I HEAR, WHAT YOU ARE DOING!

Workspace Awareness in Collaborative Virtual Environments

GERHARD SCHUBERT[1], IVAN BRATOEV[2],
VADYM STRELCHENKO[3] and FRANK PETZOLD[4]
[1,2,3,4] *TU Munich / Chair of Architectural Informatics*
[1,2,3,4] *{schubert|ivan.bratoev|vadym.strelchenko|petzold}@tum.de*

Abstract. Nowadays Virtual Reality application are mostly used for the presentation and discussion of ideas. To enhance the discussion / creative process, participants require the possibility to make changes to the presentation. Similar to the established processes (e.g. model and sketch) a simultaneous editing of several users will occur in the virtual environment, what leads to problems of user comfort. Such issues are explored in the filed of Workspace Awareness. Possibilities for improving workspace awareness for interactive collaboration in virtual worlds were investigated within an interdisciplinary research project between the fields of computer science and architecture. The basis for this research is the development of various test scenarios, taking into account the requirements of the urban planning context of architecture.

Keywords. Virtual Reality; Collaboration; Architectural Design; Awareness.

1. Introduction

The use of virtual reality (VR) for the presentation and communication of architectural planning ideas and results has become widespread. However, this medium is in most cases only used for pure presentation of final thoughts and not for joint collaboration and thus discussion and identification of ideas in early design stages. Particularly in an architectural, planning context, the manipulation of objects, the changing of geometry, position and spatial location is indispensable. Using different tools for editing and presenting leads to a disruption of the design process (Schubert 2014). To overcome these mediabreaks in virtual space would require simultaneous editing of virtual objects for different users. In contrast to the real world, there can be considerable problems and irritations in the communication between the different participants in a virtual context. This occurs in particular when changes are made by one user and need to be visible to one or more other users in real time. Just imagine that you are in VR space and suddenly virtual objects change their position or shape. You would experience an interruption in your creative process and to an extent discomfort in further working in the environment.

Intelligent & Informed, Proceedings of the 24th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2019, Volume 1, 583-592.

The cause for this misconception can be found above all in the reduced sensory perception in virtual space. In real, physical space, several senses are automatically subconsciously involved in perception. Through our sensory organs, especially seeing and hearing, changes in the physical world are perceived unconsciously or incidentally, which directly indicates the activity of the another person. In contrast, in VR this is reduced to a few, individual and often reduced impressions, e.g. missing personal avatars. Furthermore, there is a big difference between the real and the virtual world in the way objects can be manipulated. There are limitations in the real world purely due to physics. On the one hand, objects can only be changed by direct contact. Beyond that, there are limits by the pure size of objects or materials. These limitations do not exist in virtual space. Based on the programmed application, arbitrary objects can be manipulated, an/or be moved independently from most different positions of the user.

Since virtual worlds only provide limited information strong perception problems and irritation can occur. Situations in which a person makes a change and the virtual scenario of another user changes ad-hock are particularly affected.

1.1. AIM

The focus of the presented project is the investigation of this problem and the development and validation of a solution for collaboration. The application of the solution is for the early urban design phases. The aim was, to provide users with direct feedback on changes made by other users in the event of possible changes and thus to improve awareness and understanding in virtual, three-dimensional space.

1.2. PRE-TEST: DISCOMFORT

To confirm the assumption that sudden changes in the VR environment, which are not caused by the user themselves, would cause discomfort for said user, a small user study was done in advance. Participants were placed in a 3D representation of a test-scenario of an urban district. They were tasked with finding a specific landmark of the environment. At random points during the navigation a building would appear instantly in front of the user. Users with previous VR experience described the event as irritating, while those without found it unexpected. A second small user study focused on the issue of how the movement of a building would affect the comfort of the user. New participants were selected for the purpose. They were placed in the same environment, as in the previous test, but at a specific viewing point, where the user could already see the landmark. Their main task was to describe the differences of the digital representation of the landmark to that of the real one (the digital representation was purposefully changed for this test). During this description the landmark would start moving towards the user at 200 meters per second, the initial distance from the viewing spot was approximately 69 meters. When the landmark would stop 0.5 meters in front of the participant the test would be concluded and they where asked about their experience. Half of the users were startled and had a general unpleasant experience, while the other half reported not being negatively affected by the sudden change. In spite of the informal nature of the preliminary prototype tests, the results indicated disagreement among the participants, and the problem was deemed worthy of further exploration.

2. Application Scenarios

Based on the goals and aims of this paper a few areas of application are selected based on their different focus points and the extent to which they implemented collaborative elements. They were observed and analyzed on how they approach the different challenges in collaborative virtual environments.

The CocoVerse is a VR application that utilizes a HMD and controllers. It offers the users very broad options for creative and design thinking through the utilization of a "tool belt" positioned at the waist level of the user. Although the preliminary tests were done in pairs, the focus was on the user interface and actual collaboration was only observed in some pairs, but was not explored further. (Greenwald et al. 2017)

GreenSpace II offers a shared environment for architectural design review. It uses a 6 DoF tracked HMD and controllers, and focuses on the exploration of user experience and interface. Although the work is mostly explorative it observed that audio signals and queues where more critical for the social and collaborative work between users. (Davidson and Campbell 1996)

The work of (Lena S., 2016) explores the real-time collaboration in urban design with the use of a telepresence system. In the setup, a heterogeneous system is used, where one user is in a CAVE system that represents the planned area in a 1:1 Scale virtual environment, and the other user is utilizing a touchscreen interface to make global changes to the environment. The focus of the work is on the different type of interactions a user can perform in the CAVE system.

3. Theoretical Approach

The shared chalkboard application proposed in (Gutwin et al. 2011) utilizes the concept of Workspace Awareness (WA). The work explores how the application of dynamically synthesized chalk sounds in addition to visual assistance (represented as a radar) in a distributed groupware would improve the WA of users. The study highlighted empiric evidence, that the addition of sound to such environments greatly improves the awareness of participants.

Based on the presented problem and in respect of the application scenarios the focus of this research project is the improvement of collaboration in virtual environments. Key for these improvements is the Workspace Awareness that builds upon the basic principles of Situation Awareness developed by Endsley (1998). WA is defined as "the up-to-the-moment understanding of another person's interaction with the shared workspace" (Gutwin and Greenberg 2002). It not only focuses on a single user in an environment but on a multitude of participants performing similar tasks affecting the same environment. While this is a trivial task in real-world face-to-face collaboration, Workspace Awareness becomes a bottleneck for productive collaboration in VR because of the limited information that systems provide about their state changes, and their unfamiliar interaction interfaces. Workspace awareness is an important point when you focus on “real” collaboration in virtual environments. By the term “real” we mean, that people can meet in VR and for example change the position and shape of virtual objects like building blocks. Gutwin and Greenberg (2002) introduced the WA framework, which provides a toolset for describing and analyzing medium-sized workspaces. In 2011 they proposed the shared chalkboard application which utilizes the concept of Workspace Awareness (WA)(Gutwin et al. 2011). The work explores how the application of dynamically synthesized chalk sounds in addition to visual assistance (represented as a radar) in a distributed groupware would improve the WA of users. The study highlighted empiric evidence, that the addition of sound to such environments greatly improves the awareness of participants. The framework is applicable to real-time distributed groupware, and small mixed-focus collaboration groups performing generation and execution tasks. The WA framework is split in three parts:

- Information: the required information that collaborators might require from the workspace. This allows for the creation of a fixed vocabulary with which to describe the system.
- Communication: focuses on ways, in which the information can be delivered to the users. Three proposed ways for this communication are through “bodies and consequential communication”, “artifacts and feedthrough”, and “conversation, gesture, and intentional communication”.
- Use: focuses on how a user would utilize the obtained information from the previous two parts. This allows designing the workspace in such a way that only the relevant information is communicated in the relevant time points.

Even if these investigations refer to the two-dimensional space, it can be assumed that a transfer of the methods into the three-dimensional workspace is possible. Within this project, methods for improving workspace awareness in 3D spaces were developed and investigated. The aim was, to provide users with direct feedback on changes made by other users (Represented by the Communication part of WA) and thus to improve awareness and understanding in three-dimensional space.

3.1. SONIFICATION

Sound cues were chosen as an awareness presentation, because they, speaking in Workspace Awareness terms (Gutwin and Greenberg 2002), can: promote perception of the elements in the environment, comprehension of their meaning, and allow projection (or prediction) of their state into the future; all this while not overloading the visual channel with additional information.

In this context, the relevance of the two major sensory organs of seeing and hearing is briefly discussed. Both senses have relevant differences in their role in human perception. The visual sense is clearly the dominant sensory organ in comparison to all other. In this way ca. 80% of all Information is perceived through sight (Weidlich und Trost 1995). In contrast to this, the role of the ear can be understood as a guardian: “The ear is always open, ready to receive information from all sides and, if necessary, to warn the organism of dangers, even when it is asleep” (Hellbrück und Ellermeier 2004).

On the basis of this, a potential way of implementing the Communication aspect of WA is achieved through the Sonification of changes to the design environment. This type of Sonification is represented using earcons, the aural analogy of a visual icon in modern day operating systems. (M. Blattner et. al 1989) divides them into three main groups:

- Representation earcons rely on digitalization of the natural sounds and their mainly metaphoric use in computer systems.
- Abstract earcons rely on the notion of "elements" that can be combined into structures to explain the relationship between some actions.
- Semi-abstract earcons are a combination of the previous two classes.

Gaver (1989) argues that the representation earcons bring implicit benefits by exploiting the natural way humans react to sound, (M. Blattner et al. 1989) shows that this type of earcons works only with few distinct sounds, since each sound would require a unique audio representation. Abstract earcons on the other hand would require far less, but the user would need to familiarize themselves with the different sounds.

3.2. PRE-TEST: SONIFICATION

To establish the correct type of earcons for a virtual environment within an architectural context a small user study was performed. The goal of this study was to determine not only the speed at which buildings should move through the environment (slow 20m/s, medium 40m/s, 60 m/s), but also the type of earcon that should be used: abstract, represented by a sine wave, or representational, using a concrete on concrete sound. Eight participants were selected for the test (7 male, 1 female).

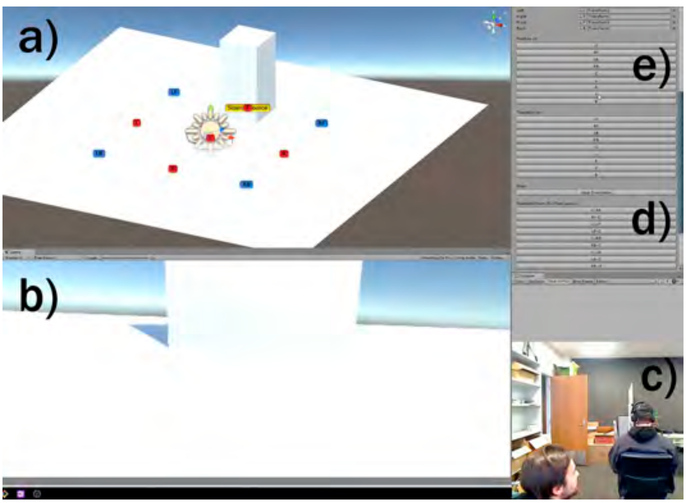

Figure 1. Experiment setup: a) VE, b) participant's perspective; c) participant and experimenter; d) predefined translation paths; e) arbitrary translation panels.

A participant would be placed at the Center (C) of the environment. A real-sized building, approximated by a cuboid, was present in the environment. It could translate from and to any of the predefined positions: Left Front (LF), Right Front (RF), Left Back (LB), Right Back (RB), Front (F), Back (B), Left (L), Right (R), C. When the building moved, it would emit sound. The building would be initially placed in front of the user. Participants were put through a tutorial to build a correlation between the visually perceived movements of the building and the emitted sound: the building was translated along the major directions (C-F, C-L, L-R, L-C, C-B), and then additionally between randomly selected positions from the predefined set.

At the start of the actual experiment, participants were asked to close their eyes and guess the path that the building traversed. It was silently positioned at a randomly selected position and then translated, with audio feedback turned on, to a newly selected random position (Fig. 1 e)). This was repeated 5 to 10 times for each of 6 type of sound-speed combinations. After going through all the sound types and speeds, participants were asked, as to which sound was the best, and what speed was the most appropriate. 6 out of 8 participants preferred the concrete sound to the sine wave, 1 was indifferent as long as the speed was slow, and 1 preferred the sine wave. Figure 2 shows that the precise preferred speed was 30 m/s.

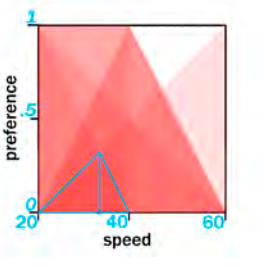

Figure 2. Preferred speed according to analysis.

4. Main Setup and Test Procedure

The main user study investigates the application of sound feedback for the improvement of WA in a virtual environment. It is developed utilizing the results of the previous user tests. The user is positioned in the center of the virtual environment, where they can move around and have to perform a primary task. In this test scenario they are presented with different outlines of objects and are tasked with filling these silhouettes out (Fig 3. a)). This mimics a typical design activity of architects in the early design stages of sculpting their models. Representing the activities of a second user, that is exploring the environment and is performing similar design tasks, a random building in the environment is selected and moved to another arbitrary location around the original user at irregular time intervals. The users, performing the primary task, are given a secondary task to locate and catch the moving building.

The user tests was implemented with the help of Unity3d 2018.1.5f1. The Resonance Audio SDK for Unity version 121 was used for the audio spatialization. The hardware for the test was the HTC Vive HMD, 6 DoF controllers, on-ear stereo headphones and a Windows 10 computer.

The virtual environment is a 1:1 Scale of an urban district (Fig 3. b)). The user can only move physically within the limitations of the actual tracked space. One Controller serves as a 3D voxel brush that draws when the trigger is pressed, while the other controller is used as a pointer and in two of the three test scenarios as an anchor for a minimap of the district.

There were 8 male and 4 female participants in total, within the age group of 22 to 32 years. One of the participants reported having "partially" good hearing, while only 3 where familiar with VR technologies and only one had a background in architecture. Participants were given a written and verbal introduction to the experiment. Next, they were placed in the immersive VR environment, which implemented 2 modes: tutorial and experiment.

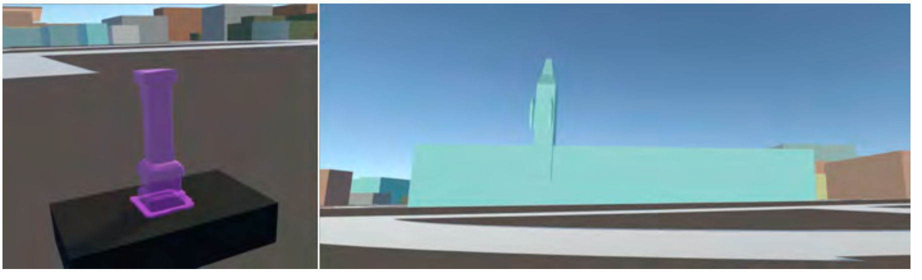

Figure 3. Virtual Environment a) the outline shape; b) the buildings.

In the tutorial mode, only one building and a test shape for filling out would be in the environment. Participants learned how to use the 3D voxel brush for tracing the shapes, the laser pointer to indicate they noticed changes to the environment, and the minimap to locate the moving buildings. They also familiarized themselves with the spatial sound of the building moving.

In the experiment mode, the participants were presented with three scenarios: "Minimap and Sound", "Sound Only", and "Minimap Only". The only difference between them was how the participant was informed of the moving building in the virtual environment. The participants had to focus on the main task of filling out the different shapes with the 3D brush and not on the moving building. When a participant was happy with the tracing of the current shape, they would voice this to the experimenter, and he would save the current trace and load the next outline. During this buildings would occasionally move through the environment and the participants were observed how well they would recognize when and where this was happening. At the end of the user test each participant was given a self-report questionnaire.

5. Evaluation

This user test followed the repeated measures design with one independent variable - the type of awareness presentation, represented by the three scenarios. Each scenario was tested for ten minutes, during which exactly eight buildings were randomly chosen by the application and moved. Due to some technical issues with the audio playback for the scenario "Minimap and Sound" and "Sound Only", the first building movement in each of the three scenarios was disregarded. 24 data points in total were captured per participant, but only 21 are used for the analysis due to the aforementioned technical issue. The collected data were the reaction speed in determining the position of a moving building ("catching"), along with the 3D drawings created by the main task. If the participant failed to catch a moving building, the reaction time was recorded as 0. 12 shapes for the experimental mode and 1 for the tutorial were prepared. Each participant went through the shapes at their own pace and had no limitation on how much time they can spend per outline. The order of the shapes during the experimental mode was consistent among participants.

Table 1. Statistical Summary.

	Scenario 1: Minimap and Sound	Scenario 2: Sound Only	Scenario 3: Minimap Only
Mean time to catch building (sec.)	4.662725	4.700911	6.61726
Standart deviation (sec.)	2.266442	2.483499	2.73528
Range (sec.)	[1.865601, 11.195500]	[1.567505,13.192500]	[1.890381, 14.635010]
Caught Cout	73	65	55

Table 1 displays the performance of all participants during the user test and the total mean time to catch buildings. While the mean difference between Scenario 1 and 2 is minimal (4.66 seconds versus 4.70), there is a drastic drop in caught buildings, when only sound was used as indicator of a moving building and even a further drop, when only the minimap was used. The proportion of caught / missed buildings can be more clearly seen in Figure 4:

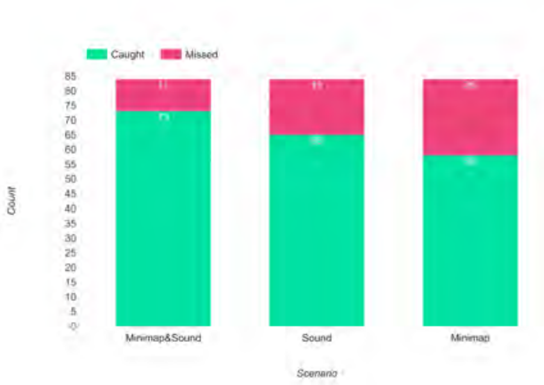

Figure 4. Caught and missed translations.

The implementation of the voxel drawing system was not highly optimized for performance and as a result, the frame rate dropped to approximately 70 frames per second. One participant indicated suffering from a slight motion sickness during one such instance.

The user test clearly shows, that providing the user with interface tools (the Minimap) and enhancing their senses (the Earcons) greatly improves their workspace awareness and leads to increase in comfort when working in a collaborative virtual environment, where not all participants are working on the same tasks. This research provides a guideline on how a collaboration in VR could be performed in an architectural context.

6. Summary

The paper presented here deals with the question of possible collaborations in virtual space in an urban planning context. Based on the problem of a surprise effect in sudden changes of geometries, different attempts to improve workspace awareness were investigated. Our work is based on the hypothesis that additional auditory cues promote workspace awareness in collaborative virtual environments. Based on this, several test arrangements and test series were defined and performed. Namely, sound cues can convey: the nature of an event (e.g. sound of concrete sliding on concrete for moving buildings), the position where an event occurred, the directional change of the source of the event, and the speed of its positional change. Additionally, sound is known to have alarming quality, and helps inform the user about the fact that something is happening. The evaluation of these test scenarios has clearly shown the potential of the sonification of actions in virtual space. In summary, it can be said that audio cues significantly improve workspace awareness in VR.

Figure 5. LEFT: Discomfort and irritation due to unexpected changes in the virtual space. RIGHT: Soundcues improve the WA.

This was proven in the context of the project on the one hand by the subjective opinion of the test persons: The background of changes by sound cues with the simultaneous display of a mini map was chosen as the favourite by far. In addition, the measurement of the reaction time to a change and the recognition and localisation of this in the VR room confirmed the subjective statement of the test subjects. Here, too, Minimap + Sound Cues improved the response time as well as the localization in the room compared to the other variants without Sound Cues.

Applied to the application cases, it can therefore be assumed that in collaborative, creative work in virtual space, the use of acoustic signals greatly reduces the irritation of other users. As the experiments have shown, the time spans as well as the speed of movement are of importance. Even if the improvement potential through sound cues could be proven within the scope of the project, the investigation will be deepened in further investigations. The focus is on perceived importance, annoyance and intrusion, as well as better visual displays of users (e.g. inter- actions, user representation, presence, etc.).

Acknowledgement

This study was performed in cooperation with "Human Computer Interaction" Research Group at the University of Otago, New Zealand. Many thanks to Professor Regenbrecht and his research team.

References

Blattner, M.M., Sumikawa, D.A. and Greenberg, R.M.: 1989, Earcons and Icons: Their Structure and Common Design Principles, *ACM SIGCHI Bulletin*, **21**(1), 123-124.

Davidson, J. and Campbell, D.: 1996, Collaborative Design in Virtual Space - GreenSpace II: a Shared Environment for Architectural Design Review, *ACADIA Conference Proceedings*, 165-179.

Endsley, M.R.: 1988, Design and Evaluation for Situation Awareness Enchancement, *Proceedings of the Human Factors Society Annual Meeting*, 97-101.

Gaver, W.W.: 1989, The SonicFinder: An Interface that Uses Auditory Icons, *ACM SIGCHI Bulletin*, **21**(1), 124.

Greenwald, S.W., Corning, W. and Pattie, M.: 2017, Multi-User Framework for Collabortaion and Co-Creation in Virtual Reality, *12th International Confference of Computer Supported Collaborative Learning (CSCL)*.

Gutwin, C. and Greenberg, S.: 2002, A Descriptive Framework of Workspace Awareness for Real-Time Groupware, *Computer Supported Cooperative Work (CSCW)*, **11**(3-4), 411-446.

Gutwin, C., Schneider, O., Xio, R. and Brewster, S.: 2011, Chalk sounds: the effects of dynamic synthesized audio on workspace awareness in distributed groupware, *CSCW '11: Proceedings of the ACM 2011 conference on Computer supported cooperative work*, 85-94.

Hellbrück, J. and Ellermeier, W.: 2004, *Hören: Physiologie, Psychologie und Pathologie*, Hogrefe, Verl. für Psychologie, Göttingen [u.a.].

Schubert, G.: 2014, *Interaction forms for digital design: a concept and prototype for a computer-aided design platform for urban architectural design scenarios*, Ph.D. Thesis, Technical University of Munich.

Streppel, L.: 2016, *Real-time Collaboration over Distance: using Motion Capture and VR Interaction in a Telepresence System*, Master's Thesis, Ludwig Maximilian University of Munich.

Weidlich, R. and Trost, M.: 1995, Sehen und Bildschirmarbeit aus augenärztlicher Sicht, *Sehen und Bildschirmarbeit*, Bremerhaven, 70-75.