wraga, M., Creem-Regehr, S. H. & Proffitt, D. R. (2004). Spatial updating of virtual displays. Memory & cognition, 32, 399–415.