International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2015)
M. Imura, P. Figueroa, and B. Mohler (Editors)

Global Landmarks Do Not Necessarily Improve Spatial Performance in Addition to Bodily Self-Movement Cues when Learning a Large-Scale Virtual Environment

T. Meilinger[1], J. Schulte-Pelkum[1], J. Frankenstein[1,2], D. Berger[1,3] and H.H. Bülthoff[1,4]

[1] Max Planck Institute for Biological Cybernetics, Tübingen, Germany
[2] University of Mannheim, Mannheim, Germany
[3] Harvard University, Cambridge, MA, USA
[4] Korea University, Seoul, Korea

Abstract
Comparing spatial performance in different virtual reality setups can indicate which cues are relevant for a realistic virtual experience. Bodily self-movement cues and global orientation information were shown to increase spatial performance compared with local visual cues only. We tested the combined impact of bodily and global orientation cues by having participants learn a virtual multi corridor environment either by only walking through it, with additional distant landmarks providing heading information, or with a surrounding hall relative to which participants could determine their orientation and location. Subsequent measures on spatial memory only revealed small and non-reliable differences between the learning conditions. We conclude that additional global landmark information does not necessarily improve user's orientation within a virtual environment when bodily-self-movement cues are available.

Categories and Subject Descriptors (according to ACM CSS): H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems —Artificial, augmented, and virtual realities

1. Introduction and previous work

Virtual environments (VEs) are used in a rapidly increasing range of applications, one of which is the study of navigation. The current experiment makes use of the versatility of modern virtual reality technology in order to systematically vary the information available during a navigation task. Specifically, we examine how multiple cues contribute to the perception of the surrounding environment and a navigator's position and movement. These cues can be categorized into visual cues, bodily cues, and other cues such as audition or haptics. Visual cues are potentially available, for example, during desktop navigation, and include landmarks relative to which navigators determine their position and orientation, optic flow (i.e., the movement pattern on the retina indicating self- and object movement), as well as depth cues, for example, texture gradients (distant tiles of the same size occupy a smaller space on the retina), motion parallax (distant objects move slower across the retina than close-by objects during self-movement), or stereo vision. Bodily cues as understood here consist of vestibular and proprioceptive cues. Vestibular cues from the inner ear provide rotational and translational acceleration information to infer self-movement and navigator's orientation relative to gravity. This information is provided, for example, in movement simulators. Proprioceptive cues give a sense of the position and movement of the parts of a navigator's body and are complemented by "efference" copies from planned movements. Proprioceptive cues are available when physically turning and/or walking, for example, on a treadmill or in a tracked space.

Several experiments examined the usefulness of vestibular and proprioceptive information in addition to mere visual information for spatial learning. In general, providing vestibular information in addition to visual information does not seem to help: measures of metric relations typically do not differ between learning conditions [WLS03, WG07, CW13]. Contrarily, proprioceptive and vestibular information together seem to provide small, but reliable advantages relative to visual only learning [WLH04, WG07, RVB11] and even to visual and vestibular information together [CW13]. However, this advantage is not always found and might require a sufficiently complex environment [WG07, MLP*10]. Results from Ruddle et al. [RVB11] suggest that

DOI: 10.2312/egve.20151306

physical translation might be more important than physical rotation. However, their experiment used a grid like environment with global orientation cues: surrounding walls had individual textures and at least one wall was visible from almost every location within the environment. Participants might have relied on global landmarks rather than physical rotation.

Global landmarks were shown to help orienting in VEs experienced visually only. Navigators used them to guide their local route decisions [IDB*10, SM00]. Similarly, navigators exploited global heading information from virtual slant both for route decisions and in configurational learning [RSMM04].

When learning complex VEs navigators profit both from proprioceptive information and global heading cues provided in addition to local visual cues. However, it is an open question how proprioceptive and global heading cues interact. In the present experiment, we examined whether global landmarks aid spatial learning if provided in addition to proprioceptive cues. Furthermore, we differentiated between two kinds of global landmark information. Firstly, a global landmark (i.e., a mountain silhouette) placed in infinity providing heading information only. Secondly, a global landmark (i.e., a factory hall) providing heading and distance information. Contrary to the distant mountain scenery, participants could locate themselves relative to the hall. We expected navigators to profit from both kinds of global information, but more so from the heading and distance information within the hall.

Figure 1: *A participant equipped with tracking helmet, hmd and cover using the pointing device to point towards previously learned locations.*

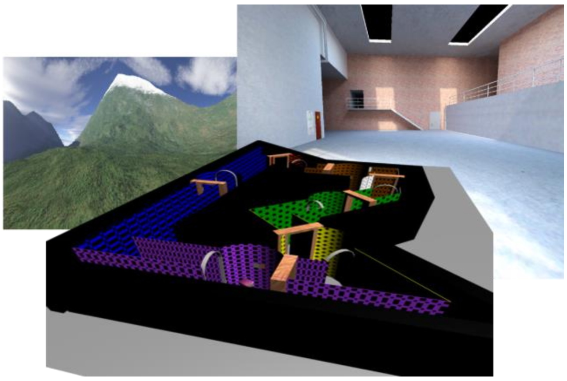

Figure 2: *Snapshots of the labyrinth (front), and parts of the mountain silhouette (back left) and the factory hall (back right), which surrounded the labyrinth.*

2. Technical description

Users freely walked through a 15x12m large space with their head coordinates tracked by 16 high-speed motion capture cameras at 120 Hz (Vicon® MX 13). Coordinates were transmitted wirelessly (using WLAN) to a notebook computer (Dell XPS M170) mounted on a backpack. This notebook rendered an egocentric view of a virtual environment in real-time. Participants viewed the scene in stereo using a very light-weight head-mounted display (eMagin Z800 3D Visor) providing a field of view of 32×24 degrees at a resolution of 800×600 pixels for each eye. The interpupillary distance was fixed at 6.5 cm. We adjusted the fit of the head-mounted display and the placement of the display screen individually for each participant. Frame rate was 30Hz with a latency of approximately 140ms. The setup provided important depth cues such as stereo vision and motion parallax, as well as all bodily cues important for orientation such as efference copies, vestibular, and proprioceptive information.

For testing, participants sat on a chair and estimated directions using a custom-build pointing device (Figure 1). The joystick-like device consisted of two tubes connected with a flexible hose. The lower tube was mounted on a tripod and had a button, which was used to estimate latency measures. The end-tube contained two orthogonal acceleration sensors indicating direction, when deflected from vertical orientation. The device provided a resolution of approximately 2°. It is, therefore, not only more precise than many common joysticks, but also gives much more direct estimation of target orientation as it can be fully deflected into horizontal orientation.

The virtual environment consisted of a labyrinth, and optionally an additional mountain silhouette, a virtual room, or a black background (Figure 2). The labyrinth was used in a previous experiment [MRB14] and encompassed seven connected straight corridors, each with a different wall color and texture. The corridors formed one closed loop without any junctions (Figure 3). In the middle of each corridor a cylindrical room was located with a wooden entrance and a metal exit door which were clearly distinguishable. Seven distinct target objects were placed at a height of 1.3 m, one in each corridor room. The height of each corridor and door were fixed individually to participant's eye-height plus 15 cm. This ensured that participants could not look over the walls even when standing on their toes, but had a good view on any distant information. When looking straight within a corridor the background covered roughly 20-25% of the visual field. The mountain silhouette was used in previous experiments as an orientation cue [DB08] and was placed inside a box virtually at infinity, thus providing no reliable distance cues to the mountains. The room was an industrial hall providing distance and familiar size cues in addition to orientation information.

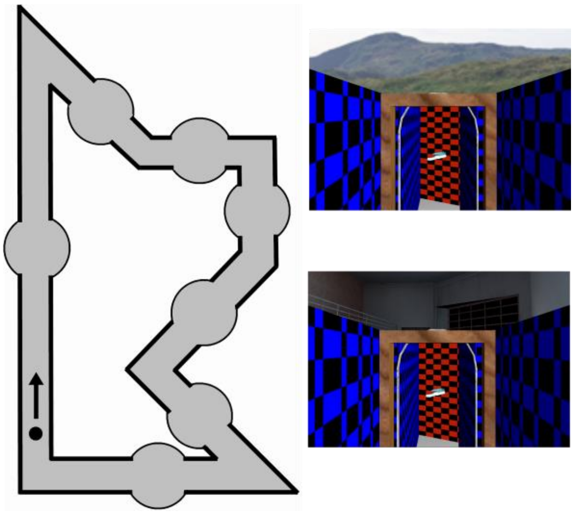

Figure 3: *Left: Schematic drawing of the labyrinth layout with the starting point and walking direction marked. Right: Egocentric snapshots displaying a participants view from within with the mountain and the hall in the background.*

3. Experiment

33 Participants (19 female; age mean: 24 years, standard deviation: 3.3 years) participated within a between-subject design with the factors 'labyrinth only', mountain, and 'factory hall' (Table 1). They were randomly assigned to the three conditions (11 per group).

To learn the environment, participants were asked to walk ten times clockwise through the corridors. At the end of the tenth passage, participants were shown the wall texture of a corridor and were asked to name the object located in that corridor room. Participants who did not name all objects correctly could walk two extra rounds before being asked again. In average learning lasted 22.5 minutes.

In the following test phase, participants were seated on a chair in front of the pointing device. Through the HMD, they were presented with a view of one of the seven corridor rooms with the objects removed and the room doors closed. Participants' positions were exactly where target objects had been situated during the learning phase, but in different visually simulated orientations for different trials. Each participant performed 56 trials, consisting of a factorial combination of seven locations (one for each corridor) × eight different visually simulated body orientations (0°, 45°, 90°, 135°, 180°, 225°, 270° and 315°). Trial order was randomized for each participant. In each trial, participants were asked to first confirm their location and heading and afterwards point towards a randomly chosen target object. The time for self-localization was recorded as the time between the initial presentation of a new view and the button press by which participants confirmed that they knew their location and orientation. Immediately afterwards, participants were asked to point as accurately and quickly as possible to a target (one of the seven learned target objects randomly chosen) whose name appeared on the screen. Targets were always occluded by the room walls. During self-localization participants were free to look around. However, during pointing they had to keep a straight head orientation, otherwise the pointing was not accepted. Error was computed as the absolute deviation between correct and estimated pointing direction. Including briefing and debriefing the experiment lasted about 1.5 hours.

Table 1: *Overview of the experimental design.*

4. Results

To obtain estimates robust against outliers we computed median values per participant for self-localization time, pointing time and absolute errors. Participants learned the spatial layout; performance in each group was better than the chance level of 90° obtained when randomly pointing in any direction, all three $t(10)$'s > 5.20, p's $< .001$. As shown in Figure 4, performance was highly similar in the learning groups. Results of a one-way ANOVAs indicated similar performance in the groups for self-localization time, $F(2,30)=0.14$, $p=.866$, $\eta_p^2=.01$, pointing error, $F(2,30)=0.66$, $p=.526$, $\eta_p^2=.04$, and pointing latency, $F(2,30)=0.86$, $p=.433$, $\eta_p^2=.05$

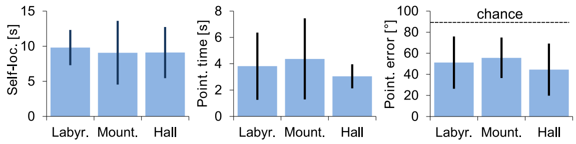

Figure 4: *Performance in self-localization and pointing as a function of the three learning conditions. Means +- 1 Standard deviation are displayed.*

5. Discussion

Prior studies showed that proprioceptive cues [WG07, CW13, RVB11, WLH04] and global heading cues [IDB*10, RSMM04, SM00] enhance spatial learning in addition to learning with local visual cues only. The present study examined how proprioceptive and global landmarks together influence spatial learning compared to proprioceptive and local visual information only. Surprisingly, participants did not profit much from additional global landmarks. For learning the environment examined, proprioceptive cues seem to be sufficient. Providing users with proprioceptive information seems to be sufficient to keep them oriented. Additional global orientation cues might not necessarily improve performance.

Thirty-three participants conducted the present experiment. Was this number too small to establish reliable differences? We think that if there exists a substantial advantage of additional global orientation cues, it should be more prominent in the results, as we analyzed three different measures and 56 trials per participants summing up to 5544 data points.

The VE consisted of several partly oblique corridors; it was looped and participants experienced it multiple times. It might be possible that participants exploring larger, linear environments once or twice only might indeed profit from global landmarks such as towers or hills. This question is subject to future research.

A broad variety of cues within a VE help navigators to orient themselves. The present work shows that proprioceptive information can be sufficient for navigators to stay oriented. Additional global orientation cues do not necessarily improve performance.

6. Acknowledgements

Part of Heinrich H. Bülthoff's research was supported by the Brain Korea 21 PLUS Program through the National Research Foundation of Korea funded by the Ministry of Education. Correspondence should be directed to Heinrich H. Bülthoff.

References

[CW13] Chrastil, E.R., Warren, W.H.: Active and passive spatial learning in human navigation: acquisition of survey knowledge. *Journal of Experimental Psychology: Learning, Memory & Cognition* 39, 5 (2013), 1520–1537.

[DB08] Doeller, C.F., Burgess, N.: Distinct error-correcting and incidental learning of location relative to landmarks and boundaries. *Proceedings of the National Academy of Sciences U. S. A.* 105, 15 (2008), 5909-5914.

[IDB*10] Iglói, K., Doeller, C.F., Berthoz, A., Rondi-Reig, L. Burgess, N.: Lateralized human hippocampal activity predicts navigation based on sequence or place memory. *Proceedings of the National Academy of Sciences U. S. A.* 107 (2010), 14466-71.

[TRB14] Meilinger, T., Riecke, B.E., Bülthoff, H.H.: Local and global reference frames for environmental spaces. *Quarterly Journal of Experimental Psychology* 67 (2014), 542–569.

[MLP*10] Mellet, E., Laou, L., Petit, L., Zago, L., Mazoyer, B., Tzourio-Mazoyer, N.: Impact of the virtual reality on the neural representation of an environment. *Human Brain Mapping* 31, 7 (2010), 1065–1075.

[RSMM04] Restat, J.D., Steck, S.D., Mochnatzki, H.F., Mallot, H.A.: Geographical slant facilitates navigation and orientation in virtual environments. *Perception* 33 (2004), 667–688.

[RVB11] Ruddle, R.A., Volkova, E., Bülthoff, H.H.: Walking improves your cognitive map in environments that are large-scale and large in extent. *ACM Transactions on Computer-Human Interaction* 18, 2 (2011), 1-20.

[SM00] Steck, S.D., Mallot, H:A.: The role of global and local landmarks in virtual environment navigation. *Presence Teleoperators Virtual Environ.* 9, 1 (2000), 69–83.

[WG07] Waller, D. & Greenauer, N.: The role of body-based sensory information in the acquisition of enduring spatial representations. *Psychological Research* 71, 3 (2007), 322–332.

[WLH04] Waller, D., Loomis, J.M., Haun, D.B.M.: Body-based senses enhance knowledge of directions in large-scale environments. *Psychonomic Bulletin & Review* 11, 1 (2004), 157–163.

[WLS03] Waller, D., Loomis, J.L., Steck, S.D.: Inertial cues do not enhance knowledge of environmental layout. *Psychonomic Bulletin & Review* 10, 4 (2003), 987–993.