Improving Co-located Collaboration with Show-Through Techniques
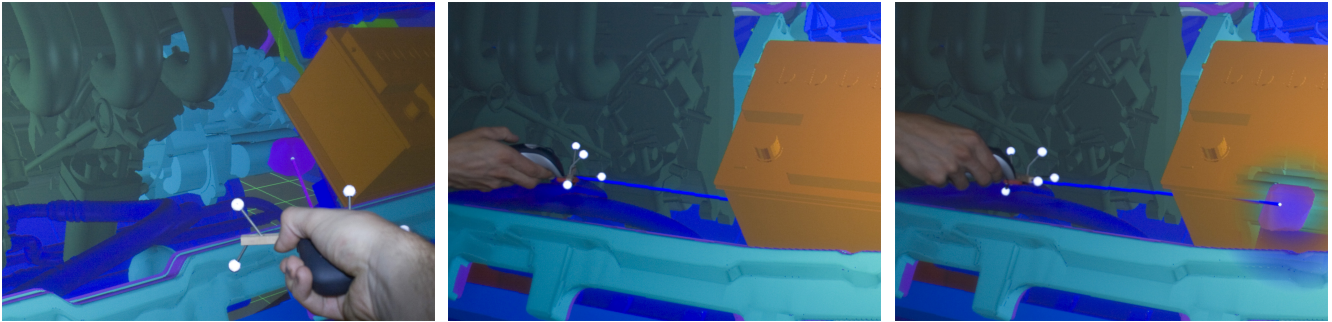
Category: Research

Figure 1: The first two images (a, b) illustrate the issue of interpersonal occlusion between two tracked users in a multi-user VR system: an object that is fully visible to one user (a) can not or only partially be seen from other viewpoints (b). Show-through techniques can improve target discovery in such situations by showing the indicated object through the occluding environment (c).

ABSTRACT

Multi-user virtual reality systems enable natural interaction with shared virtual worlds. Users can talk to each other, gesture and point into the virtual scenery as if it were real. As in reality, referring to objects by pointing, results often in a situation whereon objects are occluded from the other users' viewpoints. While in reality this problem can only be solved by adapting the viewing position, specialized individual views of the shared virtual scene enable various other solutions. As one such solution we propose show-through techniques to make sure that the objects one is pointing to can be seen by others.

We analyzed the influence of such augmented viewing techniques on the spatial understanding of the scene, the rapidity of mutual information exchange as well as the social behavior of users. The results of our user study revealed that show-through techniques support spatial understanding on a similar level as walking around to achieve a non-occluded view of specified objects. However, advantages in terms of comfort, user acceptance and compliance to social protocols could be shown, which suggest that virtual reality techniques can in fact be better than 3D reality.

Index Terms: H.5.2 [User Interfaces]: Input devices and strategies—Evaluation/methodology;

1 INTRODUCTION

Multi-view virtual reality systems enable the collaborative experience of a shared 3D space in a similar way as in real life. Consequently, natural forms of gestural communication, such as pointing, can immediately be used for the collaborative inspection of computer-generated 3D models. However, real-world correspondence also results in real-world problems. One might want to show a virtual object to colleagues, which may be occluded from their respective viewpoints. To solve this problem in reality, people have to walk around the occluding objects to obtain a suitable viewing position. Often they move close to the person who is pointing in order to see the specified object (e. g. by looking over his shoulder). This behavior can result in physical proximity of users that does not comply with social protocols of formal presentations.

Sophisticated multi-user virtual reality systems provide the possibility to generate user-dependent individual views of the shared 3D scene (e. g. [1]). Building on such display technology we can equip observing users with augmented viewing capabilities to allow them to look through objects if another person points at an occluded target. In particular, cut-away views and x-ray vision are promising candidates, considering that they can be used to reveal hidden objects while maintaining most of the local 3D context. We refer to these novel type of collaborative interaction techniques as "show-through" techniques, since the objects being pointed at are showing through occluding objects.

We assumed that users would greatly appreciate show-through techniques and also that they would maintain larger distances to each other in order to feel more comfortable. On the other hand, we were concerned that the understanding of spatial relations in the presented 3D scene would suffer when users can see specified objects without being actively involved in their visual discovery. Though objects can be identified more rapidly if occluding objects vanish automatically, it is likely to be more difficult to understand and memorize their exact 3D location in a complex environment. In addition, most humans are not experienced with such "magical" viewing capabilities. However, our user study revealed that users become rapidly proficient with augmented vision, have a similar spatial understanding of a 3D model as when required to move around and that they make use of the show-through techniques to adhere to well-established social protocols.

2 RELATED WORK

2.1 Co-Located Multi-User Virtual Reality

Multi-user Virtual Reality can enhance collaborative work, in scenarios such as presentations and joint reviews of architectural and mechanical design. Multi-user applications require that the user interface becomes transparent not only to one user, but to multiple collaborators. As such, the interface must not constitute an obstacle to the direct communication between them. Unfortunately, most virtual reality display systems do not support multiple co-located users appropriately. Head-mounted displays strongly compromise the users visual perception of others and the operational environment. Projection-based systems, on the other hand, are commonly limited to displaying only one stereoscopic view. Thus, just one user can see the virtual world perspectively correct, while surrounding users perceive distorted views. It is impossible in such settings to communicate locations of interest through natural pointing.

Nevertheless, specifically projection-based VR systems are often

used by groups of people rather than a single person since their spatial layout affords collaboration [2, 21]. To meet the requirements of group interaction in single view VR setups Simon proposed rendering the scene for a fixed viewing position to be unaffected by the users' head movement [21]. He argues that basic manipulation techniques can still be well supported in such a setting if the virtual interaction tools (pointing ray, menus, etc) are rendered perspectively correct for the operating user. This workaround tries to cope with the limitations of single-view projection systems, but it implies that all users perceive a distorted view on the actual content of the scene. The approach also inhibits natural viewpoint maneuvering in terms of head movement. We argue that co-located collaboration in virtual environments requires to render perspectively correct views of the scene individually for all involved users. Only true multi-view setups enable the user to build on interaction and communication skills gathered in the real world.

Head-Mounted Displays (HMDs) do provide individual views for each user but most of them also impede the perception of the physical environment including one's own body and that of the others. This complicates direct interaction and largely inhibits gestural communications between team members. If one cannot perceive the physical actions of the others, coordinated activities are hardly possible. Elaborated avatar representations of the users can improve this situation [19], but this approach cannot compete with the quality of nonverbal communication in a shared physical reality. See-through HMDs allow to superimpose (or augment) the visual perception of the real world with computer generated graphics. Such augmented reality (AR) displays allow collaborative interaction with 3D computer graphics in a natural way, e.g., direct pointing gestures [3, 17] Though, the quality of the visual perception of both, the real and the virtual environment can only be traded against each other in that case.

More recently, several other concepts to provide a shared 3D virtual reality to multiple users have been proposed. Advances in the development of autostereoscopic (see [10] for an overview), volumetric [16, 14] as well as holographic displays [18] promise highly realistic presentations of 3D geometry that multiple users can observe from their respective viewpoints without the use of any special headgear. In our studies, we followed the approach of time-multiplexing the output from multiple video projectors using high frequency LC-shutter technology [13]. Projection-based multi-user VR systems require the users to wear shutter glasses, but compared to HMDs, the impediments of interface hardware to immediate gestural communication between users is minimized.

Salzman et al. [20] demonstrated the feasibility of bare-handed pointing and tracing to guide the attention of others in a comparable setup. They also found that pointing at an equivalent physical model is still more accurate – particularly for "touching" virtual objects in proximity. This can be explained with the lack of haptic feedback and different focal distances of the real finger and virtual objects which are stereoscopically displayed at the same depth. In order to compensate this, they suggest displaying the pointing direction with a virtual ray if high accuracy is required.

In the collaborative exploration and inspection of complex 3D scenes, we observed that objects and features are not always visible from all involved viewpoints. The referential awareness of other users can be compromised due to occlusion. Consider the set of visible objects for one user V_1 and the set of visible objects for a second user V_2. Objects visible for both users, $V_1 \bigcap V_2$, do not pose a communication problem. However, one user may point at an object that cannot be seen from the others's viewpoint $V_1 - V_2$ or $V_2 - V_1$.

This issue has already been observed by Chastine and Zhu in the context of video-based tele-collaboration [5]. They suggest occlusion handling techniques to cope with it. However, Chastine and Zhu do not explicitly analyze the suitability of different techniques

nor do they consider the corresponding impact on mutual communication and collaboration.

In our scenario, users may also alleviate the occlusion problem by walking over to the person who indicates a point of interest. Just as in reality, one may look over the shoulder of an other to see what he/she is referring to. However, this may interfere with social protocols. We assumed that being in such close proximity to one another could result in the users feeling uncomfortable, and that occlusion-handling techniques could help to avoid that. Thus, we were interested in the mutual distances of users during our experiments in comparison to typical behavior in less constrained situations.

2.2 Proxemics

The extent of open space a person needs surrounding him or her depends on several factors, including physical ones such as temperature, lighting and noise. But one's culture and social relations are also encoded in terms of space. Edward T. Hall coined the term proxemics [15] "for the interrelated observations and theories of man's use of space as a specialized elaboration of culture." Proxemics deal with social and personal space and man's perception of it.

In "The Hidden Dimension" [15], Hall defines four classes of interpersonal distances and their respective meaning in regards to social relations. All four are further subdivided into their respective close and far phases: *intimate distance* (0 cm – 5 cm – 46 cm); *personal distance* (46 cm – 76 cm – 120 cm); *social distance* (1.2 m – 2.1 m – 3.7 m); and *public distance* (3.7 m – 7.6 m – more). Hall derived these measurements from experiments and interviews he conducted with adult natives from the northeastern seaboard of the US. Since they mainly depend on culture and perceptional channels involved in the communication among one another, they also serve as rough approximations for interpersonal distances in other regions of western culture.

Intimate distance involves touch and olfaction. Visual perception becomes distorted. It occurs only in certain situations such as "wrestling or making love". Otherwise, people generally try to avoid such closeness. *Personal distance* is mostly observed between family members, partners and close friends. The distinction between the close and the far phase of personal distance corresponds to one's arm reach at approximately 76 cm. In more formal relations among acquaintances (e. g. in collaborative work settings), people tend to maintain a *social distance*. Hall describes the close phase of *social distance* (up to 2.1 m) as the generally observed situation in collaborative work settings. He also notes that important people often try to keep their subordinates at this far phase. *Public distance* can be observed in settings like public speeches. It is not adequate for most kinds of collaboration and therefore, not relevant to our studies.

3 FACILITATING MULTI-USER POINTING WITH SHOW-THROUGH TECHNIQUES

Whenever a group of people is inspecting a complex 3D scene, interpersonal occlusion is a frequently occurring issue. We propose *show-through* techniques to improve the gestural communication of people dealing with such problems in multi-user VR setups. As an exemplary application, we refer to the design review of a car's engine compartment. Using the design space proposed by Elmqvist and Tsigas [12], we defined the constraints for appropriate techniques to handle interpersonal occlusion issues:

- **Primary Purpose** : We aim to improve gestural communication between users by facilitating the discovery of occluded objects if one is trying to show them to others.

- **Disambiguation Strength** : We want to support occlusion handling for objects that are strongly interacting with other

objects in the virtual environment. This includes enabling the visibility of objects that are enclosed in others.

- **Depth Cues** : Depth cues ought to be maintained as far as possible in order to best support the users' spatial perception.

- **View Paradigm** : Since our primary purpose is to improve interpersonal communication, the notion of a shared 3D environment must not be destroyed. Furthermore, we argue that target objects ought to be seen embedded in the shared 3D space to make sense of pointing gestures and to better understand the spatial layout of the scene. Thus, we do not consider solving the occlusion issue by providing multiple viewports to individual users.

- **Interaction Model** : Techniques for handling interpersonal occlusions must implicitly be triggered by a user pointing at certain objects in the scene. However, the resulting changes are applied to the view of others. Thus, techniques to handle interpersonal occlusions can be considered as *actively* triggered by one user (but without receiving visual feedback), while *passively* perceived by others.

- **Target Invariances** : For the collaborative inspection of 3D graphics it is generally very important to *preserve the location, the geometry and the appearance* of target objects. Otherwise users would perceive different objects and, thus, communication regarding the presented 3D environment would be hindered.

Regarding the properties of different occlusion management techniques, we found that the *Virtual X-Ray* fulfills nearly all of our requirements better than other techniques. Only in terms of provided depth cues, some alternatives may offer better performance. However, those techniques, namely *Multiple Viewports, Tour Planner and Interactive Exploder*, have severe drawbacks with respect to the other dimensions in the design space of occlusion management techniques.

Multiple Viewports can provide an additional non-occluded view at a target object, but this approach does not support its correct localization, which is important for the collaborative inspection of a 3D scene. *Tour Planner* refers to techniques that move the virtual viewpoint of users in order to gather a non-occluded vision at target objects. This approach is not adequate for multi-user VR applications considering that it would destroy the notion of a shared 3D environment. Using *Interactive Exploders*, users can move occluding objects aside, but this requires quite a bit of effort. It takes time to manipulate objects and as the layout of the scene is changed, users loose important information about the relative placement of objects.

Virtual X-Ray techniques can solve visibility issues in that occluding objects are removed or rendered transparently [4, 6, 9, 11, 8]. In contrast to other approaches, such "see-through" techniques can solve occlusion issues while preserving the layout of the scene as well as the coherence of a virtual environment which is shared by multiple users. To facilitate collaborative inspections of a virtual 3D scene, *Virtual X-Ray* techniques are applied to ensure that for all other users the object is showing through that one is attempting to refer to. We therefore suggest using the term "show-through" techniques for this particular application of the *Virtual X-Ray*.
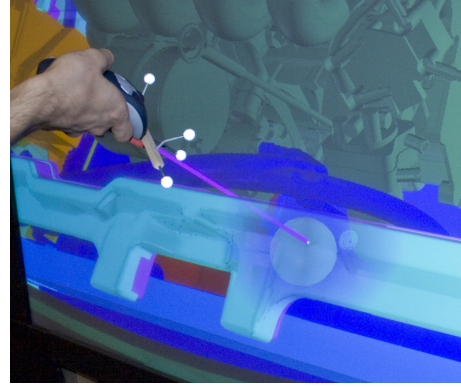
3.1 Implementation

We implemented two show-through techniques (see Figure 2) with different visualization parameters: *Cutaway* and *Transparency*.

In both cases, we define a cutting volume stretching from a user's viewing position to the indicated parts of the scene. All fragments of graphics objects that fall within that volume are either removed completely or displayed semi-transparently.
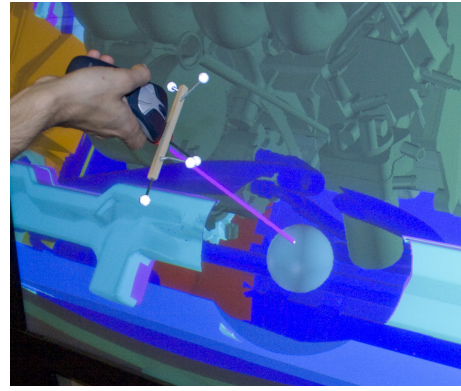
(a) The *Cutaway* technique.

(b) The *Transparency* technique.

Figure 2: Show-through techniques can improve target discovery by completely or partially removing the occluding environment.

The implementation of both show-through techniques is largely similar. First, a volume containing the selected object as well as both eyes of the user, is computed. For simplification, we constrained the volume to the smallest possible cylindrical shape. In our study, we were only considering the mutual presentation of simple, sphere-like objects. The axis of the cylinder is defined by the center of the object and the center between the user's eyes. For the description of a more general approach refer to previous work of Burns and Finkelstein [4].

Once the cutting volume is defined, the fragments falling inside that area can be determined and their appearance can be transformed accordingly. For the *Cutaway* approach, all fragments inside the cutting volume are discarded using a fragment shader. For the *Transparency* approach, alpha values of each fragment are modified, taking into account the distance to the axis of the cylinder. Occluding objects were only sorted in depth on a per-object basis. For the engine compartment of our car model, the transparency effect was correct for most viewpoints. In a different scenario, however, sorting on a per-triangle basis might be required [11].

The lightweight implementation of the proposed show-through techniques introduced a negligible impact on rendering performance. Our sample application bottleneck was mainly the rendering of the engine model, the application had to generate two stereoscopic image pairs. In average it was able to provide an update rate of 40 fps.

4 USER STUDY

We implemented a two-user pointing task to analyze the usability aspects of the proposed show-through techniques (see section 3). The experimental task was designed to investigate user performance regarding the identification of indicated targets in a dense 3D environment and the memorization of their respective locations. Both show-through *techniques* (*Cutaway* and *Transparency*) were compared to a baseline condition (*None*) in which users had to obtain adequate viewpoints by walking around to see the otherwise occluded objects. In addition to these pure performance measures, we studied the impact of our visualization techniques on the users' behavior in such a collaborative work setting. We were interested to see whether users could benefit from the proposed show-through techniques in terms of maintaining more comfortable distances to each other without decreasing the efficiency of their collaborative interaction.

The experimental task we implemented for our studies refers to a collaborative design review in the automotive industry, where variations of a design are evaluated by a group of experts. Showing certain features of the model to colleagues is a frequently occurring subtask in such collaborative work settings. Multi-user VR systems are a promising technology to facilitate this as they enable immediate information exchange about features in a shared 3D environment. Following the situation of a design review as a reference, we presented the engine compartment of a VW Golf in its original size on a multi-view projection screen (see Figure 4). Our experimental system supported two tracked users that could individually walk around the virtual model to observe it from different viewpoints. The perceived stereoscopic view was always corresponding to the users' respective viewing position. Thus, the model was perceived to remain at a fixed location in the shared environment while users were walking around it. The 3D model was tilted about 40° in order to provide users with a good overview that is similar to the visual experience in the real world where people look from above at a car's engine compartment.
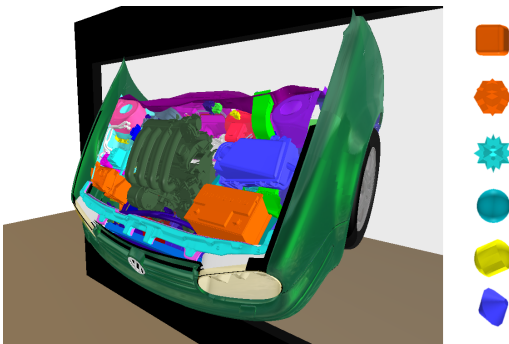

Figure 3: Left, engine model used for the user study, the black frame represents the borders of the real screen. Right, during the test participants had to discover and point to six synthetic targets (not in same scale as the engine model).

We assigned different tasks to the two involved users of our multi-user system: the experimenter (hereinafter referred to as the *presenter*) was pointing to certain objects in the model that had to be identified and located by a second user, the *observer*. The role of the *presenter* was always assumed by one of the authors. For the role of the observer, we invited volunteers to analyze how show-through techniques would affect their ability to identify, locate and memorize the indicated objects and also their behavior in terms of proxemics (see section 2.2). Abstract 3D geometries that were placed at 15 predefined positions in the model served as target

objects for presentation and identification (see Figure 3). Therefore, all target objects were of the same size. We decided to use abstract target objects that had no semantic relation to the model in order to to minimize the potential bias from different levels of knowledge users may have about the structure of a car's engine compartment. Only one target object was shown at a time to ensure that all experimental conditions were evaluated using the same sequences of target objects. Two different investigators were alternately running the experiments. We made sure to fulfill the role of the *presenter* in compliance with a strictly defined storyboard to minimize the influence of personality. Additionally, we prevented that the *presenter* would have to search the next target object in line by providing a visual localization aid. The high quality of the shutter elements we used in our setup effectively prevented these hints being seen by the *observer*.

Correspondingly to the different roles of the *presenter* and the *observer*, each *trial* also consisted of two phases. During the presentation phase, the *presenter* was pointing to three target objects in a row, while always waiting until the *observer* confirmed that he had located the respective object in the scene. In the following retrieval phase, the *observer* was asked to prove that he had memorized the three objects and their location by pointing to them in the same sequence as they were presented. In both show-through conditions, *X-Ray* vision facilitated the identification of presented objects for the observer, but only during the presentation phase. In the following retrieval phase, users had to rely on the information they gathered in the presentation phase to once again find the three objects in the scene.

During the experiments, different types of data were recorded. Primarily, we logged discovery time during the presentation phase and the time required for the retrieval of target objects. The logged discovery time did not include the time required by the *presenter* to approach the next target object. Instead, only the time that the presenter was pointing at it until the *observer* confirmed that he could localize it, was recorded. Additionally, we logged the users' (head) positions throughout the experiments. This enabled us to sum up the *covered distance* of the participants during the presentation phase as well as to track the distance both users maintained from each other during the experiment.

4.1 Experimental Setup

The experiments were conducted on a projection based two-user VR setup (see Figure 4). The system is built on time-multiplexing of individual views using LC-shutter technology as described in [13]. The users were required to wear tracked shutter glasses. We used custom made double-cell shutter glasses sizing 65 mm in width and 45 mm in height for the shutter glasses. Square 80 mm shutters were mounted in front of the projectors. While single LC-shutter elements provide a contrast ratio ranging between 1000:1 and 5000:1 between the open and closed state, we achieved a contrast ratio of 25000:1 with our double-cell approach. As a result, absolutely no crosstalk between the views of both users was perceptible.

The physical dimensions of the projection screen were 3 m in width and 2 m in height with a resolution of 1800 px × 1200 px. The tracked workspace, wherein people could walk around to examine and gather information about the scenery from different angles, covered an area of approximately 4 m by 4 m. Note that these spatial constraints of the operational environment also affect the users' behavior in terms of proxemics.

The visual stimuli were presented at a stereo depth range of ±0.8 m from the projection screen (see Figure 3). The target objects inside the simulated engine compartment were situated in places which were hidden from most possible viewing positions (e. g. behind the engine block). Without using show-through techniques, the observer could not see them.

4

Figure 4: Our projection-based two-user setup displayed superimposed stereo perspectives on the screen.

4.2 Participants

Twenty-four paid users, aged between 19 and 31, participated in the study. All of them were students of varying disciplines ranging from engineering to computer science and to design and humanities. Seventeen participants had prior experience with VR applications while 7 did not. We organized the experiments as a competition so that the three participants with the shortest retrieval times of targets, won a ticket to the movies. Note that the differences regarding social hierarchies between presenter (research assistant) and participant (students), were not that distinctive as they certainly are within formal presentations held for an executive committee.

4.3 Design and Hypotheses

Each participant of our study was involved in two successive experiments. First, we compared the target retrieval performance and the learning progress of user groups that were operating with different occlusion management *techniques* (*None, Cutaway, Transparency*). Thereafter, we introduced each user to the two other conditions they had not been exposed to during the first experiment. Another set of trials was performed with each *technique* condition in order to analyze user behavior with different *techniques* and to provide a basis for the users' subjective ratings.

We expected stronger effects of learning the task during subsequent *blocks* than that of *techniques*. Therefore, *technique* was compared between-subjects with repetitive *blocks* as a within-subjects variable. Each user group consisted of 8 people. Before the experiment started, users were given written descriptions of the task including the advice to strive for the best possible target retrieval performance, and the hint that the sequence of target object was repeated during each *block*. We provided a short training session to make sure users understood the task, the respective occlusion management *technique* they were assigned to and also that they could effectively handle the involved interaction techniques (e. g. head tracking, ray selection). We trained the interaction procedure of our experiment using a special sequence of six target objects which did not occur during the recorded trials.

The following recorded *trials* also consisted of the presentation of three subsequent target objects that had to be found by the *observer*. Five different *trials*, with unique object locations, were performed during one *block*. The combination resulted in 15 different target locations inside the engine compartment which were repeated in three subsequent *blocks*. In total, 15 *trials* were per-

formed by each user. Participants were encouraged to take breaks between *blocks* to minimize fatigue.

The dependent variables for the first experiment were: the *discovery time* (time during which targets were shown by the *presenter* until the confirmation that the *observer* localized them); the *retrieval time* (duration of the retrieval phase); and the *covered distance* during the discovery phase. The independent variables were the occlusion management *technique* and *block*.

We expected users to learn the task sequence rapidly, and thereby successively optimizing the effectiveness of their operations. We also assumed that they needed to move less when show-through *techniques* were applied. We reasoned that this could enhance the discovery phase by reducing *discovery times*. However, the reduced movement in the scene's surroundings and the unfamiliar visual appearance of objects that show through other geometries can also have a negative impact on the users' spatial understanding of the scene and thus affect *retrieval times*. From these considerations, we derived the following hypotheses about the results of the first experiment:

- H1.1 : Decreasing mean of *discovery time* over successive *blocks*,

- H1.2 : Decreasing mean of *retrieval time* over successive *blocks*,

- H1.3 : Decreasing mean of *covered distance* over successive *blocks*,

- H1.4 : Larger mean of *covered distance* in the *None* condition as compared to both of the other *techniques* and

- H1.5 : Larger mean of *discovery times* in the *None* condition as compared to both of the other *techniques*.

During a break of approximately 15 minutes users were asked to provide information about their age and gender, as well as previous experience with VR-systems. Using a short questionnaire, we checked whether any particular problems were experienced during the experiments.

In a second experiment we wanted to compare the three occlusion management *techniques* within subjects. We assumed that after having gained experience during the first experiment, users would have acquired a robust cognitive model, meaning that they were well trained with the task procedure. Thus, we estimated expert performance, in that effects of further learning would be negligible. However, since users were trained with only one of the *techniques*, we refrained from taking performance measures. Instead, we focused on subjective preference of users and user behavior in terms of proxemics.

We logged the distance between users as recorded during the presentation phase of each trial. As in the first study, the experimenters performed the role of the *presenter*, making sure to follow a strict storyboard in order to avoid biasing the distance data.

We conducted three additional *blocks* in that second experiment conforming to those from the first one – aside from this, users were now testing other occlusion management *technique* in each *block*. The order of the *techniques* was balanced among the 24 test subjects.

After having completed three *blocks* with different *technique* conditions, we asked users to score *techniques* on a 5 point Likert scale in the parameters of *spatial understanding* (how *techniques* support gathering information about the position of target objects in the scene), *collaboration* (how *techniques* support interpersonal communication about the scene) and *comfort* (the perceived comfort while interacting in presence of different occlusion management *techniques*).

Without the use of visual aids, close proximity to the *presenter* is generally required in order to see an indicated object. Our proposed show-through techniques, instead, enable to observe indicated objects in the scene from every viewpoint that the operational environment permits. Following the findings of Hall [15], we assumed that users would tend to remain in the close phase of social distance rather than in closer proximity. The size of the operational environment, however, hardly allowed users to interact with the virtual scene, while remaining in social distance to each other. Nevertheless, we expected much shorter average distances between users in the *None* condition as compared to both show-through *techniques*. Consequently, we estimated a user preference for show-through *techniques* with respect to *comfort*. Regarding *spatial understanding* we felt that the localization of objects in the scene was cognitively more difficult in cases in which they are only perceived as showing-through other geometries than if an appropriate viewpoint has to be obtained by walking around the virtual scene. We therefore expected the best scores for the *None* condition in the domain of *spatial understanding*. On the other hand, since show-through techniques allow users to see indicated objects from other perspectives than that of the *presenter*, the pure amount of information that can be gathered as a group of users is considerably increased. We considered that users may have become aware of this advantage, and we therefore took it into account when scoring *techniques* in the domain of *collaboration*. In summary, we noted the following hypotheses regarding the results of the second experiment:

- H2.1 : In average, larger *distances* will be kept between users in both show-through *techniques* as compared to the *None* condition,

- H2.2 : Strong user preference for both show-through *technique* in the domain of *comfort*,

- H2.3 : Lower subjective ratings for both show-through *techniques* as compared to the *None* condition in the domain of *spatial understanding* and

- H2.4 : Stronger user preference for both show-through *techniques* in the domain of *collaboration*.

4.4 Results and Discussion

4.4.1 User Performance

Performance data from the first experiment was collapsed and entered into a one-factor (*technique*) between subjects ANOVA considering *block* as a within-subjects variable. For all post-hoc comparisons, Bonferroni adjustment for α was applied.

Regarding *discovery time* (see Figure 5 top), we found a significant effect of *block* ($p < 0.001; F = 29.36$). Post-hoc tests revealed that all differences between the successive three *blocks* were significant (all $p < 0.05$). This confirms H1.1, stating that learning would have a significant effect on the time required to find target objects. No significant differences could be found for *technique*. Thus, we had to reject H1.5. Contradictory to that result, we were expecting longer *confirmation times* in the *None* condition. During the experiment, we observed users developing appropriate strategies for efficient interaction under different *technique* conditions. In the *None* condition, users generally followed the *presenter* to ensure a similar point of view, whereas they rarely changed their position when show-trough *techniques* were enabled. Therefore, users were able to localize indicated target objects in a comparable time for all *techniques*.

With respect to the *covered distance* (see Figure 5 bottom), we found significant effects for *block* ($p < 0.001; F = 18.15$). Post-hoc comparisons revealed significant differences between all three *block* conditions (all $p < 0.05$). This confirms H1.3, stating that

learning the task would allow users to minimize the required physical action. Additionally, significant effects were found for *technique* ($p < 0.05; F = 9.04$). Post-hoc comparisons revealed that in the *None* condition users moved significantly more than in both other conditions ($p < 0.001$). This result confirms H1.4. Without making use of show-through techniques, people are required to adapt their viewing position more often. Between both show-through techniques, we found no significant differences.
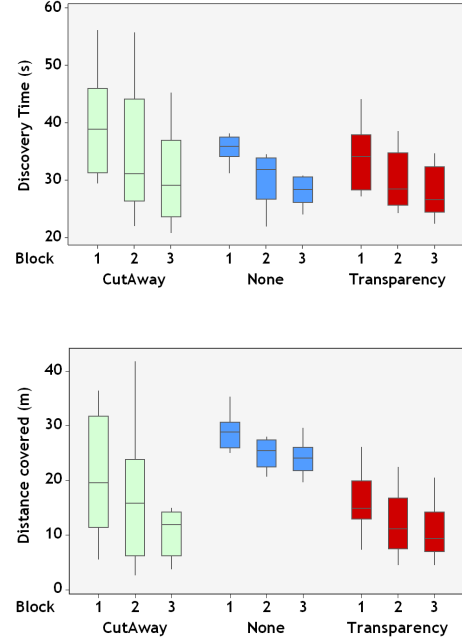

Figure 5: Boxplots of *discovery time* (top) and the corresponding *discovery distance* (bottom) during the first experiment. We can clearly see the learning effect over *blocks*, and also how users were required to move more if they were not provided with show-through *techniques*.

Retrieval time was only affected by *block* ($p < 0.001; F = 43.92$). Therefore, we can confirm H1.2, stating that learning the task would allow users to improve upon the rapidity of target retrieval. No significant effect of *technique* on *retrieval time* was found. We were assuming that the time required to retrieve target objects is strongly affected by the knowledge about the localization and appearance of these objects that the users were able to gather beforehand. We thus argue that the three tested occlusion handling *techniques* can comparably well support the users' spatial perception of objects within the 3D scene.

Overall, the results of the first experiment indicate that all tested occlusion management *techniques* allow for comparable user performance, not only in terms of fundamental task efficiency, but also with respect to supporting the users' learning progress. Regarding the required *distance covered* to localize indicated targets, we found significant benefits for both show-through *techniques*.

4.4.2 Proxemics

In the second experiment we analyzed the users mutual distance during the presentation phase. We compared the techniques between subjects because we expected the training during the first experiment to have an important impact at the users interaction. As such we were only considering data from those users, that were well trained with the respective *technique*.

Since the length of trials varied, we first normalized the distribution of distances to obtain the relative frequency of distance ranges

during the experiments. Figure 6 shows a plot of the normalized frequencies at which distances occurred during different *technique* conditions. For all *techniques* this distribution fits accurately to a normal distribution (Anderson-Darling normality test: $p < 0.005$).

It is apparent in these graphs that show-through *techniques* allowed users to keep at larger distances. In the *None* condition, instead, users were crowding each other, frequently ending up at intimate distances, sometimes even bumping into each other. Note that the position of the user was taken from the center between their eyes. This implies, that both heads were touching at distances of $\sim 20cm$ and $\sim 40cm$ their shoulders were touching.
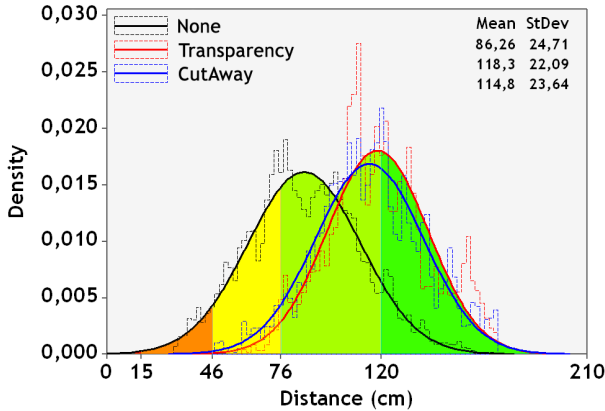

Figure 6: The distance kept between users in relation to the applied occlusion-handling *technique*. Dashed outlines represents the histogram of logged distances, shaded areas represent the "distances in man" as defined by E.T.Hall [15].

We observe, that in all conditions users kept about half of the time in the far phase of the personal distance. With show-through *techniques* applied, the other half of the time was spent in the close phase of the social distance. According to Hall [15] this is generally the preferred distance in collaborative work settings as simulated in our study. Thus, we assume, that distances between users would still be larger, if the interaction space would not be strongly affected by the following limitations of our display setup:

- Our working space was only about 3 meter wide, which clearly limited the maximum distance among users.

- The eye-wear used in the experiment consisted of relatively large shutter glasses, but still they limited the users' field of view. Particularly, peripheral vision of users was affected and as such that the subconscious attention to changes in the surrounding environment, e.g., colleagues approaching from beside, was impaired.

- There was barely any social hierarchy present between the *presenter* (Ph.D Student) and *observer* (students). In a real world setting this would be different. Recall that the more important somebody is, the larger are generally the distances kept to others.

However, in absence of show-through *techniques* users kept a lot more time in the close phase of the personal distance and, what is probably even more important, they could not avoid intruding frequently into the other's intimate distance. (see Table 1 for exact values).

For a statistical evaluation, we collapsed the data and entered it into a one-factor (*technique*) between subjects ANOVA. We found a significant effect of *technique* on *distance* ($F = 21.01, p < .001$).

Zone	Phase	None	Transparency	Cutaway
Intimate	Close	0.16%	0.00%	0.00%
Intimate	Far	5.10%	0.05%	0.18%
Personal	Close	28.46%	2.76%	5.05%
Personal	Far	57.59%	49.59%	53.09%
Social	Close	8.69%	47.60%	41.68%

Table 1: Average of time spend for users at each "distance of man" grouped by *technique*.

This effect stems again only from differences between the *None* condition and both show-through *techniques*. Between the latter two, no significant difference could be found.

We expect that if users would have had more space available, they would also have kept larger distances between the *presenter* and the *observer*, while increasing the size of the operational environment would not have had much effect on the distance observations in the *None* condition.

4.4.3 Subjective Ratings

Subjective ratings reveal a significant user preference for the *Cutaway* and *Transparency* conditions in terms of comfort. The *Friedman Rank Test* revealed significant differences among conditions ($p < 0.001$) with *None* as the bottom line, which confirms H2.2. Show-through techniques allow users to keep more comfortable distances between each other. Also, less viewpoint motion is required since even otherwise occluded geometry can be seen. In the domain of spatial understanding, we could not find significant differences among techniques ($p = 0.104$). Thus, we had to reject H2.3. It appears as if users were very confident about their ability to gather spatial knowledge in the virtual environment independently from the used occlusion handling technique.
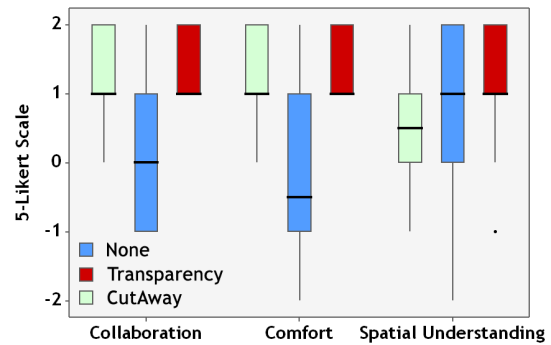

Figure 7: Subjective comparison of the three conditions, in terms of collaboration, comfort and spatial understanding of the model.

In the domain of collaboration support we have found significant differences among the subjective ratings of *techniques* ($p < 0.001$). Again, the None technique received the lowest score from users which confirms H2.4. Though our experimental task did not require much collaborative interaction but rather only fundamental information exchange, users seemed to predict further benefits for collaborative interaction from show-through techniques.

5 CONCLUSION AND FUTURE WORK

We suggested the use of show-through techniques for dealing with inter-user occlusion problems in multi-user virtual reality systems. The most important result of our user study, was that during presentations, show-through techniques reduced the number of cases in

which users needed to get very close or even bump into each other. Our findings also showed that they can maintain a socially convenient distance for most of the time while moving much less than without the techniques. Since spatial understanding of the scene does not seem severely affected by our techniques, users made use of them to consider social protocols as much as possible.

Regarding quantitative performance we found no significant difference between both tested show-through techniques. Considering the subjective ratings, the collaboration- and comfort- attributes were rated comparably well. Nevertheless, users rated the transparency technique better in terms of spatial understanding. We assume that this can be explained with the better preservation of depth cues through transparency than in case of rigorous cut-aways. The development and more detailed evaluation of multi-user interaction techniques that improve the trade-off between mutual information exchange, visual complexity and visual reliability of the belabored geometrical model remains to be an exciting research topic for future work.

In our work, we involved only two users with distinct roles – the presenter and the observer. However, in real life these roles change and users are taking turns at showing and observing, which needs to be supported appropriately. Furthermore, with show-through techniques, the observer sees a different perspective of the object being pointed to than the presenter, which can prove to be a limitation in certain situations. Often, for example, more than two people are involved in a collaborative discussion of a design scenario. The questions are if our techniques are similarly effective in such cases, how do they scale, and how do they handle situations wherein more than one person is pointing? One issue already became apparent: the size of the display needs to be appropriate in regards to the number of users involved. A three-meter wide display allows at most two users to maintain a comfortable distance to each other.

Show-through techniques are also beneficial for the *presenter* as he or she does not need to worry if an object is occluded from the *observer*'s point of view. In physical reality or when VR-systems offer only corresponding reality-based interfaces, showing objects is not always straight forward. One needs to consider the viewing position of observers. In cluttered 3D environments we experience an important difference between pointing or actually showing something. Pointing does not necessarily mean that the observer is seeing the indicated feature. Instead, pointing to an object often implies for colleagues to move around in order to arrive at an appropriate view position. Showing, on the other hand, often involves the effort of moving something physically into the field of view of others. In multi-user VR, show-through techniques can bridge this gap.

REFERENCES

[1] M. Agrawala, A. C. Beers, I. McDowall, B. Fröhlich, M. Bolas, and P. Hanrahan. The two-user responsive workbench: support for collaboration through individual views of a shared space. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 327–332, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

[2] V. Bayon, G. Griffiths, and J. R. Wilson. Multiple decoupled interaction: An interaction design approach for groupware interaction in co-located virtual environments. *Int. J. Hum.-Comput. Stud.*, 64(3):192–206, 2006.

[3] M. Billinghurst, S. Baldis, E. Miller, and S. Weghorst. Shared space: Collaborative information spaces. In *HCI International '97: Proceedings of the Seventh International Conference on Human-Computer Interaction-Volume 1*, pages 7–10, New York, NY, USA, 1997. Elsevier Science Inc.

[4] M. Burns and A. Finkelstein. Adaptive cutaways for comprehensible rendering of polygonal scenes. *ACM Trans. Graph.*, 27(5):1–7, 2008.

[5] J. Chastine and Y. Zhu. The cost of supporting references in collaborative augmented reality. In *GI '08: Proceedings of graphics inter-face 2008*, pages 275–282. Canadian Information Processing Society, 2008.

[6] C. Coffin and T. Hollerer. Interactive perspective cut-away views for general 3d scenes. In *VR '06: Proceedings of the IEEE conference on Virtual Reality*, page 118, 2006.

[7] J. E. Cutting. How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments and Computers*, 29(1):27–36, 1997.

[8] J. Diepstraten, D. Weiskopf, and T. Ertl. Transparency in interactive technical illustrations. *Computer Graphics Forum*, 21(3):317–325, 2002.

[9] J. Diepstraten, D. Weiskopf, and T. Ertl. Interactive cutaway illustrations. In *Computer Graphics Forum*, volume 22, pages 523–532, 2003.

[10] N. A. Dodgson. Autostereoscopic 3d displays. *Computer*, 38(8):31–36, 2005.

[11] N. Elmqvist, U. Assarsson, and P. Tsigas. Employing dynamic transparency for 3d occlusion management: Design issues and evaluation. *Human-Computer Interaction –INTERACT 2007*, pages 532–545, 2008.

[12] N. Elmqvist and P. Tsigas. A taxonomy of 3d occlusion management for visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(5):1095–1109, 2008.

[13] B. Fröhlich, R. Blach, O. Stefani, J. Hochstrate, J. Hoffmann, K. Klüger, and M. Bues. Implementing multi-viewer stereo displays. In *WSCG (Full Papers)*, pages 139–146, 2005.

[14] T. Grossman and R. Balakrishnan. Collaborative interaction with volumetric displays. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 383–392, New York, NY, USA, 2008. ACM.

[15] E. T. Hall. *The Hidden Dimension*. Doubleday, 1966.

[16] A. Jones, I. McDowall, H. Yamada, M. Bolas, and P. Debevec. Rendering for an interactive 360° light field display. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*, page 40, New York, NY, USA, 2007. ACM.

[17] K. Kiyokawa, M. Billinghurst, S. E. Hayes, A. Gupta, Y. Sannohe, and H. Kato. Communication behaviors of co-located users in collaborative ar interfaces. In *ISMAR '02: Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, page 139, Washington, DC, USA, 2002. IEEE Computer Society.

[18] M. Lucente. Interactive three-dimensional holographic displays: seeing the future in depth. *SIGGRAPH Comput. Graph.*, 31(2):63–67, 1997.

[19] H. Salzmann and B. Froehlich. The two-user seating buck: Enabling face-to-face discussions of novel car interface concepts. In *IEEE Virtual Reality Conference 2008 (VR 2008), 8-12 March 2008, Reno, Nevada, USA, Proceedings*, pages 75–82, 2008.

[20] H. Salzmann, M. Moehring, and B. Froehlich. Virtual vs. real-world pointing in two-user scenarios. *Virtual Reality Conference, IEEE*, 0:127–130, 2009.

[21] A. Simon. Usability of multiviewpoint images for spatial interaction in projection-based display systems. *IEEE Transactions on Visualization and Computer Graphics*, 13(1):26–33, 2007.