Using Virtual Reality for Prototyping Interactive Architecture

Katrin Wolf
HAW Hamburg
Berliner Tor 5, 20099 Hamburg
katrin.wolf@acm.org

Markus Funk
TU Darmstadt
Hochschulstraße 10, 64289
Darmstadt
funk@tk.tu-darmstadt.de

Rami Khalil
ETH Zurich
Rämistrasse 101, 8092 Zurich
rkhalil@ethz.ch

Pascal Knierim
LMU Munich
Geschwister-Scholl-Platz 1,
80539 Munich
pascal.knierim@ifi.lmu.de

MUM 2017, November 26–29, 2017, Stuttgart, Germany.

DOI: https://doi.org/10.1145/3152832.3156625

Abstract

Even though, three-dimensional representations of architectural models exist, experiencing these models like one would experience a fully constructed building is still a major challenge. With Virtual Reality (VR) it is now possible to experience a number of scenarios in a virtual environment. Also prototyping interactive architecture elements, which might be very expensive, becomes possible. Thus, researchers and designers can already start to define user interfaces for interactive architectural elements, before they were even built. However, it is still an open question how exactly VR technologies can support experiencing interactive architecture. To answer this question, we compared experiencing three-dimensional architectural models and interactive architectural elements through a 2D screen & mouse+keyboard navigation, an head mounted display (HMD) & keyboard navigation, and an HMD & walking for navigation. The results of our study show challenges and opportunities regarding the immersion of these experiences.

Author Keywords

Virtual Reality; Interactive Architecture;

ACM Classification Keywords

H.5.2 [Information interfaces and presentation (e.g., HCI)]: User-Interfaces

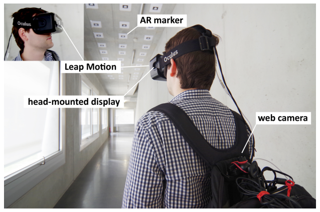

Figure 1: Our mobile VR system allowing the user to walk through an interactive archetectural prototype.

Introduction & Background

Virtual Reality (VR) allows users to immerse into digital environments through technology that dominantly stimulates their sensory system while reducing or preventing from stimulation caused by the physical environment. Perceiving VR to be a real place is creating the sensation of ‘being there’, often called ‘presence’ [10]. Presence is widely accepted to increase through using meaningful and coherent information of different modalities [2, 7, 13, 14]. The perceived level of presence of the user in Virtual Environments (VEs) is an established measurement for VR immersion [17]. Presence, which is highly related with immersion, is defined as the subjective experience of being in one place or environment, even when one is physically situated in another [17].

Early systems already allowed the user to look around in VR using 3D head mounted display (HMD) [12]. Slater et al. [11] proposed a “walking technique” aiming to increase the user’s sense of presence within VR as free physical walking covers more of the user’s senses [15]. Due to the dominance of vision over other senses [16], a multimodal experience (including haptic or proprioceptive information) can be created by using vision only. For example, Marchal et al. used purely vision and the change of perspective to generate the experience of walking up and down in VR [6]. Schmidt et al. [8] went a step further. They developed a VR system that allows the user to physically step up in VR, e.g., on a box. Their system “Level-Ups” provides visual and haptic information about stepping on virtual boxes through a HMD and motorized stilts. Cheng et al. [1] proposed a motion platform consisting of several assistants who are holding the user to simulate sitting in a swing or hang-gliding. When reflecting on their system, Cheng et al. state that getting the timing of feedback right is the main challenge in human-activated haptic feedback for VR. Simeone et al. were exploring to provide haptic feedback in VR when grasping physical objects [9]. Again, aiming for an increase of presence in VR, they created a VR experience in which the participants’ sensory perceptions (e.g., temperature, size, touch, and weight) about real life objects was affected by the design and appearance of their representation in VR.

While VR haptic feedback has been used for creating more immersive VR experiences, using VEs and spatially aligned haptic feedback for experiencing interactive architecture is still an under-explored area. To explore this topic, we conducted a user study for identifying the strengths and weaknesses of this approach.

Experiment

We designed an experiment to explore perception of virtual architecture through implementing typical actions people do in architecture under various conditions of VR fidelity.

Design and participants

Our experiment follows a 3x3 within subjects design. We controlled the architectural prototype regarding the interface technologies: 2D screen & mouse+keyboard navigation, HMD & keyboard navigation and HMD & walking for navigation (Figure 1). We assume that the sensation of "being in" the prototyped environment will lead to better ways to evaluate the architectural prototypes. Therefore, the dependent variables were presence recorded with the Presence Questionnaire [17] and qualitative opinions on aspects that supported or disturbed feeling presence in the architectural VR. We counterbalanced the order of the conditions to avoid sequence effects. We recruited 18 (5 female and 13 male) participants through our university’s email list, aged between 18 and 45 (M=22.8, SD=6.0).

Apparatus

Our three conditions had the following setups:

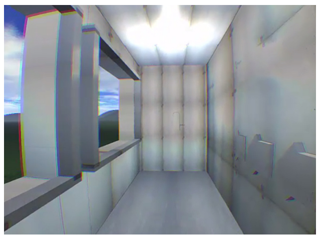

Figure 2: Visual guidance application

Figure 3: Emerging chair application

Figure 4: Grasping objects application

***(C1) 2D screen & mouse+keyboard navigation*:** The first condition used as setup a regular desktop with a screen, mouse, keyboard, and a Leap Motion sensor placed behind the keyboard. The user is seated in front of the desk. The VR is shown to the user on a monitor. The WASD-keys were used to move, and the mouse to virtually look around in the virtual room. The hands, sensed with the Leap Motion, could be used to interact with the environment. To sit down on the box (see task description below), the left mouse button had to be pressed. Once seated, pressing the left mouse button again would result in standing up.

***(C2) HMD & mouse+keyboard navigation*:** The setup was the same as in C1 except that an Oculus Rift was used as visual output for the VE. The Leap Motion sensor was mounted at the front of the Oculus Rift. A screen, which was not seen by the participants, was used to show an assistant the moment when the participants wanted to grasp the basket (see task description).

***(C3) HMD & walking for navigation*:** The third condition's setup uses again the VR headset with Leap Motion. Both was connected to a laptop which the user was carrying in a backpack. Hence, they were able to freely walk in the physical space and move accordingly within VR. Our institute's corridor was the physical surrounding environment, and the walls of this area were mapped to the dimensions of the VE. Touching the walls in the VE would cause haptic feedback from touching the real walls in our building. The position of the user is tracked using a camera which recognizes AR-markers that are mounted on the ceiling. We were using a 18x3 matrix resulting in 54 unique markers, which allowed us to cover a corridor with 2.2m times 12m in size. Grasping and looking around was implemented similarly to the setup of C2. The seat position and the basket orientation (see task desciption) were shown to an assistant through screen mirroring on a projected display that was showing the participant's view of the VE.

Task and procedure

After participants had filled in the consent form and had gotten information of the course of the study, we equipped them according to the different condition setups. For each condition, we asked the participants to explore the VE in three scenarios that represent three ways of exploring an architectural prototype and acting within such environment.

***(T1) Visual exploration*:** Participants move through VR by following arrows, which appear at the wall, see Figure 2.

***(T2) Sitting down*:** Participants move through VR and can sit down on a box, see Figure 3.

***(T3) Grasping objects*:** Participants move through VR and can grasp a basked, see Figure 4. Inspired by the Haptic Turk [1] the user can grasp the basket and feel the real object as an assistant is handing over a basket to the user in the moment he or she is grasping it.

The order of the scenarios, which serve as experimental task, within each condition was randomized. After each task, the participants filled in the Presence Questionnaire. Moreover, we collected subjective opinions about each condition regarding aspects of the system that supported and disturbed the feeling of presence in VR. Finally, we collected demographic data followed by the debriefing.

Results

We analyzed presence inferential statistic and qualitative feedback on presence regarding aspects that support or distract presence.

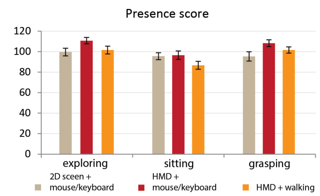

Figure 5: Presence (with means and SD).

Presence
The condition with HMD & keyboard navigation, where participants did not physically move, yielded to the highest sense of presence (considering the total score across all tasks, see Figure 5). We use a Friedman test to analyze if the different conditions influence the perception of presence during the different tasks. We found a significant effect on the presence rating ($\chi^2(8) = 32.655$, $p < .001$). Post hoc analysis with Wilcoxon signed-rank tests were conducted with an applied Bonferroni correction, resulting in a significance level of $p < .0014$. We found significant differences for presence between T2inC3 and T3inC3: $p = .001$ and between T2inC2 and T3inC3: $p = .001$).

Aspects that distracted or supported presence
Analyzing the qualitative feedback on aspects of the system that supported or distracted the sensation of being present in the architectural environment shall, for example help to better understand why C3 lacks in presence even though it combines most modalities through allowing to walk (versus C2). During our analyses, we focused on feedback about the coherence and consistence of stimuli as these aspects had been stated to be essential for the sense of presence on VR [7, 17]. Using a bottom-up analysis and open coding, we identified groups of participants' comments being clustered according input and output modalities, which therefore structure our results:

***Visual output*:** In general the 3D view using Oculus Rift was perceived more authentic then a 2D screen. P18, for example, said *"it looked very real and it was easy to accept the display as some sort of reality"*, and using the Leap Motion added to that feeling of presence as participants would see a representation of their hands and use them to interact with the environment (P1, P7, P8).

***Moving through walking*:** During the seated conditions (C1, C2), *"walking was a bit weird"* (P17) and *"turning the head did not feel real"* (P17). In contrast participants like to move in VR. That felt like *"moving in a real world"* (P10) and that they *"could freely walk around"* (P16). Further, *"walking and watching the environment (through virtual windows) was good"* (P13), *"walking around was very intuitive"* (P4), and moving through walking allows for using the hands for object manipulation as *"using mouse and keyboard for moving were keeping the hands busy and did not allow for exploring the environment with them"* (P16).

***Sitting down*:** Sitting down in the desktop conditions (C1, C2) was realized with a mouse click, which was not perceived to be natural (P12, P13). Sitting down while walking had the nice advantage of performing *"a real sitting down and standing up action"* (P16). In the mobile setup (C3), participants like to have a real perceivable seat in the scene to sit on (P5, P6, P15). Again according the consistency between the real world and VR, participants made different experiences. While P12 honored that the seat in the physical world was placed where it was shown in VR, P18 said that *"the slight difference in where the real world objects were placed compared to where they were displayed was a bit distracting"*. We assume the position mismatch of the seat may be caused by the variance in human perception and reaction that we rely here on using a Wizard-of-Oz method.

***Hand manipulation*:** In general the possibility to use hands for manipulation was appreciated (T3). Participants liked *"the visual design of the virtual hands"* (P5), the possibility to *"grab the object and move it around"* (P7), and to *"interact with objects with hand gestures"* (P7). For some participants the hand gesture recognition worked *"precise"* (P12) and was perceived to be *"authentic"* (P13) and *"feel-*

ing surface structures of the handle" (P15) was liked as well as that they could *"move objects in real time"* (P6). However, for others *"the feeling of touching something in the real world didn't match the distance of the object in the simulated world"* (P18) and some either *"couldn't grab the basket"* (P17) or *"felt the handle without grabbing it"* in the VR (P15). In summary, the Haptic Turk [1] method that we applied for handing over the basket and for placing the seat worked for some participants perfectly fine, while it failed for others, especially when a mismatch in locations or time between the physical and the virtual world was perceived.

***Coherent & consistent stimuli*:** Real walking to move in VR and using hands for object manipulation is intuitive. However, position and hand tracking errors - using web cam and Leap Motion - occur from time to time, which was sometimes causing a distraction of the authentic VR perception. P18, for example, found a *"slight difference in where real world objects were placed compared to where they were displayed"* distracting, and *"touching the basket was not in sync with the video so it was kind of confusing"* for P9. Moreover, participants observed *"stutters"* (P12), *"glitches"* (P10), and *"a shaking environment"* (P16). Even if we told the participants to consider our apparatus as prototype, they were not able to ignore technical limitations, and their sense of presence decreased if information mismatches occurred.

Discussion

In our study, we explored how different VR setups influence the feeling present in a virtual architectural prototype. In general, participants preferred VR with stimuli similar to our everyday experience when interacting with our physical environment. Accordingly, the 3D virtual experience was more appreciated than the 2D screen and moving in the VR through walking felt more natural and real than using mouse and keyboard. Moreover, sitting down through a similar action was favored versus using a mouse, and using hands for object manipulation felt intuitive as it has the advantage to provide natural haptic feedback about the grasp action as well as about surface properties of objects. Consequently, adding modalities should increase immersion as well as presence.

However, presence only increases if the stimuli set of the different modalities is meaningful and coherent [4, 7, 17]. Interestingly, the mobile 3D setup that involved most modalities led to less immersion and fewer presence ratings than the seated setup with HMD, and using the hands for manipulating a physical object (basket) significantly reduced presence sensation. We identified three aspects that can disturb the coherence and meaningfulness of the set of stimuli given in VR: *object mismatch*, *time mismatch*, and *spatial mismatch*.

***Object mismatch*:** Presence in VR increases through a real action, and its feedback should be appropriate and expected [7]. If expected feedback, which usually is a consequence of what we learned when acting in the physical world, and feedback given by a system differs, the sense of presence will decrease. With other words: feedback provided in VR should be similar to the feedback the physical world is providing us with. For example, virtual rooms should be perceived to have the same size like the real ones. That is important when walking through VR as well as when users touch the real walls when reaching them in the VE. Moreover, virtual objects should have same sizes, shapes, and textures than physical ones provided by a system as passive feedback, like our seat or basket.

***Temporal mismatch*:** Latency is known to create motion sickness in VR [18]. Beyond that, immediacy of control is also important for presence, which could be reduced through a noticeable delay between an action and the resulted effect shown in the VE [4]. Hence, an interactive VR system has to respond to users' actions or changes in position with very little delay. For haptic feedback 50ms of delay is still tolerated [5]. The basket scenario requires that an assistant places the basket in the moment the user is grasping it at the same location and in exactly the same orientation as the basket is displayed in VR. However, such grasping moments may be quite predictable and thus, the assistant can prepare this situation through pre-positioning the basket close to the user's hand; any unpredictable event (grasp correction or hand replacement) most probably will cause a time mismatch. The reason is that the time to react on a situation change would be up to 500ms, which is substantially more than the tolerated delay for haptic feedback of 50ms. In Haptic Turk [1] a similar system limitation was described as: "a real time unpredictable event is difficult to handle in our current system. It does not leave enough time to properly prepare actuators for their action." Real time unpredictable events are challenging. However, such events are essential in human-computer interaction as the human behavior is hardly predictable in less than 50ms. Hence, we propose to think about alternative ways to support haptic feedback in VR, like Dexmo, a mechanical exoskeleton system for motion capturing and low latency force feedback [3].

***Spatial mismatch*:** While participants were positive about the match of the seat's position in the real world and in VR, the basket was not always where participants expected it to be when relying on the virtual position. Due to the temporal mismatch, positioning the basket seemed to be harder. Moreover, placing the seat at the baseboard has only one degree of freedom (DOF), while positioning the handle of the basket in midair has six DOF. Both challenges could also be addressed through using wearable haptic devices, such as Dexmo.

Conclusion

Aiming to better understand sensation of presence in architecture prototype using VR, we conducted an experiment with different setups and interaction scenarios. We discussed situation where immersion, measured as presence, was reduced when introducing certain modalities. We highlight that, beside temporal consistency, also other information have to match, such as object parameters and spatial coordinates, for creating a convincing feeling of presence. With this work, we provide insights and design recommendations applicable for prototyping architecture in VR.

REFERENCES

1. Lung-Pan Cheng, Patrick Lühne, Pedro Lopes, Christoph Sterz, and Patrick Baudisch. 2014. Haptic turk: a motion platform based on people. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 3463–3472. DOI: `http://dx.doi.org/10.1145/2556288.2557101`

2. Scott W Greenwald, Zhangyuan Wang, Markus Funk, and Pattie Maes. 2017. Investigating Social Presence and Communication with Embodied Avatars in Room-Scale Virtual Reality. In *International Conference on Immersive Learning*. Springer, 75–90. DOI: `http://dx.doi.org/10.1007/978-3-319-60633-0_7`

3. Xiaochi Gu, Yifei Zhang, Weize Sun, Yuanzhe Bian, Dao Zhou, and Per Ola Kristensson. 2016. Dexmo: An Inexpensive and Lightweight Mechanical Exoskeleton for Motion Capture and Force Feedback in VR. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1991–1995. DOI: `http://dx.doi.org/10.1145/2858036.2858487`

4. Richard Held and Nathaniel Durlach. 1991. Telepresence, time delay and adaptation. *Pictorial communication in virtual and real environments* (1991), 232–246.

5. Caroline Jay, Mashhuda Glencross, and Roger Hubbold. 2007. Modeling the effects of delayed haptic and visual feedback in a collaborative virtual environment. *ACM Transactions on Computer-Human Interaction (TOCHI)* 14, 2 (2007), 8. DOI: http://dx.doi.org/10.1145/1275511.1275514

6. Maud Marchal, Anatole Lécuyer, Gabriel Cirio, Laurent Bonnet, and Mathieu Emily. 2010. Walking up and down in immersive virtual worlds: Novel interactive techniques based on visual feedback. In *3D User Interfaces (3DUI), 2010 IEEE Symposium on*. IEEE, 19–26. DOI: http://dx.doi.org/10.1109/3DUI.2010.5446238

7. Michael W McGreevy. 1992. The presence of field geologists in Mars-like terrain. *Presence: Teleoperators and virtual environments* 1, 4 (1992), 375–403.

8. Dominik Schmidt, Rob Kovacs, Vikram Mehta, Udayan Umapathi, Sven Köhler, Lung-Pan Cheng, and Patrick Baudisch. 2015. Level-Ups: Motorized Stilts That Simulate Stair Steps in Virtual Reality. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2157–2160. DOI: http://dx.doi.org/10.1145/2702123.2702253

9. Adalberto L Simeone. 2015. Substitutional reality: Towards a research agenda. In *Everyday Virtual Reality (WEVR), 2015 IEEE 1st Workshop on*. IEEE, 19–22. DOI:http://dx.doi.org/10.1109/WEVR.2015.7151690

10. Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364, 1535 (2009), 3549–3557.

11. Mel Slater, Martin Usoh, and Anthony Steed. 1995. Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)* 2, 3 (1995), 201–219. DOI: http://dx.doi.org/10.1145/210079.210084

12. Ivan E. Sutherland. 1968. A Head-mounted Three Dimensional Display. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I (AFIPS '68 (Fall, part I))*. ACM, New York, NY, USA, 757–764. DOI: http://dx.doi.org/10.1145/1476589.1476686

13. Anne M Treisman. 1964. Verbal cues, language, and meaning in selective attention. *The American journal of psychology* (1964), 206–219. DOI: http://dx.doi.org/10.2307/1420127

14. Anne M Treisman and Jenefer G Riley. 1969. Is selective attention selective perception or selective response? A further test. *Journal of Experimental Psychology* 79, 1p1 (1969), 27.

15. Mark Ward, Ronald Azuma, Robert Bennett, Stefan Gottschalk, and Henry Fuchs. 1992. A demonstrated optical tracker with scalable work area for head-mounted display systems. In *Proceedings of the 1992 symposium on Interactive 3D graphics*. ACM, 43–52. DOI: http://dx.doi.org/10.1145/147156.147162

16. ROBERTB WELCH and DAVIDH WARREN. 1986. Intersensory interactions. *Handbook of perception and human performance.* 1 (1986), 25–1.

17. Bob G Witmer and Michael J Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments* 7, 3 (1998), 225–240. DOI: http://dx.doi.org/10.1162/105474698565686

18. Richard Yao, Tom Heath, Aaron Davies, Tom Forsyth, Nate Mitchell, and Perry Hoberman. 2014. Oculus VR Best Practices Guide. *Oculus VR* (2014).