Excuse Me: Large Groups in Small Rooms

Ephraim Schott*
Bauhaus-Universität Weimar

Tony Jan Zoeppig†
Bauhaus-Universität Weimar

Anton Benjamin Lammert‡
Bauhaus-Universität Weimar

Bernd Froehlich§
Bauhaus-Universität Weimar

Figure 1: (a) No Visibility Technique, (b) Self-Translation, (c) Transparency, (d) Scaling-Others, (e) Displacing-Others

Abstract

Standing in a large crowd can be uncomfortable and usually results in other users obstructing the view of the virtual environment. In this paper, we present four techniques designed to improve the user's view in crowded environments. Inspired by related work on various transparency and clipping techniques, as well as observed user behavior in crowded scenarios, our paper addresses the visibility problem by locally manipulating the appearance of other users. Three of our techniques define a region of interest using a handheld flashlight metaphor. Depending on the technique, occluding users are either pushed to the side, scaled, or made partially transparent. The fourth technique allows users to vertically adjust their position. A user study with 24 participants found that the transparency technique was advantageous for quick search tasks. However, in a realistic museum setting, no clear favorite could be determined because the techniques make different trade-offs and users weighted these aspects differently. In a final ranking, the vertical position adjustment and transparency techniques were the most popular, but the scaling technique and vertical position adjustment were found to be the most natural.

Index Terms: Human-centered computing—Interaction paradigms—Virtual reality; Human-centered computing—Human computer interaction (HCI)—Interaction techniques; Human-centered computing—Collaborative and social computing—Collaborative and social computing theory, concepts and paradigms—Computer supported cooperative work

1 Introduction

Collaborative virtual reality (VR) platforms allow people from different locations to come together in shared virtual spaces and engage in various social activities such as attending lectures, concerts and virtual museum tours. With the increasing popularity and accessibility of virtual reality devices, there is a growing demand for social virtual environments (SVE) to accommodate larger gatherings and foster meaningful interactions in groups. However, with an increasing number of users virtual scenes tend to be more crowded and cluttered, which can have a negative impact on user experiences.

*e-mail: ephraim.schott@uni-weimar.de
†e-mail: tony.jan.zoeppig@uni-weimar.de
‡e-mail: anton.benjamin.lammert@uni-weimar.de
§e-mail: bernd.froehlich@uni-weimar.de

In crowded scenes, the user's view is more likely to be obstructed by other users, proxemic zones [24] are more likely to be violated and users might be detracted by the many stimuli surrounding them. These problems can lead to a decreased sense of immersion, hinder social interactions, and potentially affect comfort and usability of SVEs. To deal with some of these problems [38, 50] consider other users as occluders and propose methods to turn users or user groups invisible. Earlier work on occlusion distinguishes between entities that carry information or not [19]. The challenge with occluding users is that they always carry information that contributes to the social aspect of the virtual experience. However, it is up to the user to decide whether this information is currently important or not.

In this paper, we focus on reducing the occlusion problem in crowded scenes by giving individual users control over the visibility of the surrounding scene. We built techniques that rely on known concepts, such as navigation and transparency, and propose further techniques that rely on novel concepts that locally displace or scale down the avatars of other users to improve the view for users in the rear (Figure 1). We conducted a user study (N = 24) with a quantitative and a qualitative part to evaluate our techniques. The quantitative part was designed to address research questions with respect to the efficiency and effectiveness as well as comfort of our visibility techniques. In the qualitative part users were accompanied by an interviewer on a museum tour where they could choose a technique that fitted the various occlusion situations best.

Our work was motivated by the collaboration with a local museum that is interested in making digital twins of the museum remotely accessible to larger social groups such as school classes. They wanted to be able to have a museum guide in the digital copy who guides the visitor group through the museum. The virtual tour should resemble a tour through the real museum which was the actual residence of a poet, furnished with original furniture and collection pieces. The visitor group should stay together, move from room to room and interact with the guide. As a result we designed an abstracted tour through a virtual museum for our user study, where a guide can move a user group of up to 25 people to fixed positions in different rooms. The users have no other navigation techniques beyond physical walking. They need to solve a sequence of tasks that require visibility towards objects placed in the rooms of the museum while being surrounded by other users.

Our research aimed at addressing the visibility challenges in this context and resulted in the following main contributions:

- An overview of conceptual solutions for resolving occlusion,
- four novel visibility techniques inspired by related work for resolving occlusion caused by other users,
- evidence from a quantitative user study (N = 24) revealing the

efficiency and effectiveness of the four visibility techniques in a crowded room, and
- a discussion of the advantages and issues of these techniques based on a qualitative user interview in an open museum tour (N = 24) and the quantitative user study.

2 RELATED WORK

Virtual reality is predestined for the joint exploration of virtual 3D content. The quality of the experience depends on factors such as social presence, comfort and visibility. When engaging in collaborative virtual environments with many people, these factors need to be carefully balanced to provide an optimal experience for users. Our work builds on and extends related work that has addressed single or combined factors of this problem space. In the following sections we present prior research and concepts in the domains of social presence and occlusion.

2.1 Social Presence

Museum visits have an important social factor and are not solely driven by curiosity or specific interests in exhibitions. According to Prentice et al. [37], museums are frequently visited for the purpose of acquiring general knowledge, escaping routines and enjoying time with friends and family. In fact, a survey among 105 English museums found that the average museum group size is three visitors, and only 20% of visitors go on their own [44].

In the context of virtual museums, recreating a sense of closeness and collaboration similar to real-life experiences is essential. Gutwin et al. [23] identifies three essential elements that must be conveyed to achieve an effective collaboration in virtual environments: users (who), actions (what), and locations (where). Numerous visualization methods have been explored to enhance user-to-user awareness [35, 39], which is defined as the knowledge an individual acquires through their interactions with the environment [20].

In this context, related work on collaboration examines factors influencing users' social presence, such as intimacy and immediacy between users. In CVEs, social presence refers to an individual's sense of being with others as if they were physically present and being able to engage in meaningful social interactions [40]. For this feeling to emerge the VR system should enable its users to perceive each other and communicate through verbal and non-verbal cues about shared content [11, 22].

2.2 Occlusion

The occlusion of an object by another provides a visual depth cue which allows for the estimation of a relative distance [16]. If an object is not or only partially visible from a user's point of view, then it is obscured. Although occlusion is an important feature for our depth perception, it also has a negative impact on visual tasks.

Naturally, occlusion is caused by the accumulation of objects in a 3d scene or by other users. Elmqvist et al. [19] categorize entities that carry information as *targets* and entities without as *distractors*. We follow their notion in this work, and extend it with the term *occluder*, which stands for a distractor that currently blocks a user's free view on a target. Furthermore, we distinguish between static and dynamic distractors, whereas in our case static refers to objects and dynamic refers to other users.

2.2.1 Resolving Occlusion Through Navigation

A simple way to resolve occlusion is to shift one's viewpoint to another location, for example by stepping to the side. The task of changing the viewpoint is called *travel* or, for short distances and specific views, *maneuvering* [9]. It is noteworthy that prior research on navigation techniques predominantly addresses exploration and search tasks, with limited emphasis on occlusion.

Various travel techniques exist in VR, with steering and teleportation being the most common [9, 10]. Steering allows for small changes of perspective due to its continuous direction specification. In general, steering comes with the disadvantage of causing cybersickness for some users. One assumed reason for this is the mismatch between visually perceived motion flow and the perception of the user's vestibular system [12, 29, 34]. However, related work has also shown that low translation speed also induces less cybersickness [6, 41]. In our case, we deem steering techniques suitable for use in confined spaces with numerous distractors, as they enable users to make small, slow movements that suffice to attain a better viewpoint. Especially vertically ascending can avoid obstructions caused by others and provide a comprehensive overview [1]. Steering extensions were developed, especially to address vertical height adjustment. Examples of these include elevators and magic carpets [33, 45]. In their work, Weissker et al. [49] also introduce a teleportation technique that enables precise height specifications.

Different maneuvering approaches, based on pivot points or anchors, have been presented that enable single users and groups to quickly adjust their viewpoints [8, 25, 28]. Although teleport techniques are good solutions for examining objects and bypassing static distractors, we do not consider them well suited for spaces with many dynamic distractors. With many users in a scene, as in our scenario, individual discontinuous teleport methods, such as *Anchored Jumping* [8], would result only in temporary improvements in visibility. Each jump of a user could turn themselves into an occluder for another user, which could trigger a chain reaction of viewpoint adaptations. Additionally, other works point out that the target of a jump should be indicated for the awareness of other users, which would lead to further cluttering of the scene [15, 48]. Hence, we neglected teleport techniques in this work.

2.2.2 Resolving Occlusion Through Transparency

See-through and cut-out techniques are a common approach to resolve occlusion [4, 7, 14, 18, 30, 52]. Transparency represents a powerful method for searching for and interacting with hidden objects. Magic lenses, as introduced by Bier et al. [7], established the groundwork for the use of transparent and dynamic overlays to visualize layered information in 2D graphical user interfaces. Assarsson et al. [4] presented a hand-held transparency flashlight and an implicit target transparency approach that allowed users to quickly find entities with information. The intuitive affordance of this flashlight technique inspired our own development. Other works, such as *GravityZone* and *AlphaCursor*, demonstrated that transparency techniques outperformed other techniques without transparency in search tasks and in concealed target-reaching tasks [47, 52].

Other works use cut-out techniques to improve a user's understanding of 3d scenes [14, 46]. When using transparency, it is essential to carefully design the see-through techniques. Occlusion layers offer depth cues that may be altered or reversed when employing certain cut-out or transparency techniques, potentially resulting in issues related to a user's depth perception [27]. In this context, Eren et al. [21] investigated how additional visualizations can support the depth perception in see-through scenarios.

Most research on transparency techniques primarily focuses on selection and search tasks [31], often aiming to make objects or entire spaces transparent [4, 14]. However, transparency is not only beneficial for individual users but can also simplify and enhance collaboration among multiple users [5]. For instance, Argelaguet et al. [3] introduce collaborative pointing techniques that utilize see-through technology, allowing users to indicate points of interest that would not be visible from the perspective of other users. This approach helps users to avoid violating each other's proxemic zones in confined, hard-to-see spaces [24].

Another approach that addresses both a user's private space and occlusion by other users is *Go-Through* by Reinhardt et al. [38]. In a museum setting, their work evaluates the use of avatars becoming transparent when approached closely. However, they do not consider making distant users transparent. *SocialSlider* by Wolf et al. [50], takes it one step further and allows users to manipulate the transparency of individual users, groups or all users to get a better view in rooms with dynamic distractors. In our work, we build upon this approach but consider distinct user settings unnecessary. In particular, we deem a pointing gesture to be inappropriate for making others transparent, as it may be perceived offensively or disrespectful by other visitors. Furthermore, we refrain from completely hiding users, as it has been reported in [50] that the presence of other users is quickly forgotten.

2.2.3 Resolving Occlusion Through Scaling

Another way to improve visibility of hidden objects is through scaling methods. Scaling can be applied to the user themselves, the target object, or occluders [28, 36]. The concept of scaling oneself is familiar from various navigation methods [13, 26, 28]. Similar to previous height adjustment, with these techniques, users can see over obstacles and cover distances faster due to their increased size. When scaling includes adjusting the interpupillary distance, the world becomes a miniature in this technique. Providing the user with a scaled-down copy of an object or a world-in-miniature (WIM) adds an additional view which can also resolve occlusions [42]. Nevertheless, we consider the self-scaling and target scaling approaches to be unsuitable as both distort the true scale of objects and are not well-suited for museum spaces where the original scale of objects and their spatial relation to users matter. Also, continuously displaying a larger user to all others would worsen the space problem.

What is particularly intriguing, however, is that to our knowledge, there is no technique where other occluding users in the field of view are locally scaled down.

2.2.4 Resolving Occlusion Through Displacement

Displacement can be used on targets and distractors to improve visibility [52]. Elmqvist et al. [17] introduce the *BalloonProbe*, a technique in which a user moves a sphere between objects and inflates it. This inflation causes the objects to arrange themselves more visibly along the surface of the sphere. Although displacement techniques are mostly used for selecting in cluttered object collections, we will investigate this technique further in our work.

3 Visibility Techniques

Based on the presented approaches to mitigate occlusion, we developed four techniques that are particularly suitable for smaller spaces with multiple users as occluders. These techniques are vertical navigation, transparency, scaling and displacement (Figure 2). They can all be controlled with a single controller. The *Self-Translation* technique requires direct input via a joystick, typically called thumbstick, while the other three techniques are directed through intuitive movements of the controller, similar to using a flashlight [4]. We designed the techniques for a social scenario in a museum, focusing on preserving the real spatial relations as closely as possible while enhancing the view. Consequently, we excluded approaches involving smaller copies, such as a WIM, as well as concepts that place all users in the front row, which would disrupt social formations and impede communication among visitors. The employed flashlight concept enables users to define a point of interest so that only immediate occluders are adjusted, resulting in minimal changes to user formations. In the following sections, we present our techniques in detail, along with parameter choices derived from a pilot study.

3.1 Self Translation

The *Self-Translation* technique is inspired by the natural behavior of standing on one's tiptoes to get a better view and by the arrangement of people in group photos. People stand in rows, arranged by size, so that everyone is visible in the photo. A similar principle is also used for seating in theaters. We observed similar behavior at a VR event with 21 users employing *flying* navigation: The users arranged themselves at different heights around a virtual table at which interaction techniques were demonstrated.

The objective of *Self-Translation* is to enable users to vertically adjust their position within a virtual environment. This adjustment allows users to look above other people and occluders, explore different viewpoints and gain an overview over a room.

Two versions of this technique were implemented: a discrete and continuous. The continuous technique was implemented using a frame-rate independent rate control. Displacing the thumbstick of the controller in vertical direction results in up and down movements of the user's navigation platform, which corresponded to the floor level at startup. The velocity of the up and down movement was limited to a maximum speed (max_speed) and the data values of the thumbstick were linearly mapped into the range between $[-max\_speed, max\_speed]$.

The discrete technique was implemented to limit motion sickness by reducing the visual motion flow. Moving the thumbstick up or down above a threshold triggered an elevation change of the user's navigation platform by ± 25 cm. Additionally, we included an optional fade-to-black transition during height adjustments.

For both techniques the navigation platform could not be moved below zero and the maximal height was clamped at two meters. Although suggested by Medeiros et al. [33], we did not visualize a platform because users could not move very high and a representation would have intersected other avatars in a confusing way.

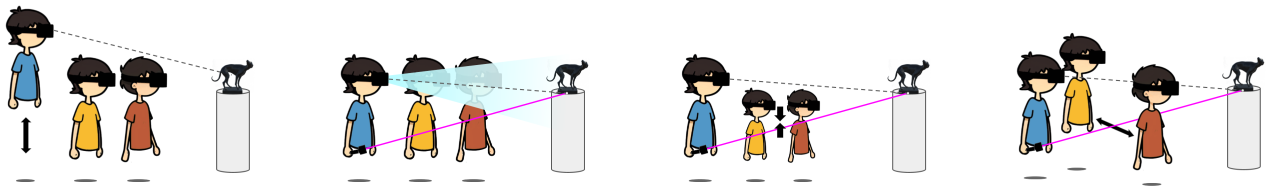
3.2 Transparency

The *Transparency* technique is inspired by the *SocialSlider* [50] concept and different see-through flashlight metaphors [4, 18, 30]. The objective of this method is to provide users with the ability to see through other users. For this purpose, the user controls a ray with the controller. The ray intersection point $\mathbf{p}$ of the ray with the environment excluding the other users is computed. A cone from the center of the head with its central axis through the intersection point $\mathbf{p}$ was created. Within this cone the other users were rendered translucent. We applied a linear falloff from the center of the cone, where other users were almost completely transparent to the boundary where they became opaque. The diameter of the cone was at the intersection point was defined by the parameter $cone\_diameter$. Even though the cone starts from the eye, for users this feels like controlling a flashlight with a transparency effect because the intersection point of the ray controls the direction of the cone. This technique was inspired by Argelaguet et al. [2] who suggested to control a ray from the eye by the direction of a handheld controller to avoid eye-hand occlusion mismatches.

3.3 Scaling Others

The idea of scaling down occluders such that one can look over them motivated the development of the *Scaling-Others* technique. Thus, with this technique the user defines a cone-shaped area of interest using the same flashlight approach as the *Transparency* technique. All occluders inside this cone are scaled down. A scale factor is calculated so that avatars inside the cone fit below the cone (Figure 3). The scaling center is the avatar's projected position on the ground plane. However, avatars are not scaled below a minimal fraction of the original height ($avatar\_min\_height$) such that they can be always easily perceived.

Since occluders in our case can also be other users, we needed to consider that a user's self-perception and embodiment is not changed by the visibility technique of another user. Having one's own scale reduced by another user could lead to discomfort and, depending on the scaling center, to a change in position and thereby potentially even induce motion sickness. Therefore, we opted for applying changes only to the avatars of other users within the view of the flashlight controlling user.



(a) Self-translation technique (b) Transparency technique. (c) Scaling-others technique. (d) Displacing-others technique.

Figure 2: The four visibility techniques.

This way, a user in the back row can scale down all users in front of them without those in front perceiving the changes. However, inconsistent perceptions of the environment among users violate the concept of a coherent workspace and can lead to collaboration difficulties. For example, user A cannot shake hands with a user B, who is scaled down, because A perceives B's hand at a different position than user B herself due to scaling. To minimize such inconsistencies and associated issues, the *Scaling-Others* technique was also designed as an easy to control handheld flashlight tool in the same way as for the *Transparency* technique. This allows users to determine where they want to direct their attention and when they wish to reduce the height of other users.

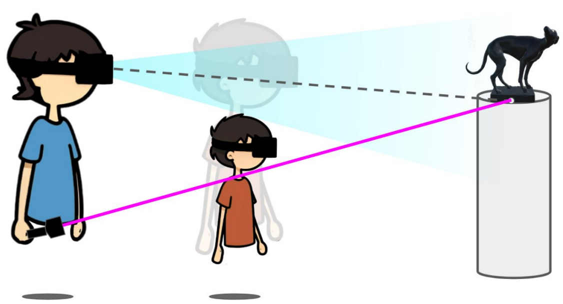


Figure 3: Illustration of local avatar manipulation (translucent avatar shows original size; red opaque avatar shows locally scaled version).

3.4 Displacing Others

In real-world situations when a small group moves through a museum together, friends naturally step aside when told they are obstructing our view. The *Displacing-Others* technique is inspired by this behavior and follows the *BalloonProbe* concept of Elmqvist et al. [17]. Instead of a sphere, however, we again use the flashlight approach, which was also used for the other techniques.

With the *Displacing-Others* technique all occluders within the flashlight cone are horizontally moved out of the cone. For this displacement approach to work with other users, we decided to reposition other users' avatars only locally.

3.5 Pilot Study

In a pilot study with four experienced VR users we compared our techniques and different implementations of them in a short search task and an exploration task. The goal was to identify appropriate parametrization of the techniques to leverage the different effects of each technique. The pilot study involved a virtual room with simulated avatars and was always conducted by two collocated users simultaneously. During the design phase and in the pilot study, it became evident that each technique had certain advantages and disadvantages. In this section, we describe the trade-offs of each technique, state our parameters for reproducibility purposes and discuss the design decisions we made.

Self-Translation: This technique is the only method that requires virtual navigation. In the pilot study we tested both the continuous and the discrete version, and decided in favor of the continuous as it provides the user with more freedom and does not require multiple steps. For the continuous implementation we had to find a balance between quick viewpoint adaptation and not making users sick. We found $max\_speed = 2m/s$ to be a suitable maximal speed.

A crucial design choice in the *Self-Translation* technique involved determining whether to distribute a user's vertical position to all other clients (consistent representation) or only locally adapt their height (inconsistent representation), which would lead to incoherent perceptions of a user's height across clients. The primary advantage of this technique is the unobstructed overview of the room it provides. However, when other users also adjust their vertical position and it is displayed consistently, the benefit of this technique diminishes as users may again obstruct the view of others. In the pilot study, we compared consistent and inconsistent representations and decided to display collocated users consistently, as inconsistent heights led to misleading communication and desynchronization between social gestures. For the search task two versions were tested; one where the simulated avatars adjusted their height depending on the distance to the exhibit, and one where they remained stationary. In the version with height adjustment, we observed that at times, it was sufficient for a participant to wait for the simulated users to adjust their height in order to see a visual target through their non-existing lower bodies. Hence, to encourage users to actively explore the technique, we opted for the version with a constant height for simulated users which always required the user to change their own height.

Transparency: This technique was the fastest among the flashlight methods in the pilot study, immediately clearing the view without the delays associated with height translation, scaling, or pushing avatars aside with other techniques. However, depending on the selected cone diameters and transparency setting a significant amount of social information from surrounding users could be lost. In the pilot study, we conducted tests with various configurations and found that the parameter $cone\_diameter = 1m$ and a transparency of 90% at the center of the cone were suitable for maintaining user visibility while also revealing occluded details.

Scaling-Others and Displacing-Others: For both techniques we use an animation of other users' avatars into their new height or position. The animation speed is an important parameter for these methods. If it is too fast, the method causes too much unrest or even flickering, while if it is too slow, fast search tasks cannot be properly accomplished. In the pilot study, we compared different speeds and decided to use an adaptive animation speed, where the speed is proportional to the difference between the current and target values or positions, i.e. slow speeds for short distances and larger speeds for larger distances. For the scaling and for the displacement we used 0.3m/s for a distance of 1m, i.e. for *Scaling-Others* and *Displacing-Others* techniques the avatars moved or changed at similar speeds. For the *Scaling-Others* technique the minimum height of avatars was set to 20% of their original height, i.e. $avatar\_min\_height = 0.2$. The cone width at the intersection point of the ray was $cone\_diameter = 1m$, identical to the *Transparency* technique.

4 Evaluation

The overarching goal of this work is to provide users with tools for improving their visibility in crowded rooms. As described in the introduction, our scenario considers all participants to be part of a loosely connected social group, like a school class or work group. To avoid isolating users within their social context, our second objective was to ensure that users remained aware of social interactions happening around them.

Within our scenario, we focused on rear user positions because that is where the occlusion is most severe. To evaluate the techniques with respect to visibility and social participation, we divided our study into two parts. In the first part, the user is alone in a museum room with simulated participants and has to complete a visual discovery task. In the second part, the participant takes part in a museum tour through three rooms together with the study conductor, who carried out a semi-structured interview, and a simulated guide.

4.1 Experimental Setup

We conducted both parts of the study in a quiet office which was equipped with two Meta Quest Pro head-mounted displays (HMDs). Both HMDs were connected to graphics workstations and operated in link mode. Our application was built with Unity3D and the networking was implemented with Unity Netcode. A 3d reconstruction of a museum served as the virtual environment of our study. It was rendered at the HMD's native resolution (1,920 x 1,800) at 90 Hz.

Users had a physical interaction space of 3.5 m x 2.5 m. In the first part users were standing in an approximately 4.5 m x 4.5 m virtual museum room. Three virtual museum rooms with different sizes were used for the second part. Besides the *Self-Translation* technique no virtual navigation was offered to the participants, limiting them to physical walking.

In both studies 25 simulated users were standing in the room with the participant. This corresponds to the average class size of all OECD countries plus one teacher [43]. The simulated users had different body heights (consistent through trials) ranging from 1.6 m to 1.9 m with static heads and bodies, since dynamic avatars could create slightly different occlusion situations for each participant.

4.2 Tasks and Conditions

We evaluated our techniques in two separate tasks. The first task investigates the visibility characteristics of our techniques and therefore follows the pattern of a controlled search task while the second aims at gathering qualitative insights. Thus, it combines a social exploration task with a semi-structured interview.

4.2.1 Visibility Search Task

In the visibility search task, participants had to search and read out numbers. They were instructed to be as fast as possible. For this participants were placed in the back of the room shown in Figure 4a. They had an occluded view onto an Achilles statue on the opposite site of the room (see Figure 4b). 15 targets were placed at different locations on the statue. We divided these locations into three regions (bottom: 0.0 m - 0.9 m, center: 0.9 m - 1.8 m, top: 1.8 m - 2.8 m).

In total, we showed the participants 33 targets for each technique, with the first three targets being warm-up trials. In general, participants were asked to apply the techniques, but they were also allowed to move. Each trial began with a sound indicating the appearance of a target and was stopped by the conductor via a button press when the number was read correctly. Only one target was visible at a time, which was additionally indicated by a guide next to the statue. The guide's hand was animated to point at the target and had a green ray emanating from the hand that was directed at the target (Figure 5). After each trial, participants were reset to their starting position and had to lower their hand. To avoid order effects, the targets had a different order for each technique.

4.2.2 Museum Tour

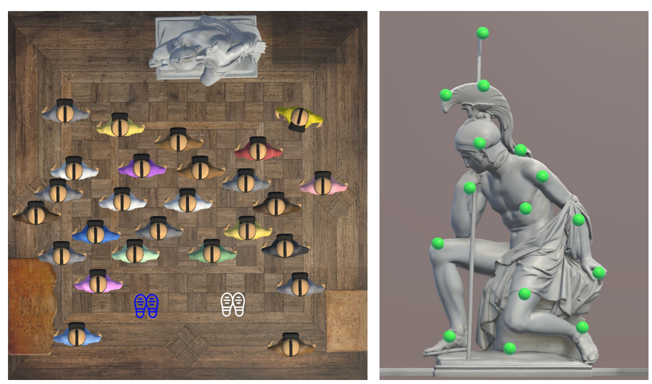
The second part of our study focused on gathering qualitative insights by looking at a more realistic scenario. Visitors often explore museums in social groups and do not have to be quick in discovering a feature on an exhibit. Thus, the study conductor joined our participant on a tour through three rooms of our museum (see Figure 6) and conducted a semi-structured interview.

In each room, the conductor could play and pause an audio file about the room. The audio source was located at the guide's head position who was placed at the front of each room, similar to the first task. The audio playback was spatial and we animated the guide's lips so that the guide was clearly identifiable as the speaker. As in the first study, 25 additional users were simulated.

The tour started in the largest room, where the participant and the conductor started right next to each other. The interview began after the experimenter had actively introduced themselves and explained how to switch techniques. The following questions were asked:

- How much of the statue do you see without a technique?
- Can you guess how many people are in the room?
- Do you feel comfortable with the distances to other users?

Following the initial questions, participants were encouraged to experiment with the techniques while listening to the tour guide, who provided commentary for approximately one and a half minutes. Subsequently, participants were asked to reflect on their behavior during the guide's presentation and express their preferences for the techniques relevant to the current room. This process was repeated as participants progressed to subsequent rooms. It is noteworthy that each successive room was smaller than the previous one, so the distances to other users decreased. Consequently, we repeated the initial questions about comfort and visibility in each space. Participant and conductor were always placed in the penultimate row.



(a) Top-down view on virtual experiment room. (b) Statue with targets.

Figure 4: Visibility search task study setup.

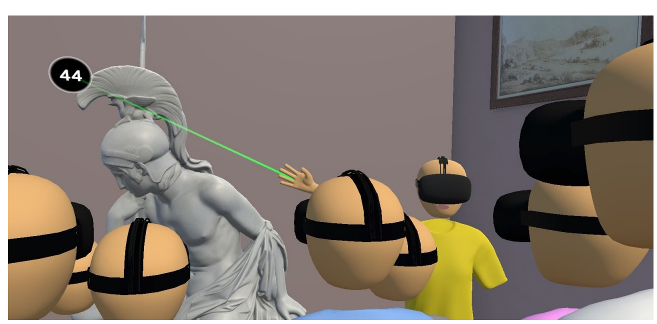


Figure 5: Guide pointing towards a target.

(a) Large sized room (4.5 m x 9.5m)

(b) Medium sized room (4.5m x 4.5m)

(c) Small sized room (2.5m x 5.5m)

Figure 6: Museum rooms visited during the museum tour.

In addition to listening to the guide and experimenting with different techniques, we prepared a small interactive game for the last two rooms. The game was verbally explained by the conductor. The objective was to create a collaborative interaction between the participant and the conductor. They took turns acting as *presenters*, each receiving an image in their hand to show to their peer, who assumed the role of a *searcher*. Simultaneously, three targets were displayed on the statues in the room. These targets displayed numbers for the presenter and images for the searcher. The searcher's task was to identify the image held by the presenter among the displayed targets and to provide information about the target's position. Ultimately, the presenter had to read the number at the described position.

This task abstracts a typical interaction between two museum visitors who draw each other's attention to shared content. All techniques were tested during this interactive game and users were interviewed for their opinions on the techniques.

4.3 Participants

Our experiment was conducted with 24 participants (12 male, 12 female) between 18 and 35 years of age (Mean (M) = 27.75, Standard Deviation ($\sigma = 4.52$)) and a height between 160 cm and 193 cm (M = 174.71 cm, $\sigma = 10.51$ cm). Eleven participants stated to be very familiar with VR; five were just familiar; four had some experience; one reported having had only a single experience and three had no prior VR experience. Participation was compensated with 10 Euros.

4.4 Procedure

The design of the study followed a within-subject design with four conditions: *Self-Translation*, *Transparency*, *Scaling-Others*, and *Displacing-Others*. To avoid order effects, we balanced the arrangement of our techniques using a Latin Square.

We welcomed our study participants to our laboratory and gave them an introduction to our study. After signing a consent form and completing a demographics questionnaire, the first part of the study was started. The controls were explained to the participants and then an HMD was put on. The task of the first part was explained and the participant got time to get familiar with their technique and the environment. Once the participants were ready, the first target was displayed. For each condition, the first three targets served as a training and were not measured. After the warm-up, participants had to perform 30 (two times each target) search tasks. After each condition a custom questionnaire was filled that asked about participants' comfort level, and about the ease of learning, ease of use, confusion, and overall impression of the technique. In addition, users indicated their perceived task load in a Raw TLX questionnaire. Once all conditions were completed, the techniques were ranked by the participants in a questionnaire. The first part of the user study took between 30 and 50 min to complete.

The second part of the study was conducted without any breaks and participants stayed in VR for 15 to 20 min. The interview was conducted within this time as described in 4.2.2. After the tour users ranked the techniques once more in a final questionnaire.

4.5 Dependent Variables

During the visibility search task, our system recorded the task completion time (TCT) and distance traveled, which was defined as the accumulated head movement distances. No quantitative data were collected in the second part, however, the interview was recorded for a qualitative analysis. Further data were collected through questionnaires. The Raw TLX questionnaire assesses a participant's perceived mental, physical, and temporal demand, and their perceived performance, effort and frustration on a scale from 0 to 100. The answers result in an overall task load score. Our custom questionnaires administered after each technique quantified a perceived change of well-being during the task (1 to 10). Additionally, users provided ratings on a 7-point Likert scale for factors such as ease of learning, ease of use, and confusion. They also rated their overall impression on a scale from 1 to 5, with 1 indicating "very good" and 5 indicating "very poor." In the final questionnaires for both parts of the study, participants were asked to rank all of the techniques and to state what they liked or disliked about each technique.

4.6 Hypotheses

Before the study, we formulated hypotheses as a basis for conducting inferential statistical tests. In the visibility search task, we wanted to find out which visibility technique was the most efficient in terms of TCT and distance traveled. We expected the transparency technique to outperform others given similar results in related work [47, 52] and because it does not require animations or adjustments.

- **H1**. The average TCT in the *Transparency* condition will be smaller than the other conditions in the search task.
- **H2**. The average distance traveled in the *Transparency* condition will be smaller than the other conditions in the search task.

With respect to the Raw TLX and our customized questionnaire, we formulated the hypotheses in an undirected manner, as all techniques are similarly controlled, have different advantages that may compensate for each other, and participants' preferences may vary. Given the unpredictable directions of the results, we assumed that each technique would influence all resulting variables.

- **H3**. The mean scores for task load will be different based on the used visibility technique in the search task.
- **H4**. The mean scores for change in feeling, ease of learning, ease of use, confusion, and overall impression will be different based on the used visibility technique.

5 Results

In this section, we present the results of our evaluation without interpretation. We begin by analyzing our quantitative data using inferential statistics and subsequently present the qualitative data obtained from questionnaires and the interview.

5.1 Quantitative Data

Each condition of the visibility search task resulted in 30 trials per user. Overall, 720 trials per technique were performed in the first task. Following our hypotheses, we conducted data analysis on the collected data from the visibility search task and the questionnaires to assess significance. A Kolmogorov–Smirnov test [32] showed that our data was not normally distributed. Consequently, we initially employed the Friedman test on all data sets to ascertain if there was general significance among the results of the techniques. If significance was indicated, we proceeded with Bonferroni corrected post-hoc analyses employing the Wilcoxon signed-rank test [51] to assess the significance of each pair of techniques.

5.1.1 Search Performance

The box plots in Figure 7 show the distribution of the measured TCT and distance traveled during the search task. We extend it by inferential analyses in the following:

Task Completion Time (H1): The Friedman test revealed that the TCT required in the visibility search task was significantly affected by the choice of visibility technique ($\chi^2(3) = 64.595$, $p<0.001$). All post-hocs tests that included a comparison between the *Transparency* technique and another technique could detect a significant influence of the choice of the visibility technique on the required TCT (all $p<0.001$), which supports **H1**.

Distance Traveled (H2): A significant influence of visibility technique on the traveled distance was indicated by the Friedman test ($\chi^2(3) = 703.722$, $p<0.001$). Again, all post-hocs tests that included a comparison between the *Transparency* technique and another technique show a significant influence of technique on the traveled distance (all $p<0.001$), thus affirming **H2**.

5.1.2 Task Load

An overview of the measured task load scores is shown in Figure 8. The Friedman test showed a significant influence for the measured overall scores ($\chi^2(3) = 35.988$, $p<0.001$) and for the respective sub scores of the Raw TLX (all $p \leq 0.003$). The subsequent post-hoc tests for the overall score revealed significance only for comparisons of the *Transparency* technique with the others (all $p<0.001$), as well as for the comparison between the *Scaling-Others* and *Displacing-Others* ($p<0.025$), thus only partially supporting **H3**.

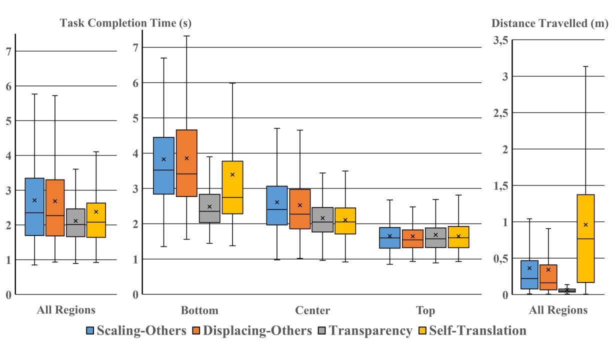


Figure 7: Box plots displaying the task completion times and the distances traveled for the visual search task.

	Scaling-Others		Displacing-Others		Transparency		Self-Translation	
	M	σ	M	σ	M	σ	M	σ
Feeling	1.2917	2.2075	1.6957	2.5270	0.3333	0.8498	2.75	3.1656
Learning	6.2083	0.9991	6.0870	1.0597	6.0583	0.1998	6.5833	0.9538
Ease of Use	4.9583	1.3687	4.6957	1.3330	6.875	0.3307	6.0833	1.5523
Confusion	5.4583	1.3222	5.4348	1.2794	6.8333	0.4714	5.8333	1.8181
Impression	2.6666	0.8498	2.9565	0.9991	1.5417	0.9119	2.5833	1.3202

Table 1: Means (M) and standard deviations (σ) of responses given to the custom single-item questions.

	Friedman		SO / DO	SO / T	SO / ST	DO / T	DO / ST	T / ST
	χ^2	p	p	p	p	p	p	p
Feeling	15.613	0.0014	0.3011	<0.001	0.0093	<0.001	0.0225	<0.001
Learning	11.439	0.0096	0.1209	<0.001	0.048	<0.001	0.0116	<0.001
Ease of Use	28.213	<0.001	0.2416	<0.001	0.0111	<0.001	0.0032	0.0011
Confusion	17.191	<0.001	0.3696	<0.001	0.0727	<0.001	0.068	0.001
Impression	17.739	<0.001	0.0703	<0.001	0.8791	<0.001	0.1961	<0.001

Table 2: p-values calculated for significance tests between all visibility techniques and technique pairs (SO = Scaling-Others, DO = Displacing-Others, T = Transparency, ST = Self-Translation)

5.1.3 Custom Scores

An overview of the resulting average and standard deviation values for the collected data from our single-item questions on change of feeling, ease of learning, ease of use, confusion, and overall impression can be seen in Table 1. The conducted Friedman tests revealed significant influences by the choice of the visibility technique for change of feeling ($\chi^2(3) = 15.613$, $p = 0.0014$), ease of learning ($\chi^2(3) = 11.439$, $p = 0.0096$), ease of use ($\chi^2(3) = 28.213$, $p<0.001$), confusion ($\chi^2(3) = 17.191$, $p<0.001$) as well as overall impression ($\chi^2(3) = 17.739$, $p<0.001$). The overview of all Friedman and post-hoc test results are shown in Table 2. Given that certain comparisons lack statistical significance, **H4** is not supported.

5.2 Qualitative Data

In this section, we present technique preference rankings, feedback, and responses from the interviews. We used axial coding to analyze the qualitative data from the questionnaires and the interview.

5.2.1 Technique Ranking

Figure 9 shows how often each technique was assigned to the different ranks after the visibility search task and after the museum tour. While the *Transparency* technique was still the most preferred after the search task, it was more evenly split across ranks 1 and 2 after the museum tour. The *Self-Translation* technique was ranked second most often in the search task, but then ranked first most often in the museum tour. The *Scaling-Others* and *Displacing-Others* techniques were ranked primarily third and fourth in both tasks.

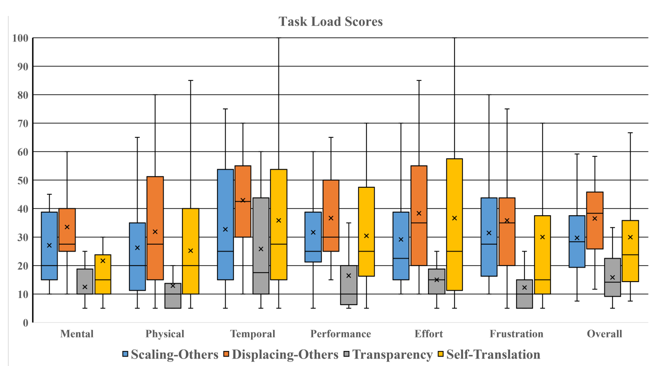


Figure 8: Box plot displaying the task load scores measured for each visibility technique during the visibility search task.

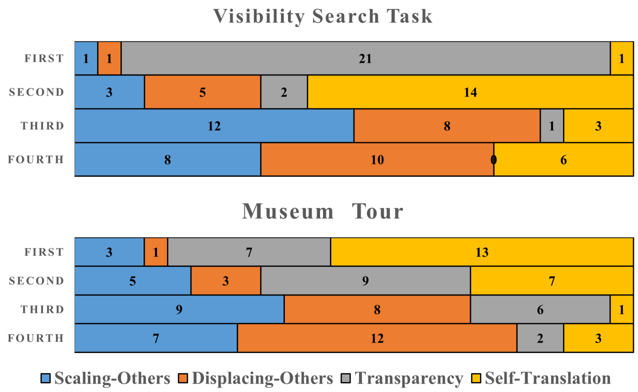


Figure 9: Rankings submitted for each technique after the search task and museum tour were completed with all of them.

5.2.2 Feedback on Techniques

To highlight the key aspects of each technique, we first summarize the participants' feedback from the questionnaires and extend these findings with insights gathered from the semi-structured interviews.

Self-Translation: Participants appreciated the overview and improved visibility that the *Self-Translation* technique provided (Comment Occurrence (CO) = 16). Furthermore, convenience and control (CO = 8), enhanced attention for experience (CO = 3), and more personal space (CO = 3) were mentioned positively. Participant (P) 6 mentioned: "I'm not that tall, but with this technique I was able to change my size to have a full overview, which was really nice." In addition, P11 wrote "I can see everything and don't have to keep my hand up. Also, the people are further away and I feel like I have more personal space."

When asked about improvements and negative comments, the 20 participants who rated this technique first or second noted that this technique did not resolve all occlusions (CO = 5) especially close to the floor, that they felt a bit of motion sickness (CO = 5), that the elevated perspective was unnatural and did not allow for a front view (CO = 4), and that they felt as "not being part of the group" (P17, CO = 3). Two participants had no negative comments (CO = 2). All participants who rated *Self-Translation* third or fourth, reported a strong cybersickness (CO = 4) and reported a "wrong" (P5) or "unnatural" (P1) room perception.

Transparency: Participants appreciated the technique for its simplicity and convenience (CO = 14), its improved visibility (CO = 11), its support of inspecting details (CO = 4), and the absence of motion sickness (CO = 2). For example, P12 stated: "You can see the objects from the front and you immediately get a better view."

When asked about improvements and negative comments, participants expressed a desire for a larger or adjustable cone size (CO = 14), a desire for full transparency rather than partial transparency (CO = 10), a feeling that the technology did not fit naturally in a museum context (CO = 4), and some users had no specific negative comments (CO = 2). P21 mentioned: "It's weird looking through someone." In the interview several users expressed a preference for a larger cone, especially in rooms with wider object arrangements.

Displacing-Others: Participants appreciated this technique mainly for the improvement of visibility (CO = 14) with P11 stating they "can see one person well from head till toes" and the feeling of realism or naturalness (CO = 6). P13 mentioned that this technique "feels natural and you can see a big part of the objects."

The participants criticized a lack of visibility and too small field of view (CO = 7), lack of control (CO = 9), and low perceived efficiency (CO = 4). P10 stated: "It was not always clear where I had to point to move the users out of the view."

Scaling-Others: Participants mainly reported that this technique improves the visibility of exhibits (CO = 14) and has a good ease of use (CO = 6). Two participants stated that this technique improves the perceived comfort (CO = 2). P22 said that it "feel[s] like there is more space in the room when people are smaller" and P1 stated that *Scaling-Others* feels similar to the *Self-Translation* technique: "It is like the self translation but it does not make me sick." In the museum tour P17 described that it took them a bit of time to discover this technique for themselves. P17 explained that they initially held the flashlight on the statue for an extended period until all users became small, and then "gently waved it to keep multiple users small".

The main criticism about this technique focused on the controls of the technique (CO = 14) and that even though the users were scaled down, they still blocked the sight on lower elements (CO = 9). P9 stated about the controls: "It was slow, I usually knew where I wanted to look, but had to wait for the scaling animation." Additionally, participants mentioned, that "Scaling others is distracting, especially if they are interacting with me" (P20, CO = 2).

5.2.3 Museum Tour Interview

To complete our report of the interview, we further describe the participants' impressions of the rooms and summarize their answers.

Room 1: In the largest room, which featured a dark exhibit depicting two statues, participants reported limited visibility. Regarding the left statue, most participants could only see the head, and for the right statue, some could see parts of the upper body or just the head. None of the users reported being able to see the guide. When estimating the number of users in the room, on average, participants estimated there were between 15 and 30 other users. In the first room, all participants described the distance to other users as comfortable. P11 mentioned, they "wouldn't like to go to the front in the middle, because it looks crowded." After the guide presented the first room, participants were asked about their behavior during the guide's presentation. Most users stated looking at the statue and others alternating between guide and statue.

Room 2: This room was square-shaped, slightly smaller than the first, and displayed the bright Achilles statue. The statue was taller and less wide than in the previous room, and users reported being able to see between 20-40% percent of it. The number of people in this room was estimated slightly lower, between 15 and 25 users. Most users still found the distance to other users comfortable.

In this room, several participants mentioned the *Scaling-Others* technique and how it worked better here because it made almost the entire statue visible. During the collaborative game, it became evident that users were comfortable with all techniques and were generally more positive about the techniques than in the search task. This could be due to the more natural task and the lack of time constraints. No clear favorite emerged from the interactive game.

Room 3: The third room was narrow and contained five broadly arranged busts. All users described the space as too crowded, with almost all participants stating that the person in front of them was too close. For instance, P19 expressed, "I'm kind of inside him, which is quite uncomfortable." P15 said, "Around me is still okay, but the user in front of me stands really close and tall in front of me, which is a bit annoying. In real life, I would take a step back."

In this room, there were a lot of positive comments about the *Self-Translation* technique, since it allowed users to avoid standing directly in the crowd and helped "peeking over the others" (P13) to view all the busts simultaneously from above. Overall, the comments largely echoed the questionnaire responses, and we encountered no issues conducting the entire interview within the VR environment.

5.2.4 Further Observations

During the interview and the collaborative game, it was observed that participants frequently adopted the technique employed by the conductor, particularly when the conductor was elevated with the *Self-*

Translation technique. Moreover, we observed that users rarely used the visibility techniques during conversations, and rarely directed them at the conductor, so that inconsistencies due to scaling or displacement occurred seldom. Since users themselves can not perceive the effect of the visibility techniques used by others, they were often unaware of the techniques and their influence on themselves. Our pilot study showed that minor inconsistencies in conversations often went unnoticed until users gestured or attempted virtual interaction; this led to uncomfortable situations such as unknowingly reaching into another user's face. However, incoherent representations often only became apparent when they were addressed verbally.

6 Discussion

In this section, we discuss the change in preferences and the overall identified advantages and disadvantages of our techniques.

6.1 Change in Preference

All participants successfully completed the search task using every technique in our study. The ranking after the first part clearly showed that *Transparency* is the preferred technique for finding features quickly. However, in a more realistic scenario, more than 50% of the participants reconsidered and prioritized a different technique.

We think that the re-prioritization has several reasons. First, while the *Transparency* technique was the fastest and virtual museum visitors certainly want to efficiently and effectively operate the available tools, there is typically no time pressure in a real museum tour. In addition, the interview showed that some of our techniques are better suited to certain situations than others. Finally, the order of the tour, which ended in a small room that favored the self-translation method, may have influenced the participants' rating ("recency effect").

6.2 Identified Advantages and Disadvantages

The different prioritization of techniques is already an indicator that some techniques are better suited for certain tasks than others.

Self-Translation, for instance, emerged as a technique offering an excellent overview of the surroundings. Users appreciated the ability to peek over other users, hover over crowded areas or to stay in elevated positions while listening to a tour guide. This was particularly evident during the guided tour, where participants reported an enhanced experience (5.2.3). However, it is important to note that in a virtual tour with real users, everyone would be able to adjust their height, which would most likely reduce this technique's advantage. Based on the results, we conclude that the *Self-Translation* technique is most suitable for experienced VR users in museum scenarios featuring medium-height content in all directions.

Transparency was favored for its ease of use and speed, making it the preferred choice for searching tasks. However, qualitative feedback revealed that it was perceived as somewhat unnatural. For a real museum application, this technique's cone size should ideally be quickly adjustable as participants expressed a desire for a larger viewing area. In contrast to the work *SocialSlider* [50], which reported that full transparency can potentially lead to isolation, we may have been overly cautious in ensuring users are always visible by using partial transparency. Nonetheless, we are confident that the flashlight metaphor is the right tool for users to select their point of interest while staying attuned to the social context.

Despite its initial learning curve and control challenges, the *Displacing-Others* technique was seen as natural by participants due to its resemblance to real-world behavior of stepping aside. Notably, it is the sole technique that achieves complete vertical clearance of occluders by horizontally pushing other avatars aside. Although receiving the lowest ranking, we believe that this technique's underlying concept is strong and that its perceived naturalness holds promise. It is worth noting, however, that there is room for improvement, particularly in reducing the movement of occluding avatars, which lead to user frustration and restlessness during searches. We believe that combining this technique with the transparency technique could balance the disadvantages of both methods and take advantage of their respective benefits.

Lastly, *Scaling-Others* effectively reduced occlusion, provided a good overview and visibility, and had the advantage of making crowds appear smaller, which reportedly improved comfort. In this context, we suggest further tests with real users to evaluate the effect of small avatars in social settings. Like the other flashlight techniques, *Scaling-Others* would benefit from an adjustable cone. Notably, since users are only scaled down and never moved away, other avatars could potentially always obscure lower parts of an exhibit with this technique.

7 Conclusion

Our paper presents four techniques designed to enhance users' views of objects in environments crowded with many people by partially adjusting local user formations or representations. Three techniques use a flashlight metaphor to obtain a clear view of an object. Obscuring user avatars are either pushed aside, scaled small, or made partially transparent. The fourth technique allows users to move up to look over the users in front of them. Our user study with 24 participants found the transparency technique to be superior to the other techniques for fast search tasks. In our qualitative part which included a semi-structured interview in a realistic museum setting, the advantages and disadvantages of each technique became apparent as users selected techniques depending on the situation. In a final ranking, the vertical position adjustment and transparency techniques were the most popular, but the scaling and translation techniques were found to be more natural than the others.

In future iterations, the flashlight techniques could also use an implicitly or automatically specified cone that aims towards points of interest (POI) that are currently being explained by a guide or simply appear in the viewing direction. These cones could better consider the shape and size of a POI. However, museum visitors would require possibilities to switch between implicit and explicit techniques since they might want to just look around and focus on other things at certain times. In certain situations, a combination of visibility techniques may prove most effective. For instance, moving up might generate an ideal perspective, but tall individuals in the front could still obstruct the sight and one of the flashlight techniques would be needed to clear the view.

Our quantitative study was limited to individual participants with otherwise simulated users in the room. Our museum tour involved at least two people and elicited valuable comments and suggestions. First observations indicate that local adjustments to other users' avatars are perceived differently in certain situations. However, to further investigate the influence of our techniques on social relationships within groups and conversations among participants we need to invite larger groups into our user study, ideally classroom size groups of pupils who are among the target audience. In such scenarios, we would like to study the effect of our techniques, and new ones such as placing everyone in the front row, on social presence and how social presence could be fostered despite the change of appearance or local adjustments of other avatars. In this context, future research could aim to identify the conditions under which inconsistent representations become noticeable and how to conceal them. Our techniques and our studies deliver first insights into solutions that explore the space between imitating the real world and what is possible in virtual reality.

Acknowledgments

This work is funded by the German Federal Ministry of Education and Research (BMBF) under the grant 16SV8716 (Goethe-Live-3D) and the Thuringian Ministry for Economic Affairs, Science and Digital Society under grant 5575/10-5 (MetaReal).

References

[1] M. Al Zayer, P. MacNeilage, and E. Folmer. Virtual locomotion: a survey. *IEEE transactions on visualization and computer graphics*, 26(6):2315–2334, 2018.

[2] F. Argelaguet, C. Andujar, and R. Trueba. Overcoming eye-hand visibility mismatch in 3d pointing selection. In *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology*, VRST '08, page 43–46, New York, NY, USA, 2008. Association for Computing Machinery.

[3] F. Argelaguet, A. Kulik, A. Kunert, C. Andujar, and B. Froehlich. See-through techniques for referential awareness in collaborative virtual reality. *International Journal of Human-Computer Studies*, 69(6):387–400, 2011.

[4] U. Assarsson, N. Elmqvist, and P. Tsigas. Image-space dynamic transparency for improved object discovery in 3d environments. *Technical report*, 2006.

[5] S. Benford, J. Bowers, L. E. Fahlén, C. Greenhalgh, and D. Snowdon. User embodiment in collaborative virtual environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 242–249, 1995.

[6] G. Bertolini and D. Straumann. Moving in a moving world: a review on vestibular motion sickness. *Frontiers in neurology*, 7:14, 2016.

[7] A. E. Bier, C. M. Stone, K. Pier, W. Buxton, and D. T. DeRose. Toolglass and magic lenses: The transparent interface. In *Proceedings of Siggraph '93 (Anaheim, August), Computer Graphics Annual Conference Series, ACM*, page 73, 1993.

[8] P. Bimberg, T. Weissker, A. Kulik, and B. Froehlich. Virtual rotations for maneuvering in immersive virtual environments. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, pages 1–10, 2021.

[9] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. An introduction to 3-d user interface design. *Presence*, 10(1):96–108, 2001.

[10] C. G. Christou and P. Aristidou. Steering versus teleport locomotion for head mounted displays. In *Augmented Reality, Virtual Reality, and Computer Graphics: 4th International Conference, AVR 2017, Ugento, Italy, June 12-15, 2017, Proceedings, Part II 4*, pages 431–446. Springer, 2017.

[11] E. F. Churchill and D. Snowdon. Collaborative virtual environments: an introductory review of issues and systems. *virtual reality*, 3:3–15, 1998.

[12] J. Clifton and S. Palmisano. Effects of steering locomotion and teleporting on cybersickness and presence in hmd-based virtual reality. *Virtual Reality*, 24(3):453–468, 2020.

[13] S. Cmentowski, A. Krekhov, and J. Krüger. Outstanding: A multi-perspective travel approach for virtual reality games. In *Proceedings of the annual symposium on computer-human interaction in play*, pages 287–299, 2019.

[14] C. Coffin and T. Hollerer. Interactive perspective cut-away views for general 3d scenes. In *3D User Interfaces (3DUI'06)*, pages 25–28. IEEE, 2006.

[15] T. J. Dodds and R. A. Ruddle. Using teleporting, awareness and multiple views to improve teamwork in collaborative virtual environments. In *Virtual Environments 2008*, pages 81–88. Eurographics Association, 2008.

[16] F. El Jamiy and R. Marsh. Survey on depth perception in head mounted displays: distance estimation in virtual reality, augmented reality, and mixed reality. *IET Image Processing*, 13(5):707–712, 2019.

[17] N. Elmqvist. Balloonprobe: Reducing occlusion in 3d using interactive space distortion. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 134–137, 2005.

[18] N. Elmqvist, U. Assarsson, and P. Tsigas. Dynamic transparency for 3d visualization: design and evaluation. *International Journal of Virtual Reality*, 8(1):75–88, 2009.

[19] 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.

[20] M. R. Endsley. Measurement of situation awareness in dynamic systems. *Human factors*, 37(1):65–84, 1995.

[21] M. T. Eren and S. Balcisoy. Evaluation of x-ray visualization techniques for vertical depth judgments in underground exploration. *The Visual Computer*, 34(3):405–416, 2018.

[22] C. N. Gunawardena and F. J. Zittle. Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *American Journal of Distance Education*, 11(3):8–26, 1997.

[23] C. Gutwin and S. Greenberg. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)*, 11:411–446, 2002.

[24] E. T. Hall. *The hidden dimension*. Anchor Books. Doubleday, New York, repr edition, 1990.

[25] S. Knödel, M. Hachet, and P. Guitton. Navidget for immersive virtual environments. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pages 47–50, 2008.

[26] A. Krekhov, S. Cmentowski, K. Emmerich, M. Masuch, and J. Krüger. Gullivr: A walking-oriented technique for navigation in virtual reality games based on virtual body resizing. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, pages 243–256, 2018.

[27] E. Kruijff, J. E. Swan, and S. Feiner. Perceptual issues in augmented reality revisited. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pages 3–12. IEEE, 2010.

[28] A. Kulik, A. Kunert, S. Beck, C.-F. Matthes, A. Schollmeyer, A. Kreskowski, B. Fröhlich, S. Cobb, and M. D'Cruz. Virtual valcamonica: collaborative exploration of prehistoric petroglyphs and their surrounding environment in multi-user virtual reality. *Presence: Teleoperators and Virtual Environments*, 26(3):297–321, 2017.

[29] J. R. Lackner. Motion sickness: more than nausea and vomiting. *Experimental brain research*, 232:2493–2510, 2014.

[30] L. Lisle, F. Lu, S. Davari, I. A. Tahmid, A. Giovannelli, C. Llo, L. Pavanatto, L. Zhang, L. Schlueter, and D. A. Bowman. Clean the ocean: An immersive vr experience proposing new modifications to go-go and wim techniques. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 920–921. IEEE, 2022.

[31] M. Maslych, Y. Hmaiti, R. Ghamandi, P. Leber, R. K. Kattoju, J. Belga, and J. J. LaViola. Toward intuitive acquisition of occluded vr objects through an interactive disocclusion mini-map. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pages 460–470. IEEE, 2023.

[32] F. J. Massey Jr. The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78, 1951.

[33] D. Medeiros, M. Sousa, A. Raposo, and J. Jorge. Magic carpet: Interaction fidelity for flying in vr. *IEEE transactions on visualization and computer graphics*, 26(9):2793–2804, 2019.

[34] H. Oh and W. Son. Cybersickness and its severity arising from virtual reality content: A comprehensive study. *Sensors*, 22(4):1314, 2022.

[35] T. Piumsomboon, A. Dey, B. Ens, G. Lee, and M. Billinghurst. The effects of sharing awareness cues in collaborative mixed reality. *Frontiers Robotics AI*, 6, 2019.

[36] T. Piumsomboon, G. A. Lee, and M. Billinghurst. Snow dome: A multi-scale interaction in mixed reality remote collaboration. In *Extended abstracts of the 2018 CHI conference on human factors in computing systems*, pages 1–4, 2018.

[37] R. Prentice, A. Davies, and A. Beeho. Seeking generic motivations for visiting and not visiting museums and like cultural attractions. *Museum management and curatorship*, 16(1):45–70, 1997.

[38] J. Reinhardt and K. Wolf. Go-through: Disabling collision to access obstructed paths and open occluded views in social vr. In *Proceedings of the Augmented Humans International Conference*, pages 1–10, 2020.

[39] E. Schott, E. B. Makled, T. J. Zoeppig, S. Muehlhaus, F. Weidner, W. Broll, and B. Froehlich. Unitexr: Joint exploration of a real-world museum and its digital twin. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*, pages 1–10, 2023.

[40] J. Short, E. Williams, and B. Christie. *The social psychology of telecommunications*. Toronto; London; New York: Wiley, 1976.

[41] A. Singla, S. Fremerey, W. Robitza, and A. Raake. Measuring and comparing qoe and simulator sickness of omnidirectional videos in different head mounted displays. In *2017 Ninth international conference on quality of multimedia experience (QoMEX)*, pages 1–6. IEEE, 2017.

[42] R. Stoakley, M. J. Conway, and R. Pausch. Virtual reality on a wim: in-

teractive worlds in miniature. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 265–272, 1995.

[43] G. A. Strizek, S. Tourkin, and E. Erberber. Teaching and learning international survey (talis) 2013: Us technical report. nces 2015-010. *National Center for Education Statistics*, 2014.

[44] The Audience Agency. Museums audience report: What Audience Finder says about audiences for museums, 2018. `https://www.theaudienceagency.org/asset/1995`.

[45] K. Vasylevska, H. Kaufmann, and V. Khrystyna. Influence of vertical navigation metaphors on presence. In *Challenging Presence-Proceedings of 15th International Conference on Presence (ISPR 2014)*, pages 205–212, 2014.

[46] L. Wang, J. Wu, X. Yang, and V. Popescu. Vr exploration assistance through automatic occlusion removal. *IEEE transactions on visualization and computer graphics*, 25(5):2083–2092, 2019.

[47] Y. Wang, Z. Hu, S. Yao, and H. Liu. Using visual feedback to improve hand movement accuracy in confined-occluded spaces in virtual reality. *The Visual Computer*, 39(4):1485–1501, 2023.

[48] T. Weissker, P. Bimberg, and B. Froehlich. Getting there together: Group navigation in distributed virtual environments. *IEEE transactions on visualization and computer graphics*, 26(5):1860–1870, 2020.

[49] T. Weissker, P. Bimberg, A. S. Gokhale, T. Kuhlen, and B. Froehlich. Gaining the high ground: Teleportation to mid-air targets in immersive virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2467–2477, 2023.

[50] A. E. Wolf, J. Reinhardt, M. Kurzweg, and K. Wolf. Socialslider: Changing the transparency of avatars. In *Proceedings of the Augmented Humans International Conference 2022*, pages 276–283, 2022.

[51] R. F. Woolson. Wilcoxon signed-rank test. *Wiley encyclopedia of clinical trials*, pages 1–3, 2007.

[52] D. Yu, Q. Zhou, J. Newn, T. Dingler, E. Velloso, and J. Goncalves. Fully-occluded target selection in virtual reality. *IEEE transactions on visualization and computer graphics*, 26(12):3402–3413, 2020.