Occlusion Management in VR: A Comparative Study

Lili Wang *[1], Han Zhao[2], Zesheng Wang[2], Jian Wu[1], Bingqiang Li[2], Zhiming He[2], and Voicu Popescu[3]

[1]State Key Laboratory of Virtual Reality Technology and Systems,
Beijing Advanced Innovation Center for Biomedical Engineering, Beihang University, China
[2]Qingdao Research Institute of Beihang University, China
[3]Department of Computer Science, Purdue University, United States

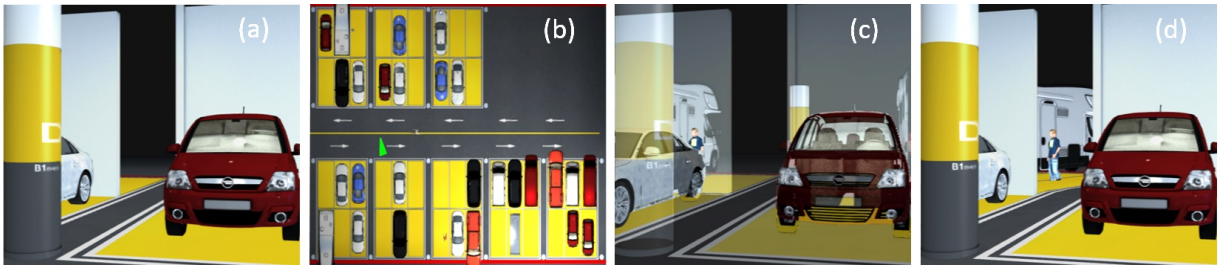

Figure 1: The task of finding a person in a garage, using conventional visualization (a), top view visualization (b), X-ray visualization (c), and multiperspective visualization (d).

Abstract

VR applications rely on the user's ability to explore the virtual scene efficiently. In complex scenes, occlusions limit what the user can see from a given location, and the user has to navigate the viewpoint around occluders to gain line of sight to the hidden parts of the scene. When the disoccluded regions prove to be of no interest, the user has to retrace their path, making scene exploration inefficient. Furthermore, the user might not be able to assume a viewpoint that would reveal the occluded regions due to physical limitations, such as obstacles in the real world hosting the VR application, viewpoints beyond the tracked area, or viewpoints above the user's head that cannot be reached by walking. Several occlusion management methods have been proposed in visualization research, such as top view, X-ray, and multiperspective visualization, which help the user see more from the current position, having the potential to improve the exploration efficiency of complex scenes.

This paper reports on a study that investigates the potential of these three occlusion management methods in the context of VR applications, compared to conventional navigation. Participants were required to explore two virtual scenes to purchase five items in a virtual Supermarket, and to find three people in a virtual parking garage. The task performance metrics were task completion time, total distance traveled, and total head rotation. The study also measured user spatial awareness, depth perception, and simulator sickness. The results indicate that users benefit from top view visualization which helps them learn the scene layout and helps them understand their position within the scene, but the top view does not let the user find targets easily due to occlusions in the vertical direction, and due to the small image footprint of the targets. The X-ray visualization method worked better in the garage scene, a scene with a few big occluders and a low occlusion depth complexity, and less well in the Supermarket scene, a scene with many small occluders that create high occlusion depth complexity. The multiperspective visualization method achieves better performance than the top view method and the X-ray method, in both scenes. There are no significant differences between the three methods and the conventional method in terms of spatial awareness, depth perception, and simulator sickness.

Keywords: Scene Exploration, Occlusion Management, Virtual Reality, Top View, X-ray, Multiperspective Visualization

Index Terms: Human-centered computing—Human-centered interaction—Virtual reality; Computer graphics—Oclusion management—Visualization

1 Introduction

When users wear tracked head-mounted displays (HMDs) to explore scenes in virtual reality (VR) applications, they naturally choose the desired view by walking and by rotating their head. However, the efficiency of such VR scene exploration is limited by occlusions and physical space constraints. Consider a shopping application, where the goal is to find a specific shelf or item. Shelves block each other and the user has to move a considerable amount to check a shelf that is occluded in the current view. If the item to be purchased is not on the first hidden shelf, the user has to continue to check the other shelves one at the time. Such sequential exploration is inefficient. In a second application where the goal is to locate a person in a parking garage, the vehicles and the garage walls limit what can be seen from any given viewpoint, and translating the viewpoint around these occluders is, again, inefficient. When the person to be located moves, sequential exploration is ineffectual.

Occlusions have been addressed in VR with several approaches, including top-view visualization, where the user relies on a comprehensive view of the scene rendered from above, X-ray visualization, which renders occluding layers transparently to reveal the scene behind them, exploded view visualization, where the scene elements are displaced away from each other, cutaway visualization where the occluding layers are discarded, and multiperspective visualization, where samples are captured from additional viewpoints and integrated into the output image.

In this paper we present a user study that compares three VR occlusion management methods to conventional visualization. The

*e-mail: wanglily@buaa.edu.cn

2019 IEEE Conference on Virtual Reality and
3D User Interfaces
23-27 March, Osaka, Japan
978-1-7281-1377-7/19/$31.00 ©2019 IEEE

three methods are top-view visualization, X-ray visualization, and multiperspective visualization. We have chosen these three methods for good coverage of the space of all possible occlusion management methods, while keeping the complexity of the user study manageable. Indeed, cutaway visualization can be seen as a special case of X-ray visualization where the opacity of the occluders is set to zero, and exploded visualization can be seen as a discontinuous form of multiperspective visualization, as any multiperspective image can be obtained by distorting the scene and by rendering the distorted scene conventionally.

Participants (n = 32) were asked to perform two tasks: finding five items in a Supermarket, and finding three persons in a parking garage (see Fig. 1), one at the time. The participants were randomly assigned to four groups. One group used top-view visualization, one group used X-ray visualization, and one group used multiperspective visualization, as they performed the two tasks. The fourth group used only conventional visualization, serving as control. The participants navigated the scene using a tracked HMD, and selected objects with a hand-held controller. Task performance efficiency was quantified using the following metrics: completion time, total distance traveled, and total view direction rotation. We also measured spatial awareness by asking participants to point to items seen earlier during scene exploration. We measured depth perception by asking participants to estimate the distance to objects visible in the various visualizations. Finally we measured simulator sickness using the standard Simulator Sickness Questionnaire(SSQ) [17].

For the first task, the participants in the top-view and the X-ray groups did not significantly outperform the participants in the control group. The group in the multiperspective visualization condition outperformed the control group in completion time, total viewpoint translation distance, and total view direction rotation. For the second task, the top-view method did not show a significant advantage over conventional visualization, while both X-ray and multiperspective did; the advantages brought by multiperspective visualization were more substantial than those brought by X-ray visualization. There were no significant differences between the three occlusion management methods and conventional visualization in terms of spatial awareness, depth perception, and simulator sickness.

To the best of our knowledge, our study is the first to directly compare approaches for occlusion management in HMD VR scene exploration. Our study compares four approaches: top-view, X-ray, multiperspective, and conventional visualization. The occlusion management techniques are compared against conventional visualization which acts as control. Occlusion management is an important approach for increasing the efficiency of VR exploration, and for overcoming the physical limitations of the real world hosting the VR application. In light of the good performance of the multiperspective approach, our paper makes the high level contribution of indicating that there are tasks of practical importance where an overt but well controlled departure from realism can benefit HMD VR visualization without inducing user disorientation or simulator motion sickness.

2 Prior Work

In VR applications, actual walking is found to be more intuitive for users [30, 41]. However, exploring complex virtual environments by walking is inefficient due to occluders that have to be circumvented one at the time in search of regions of interest (ROIs). Occlusion management techniques let the user see what is behind an occluder from the current viewpoint, without additional navigation. If the occluder does not hide an ROI, the unnecessary navigation is avoided. Furthermore, When circumventing the occluder would have required moving beyond the boundaries of the physical space hosting the application, occlusion management avoids unnecessary teleportation or walking redirection.

The most commonly used occlusion management approach is to rely on the user to find the ROIs through interactive navigation. Interactive navigation efficiency can be increased by providing the user with visual hints that indicate the path to be taken to reveal ROIs [3, 11], but the hint does not provide a preview of the ROI, and the user still has to change the view to gain sight of the ROI.

Another approach for alleviating the problem of occlusions is to provide the user with a top view of the scene. The top view removes occlusions for 2.5D scenes, i.e. height fields, and it provides global information of the scene structure as well as the current location of the user [24, 38]. In HMD visualization for VR, the top-view approach has the disadvantage that it switches from the first-person view to the third-person view, which breaks the sense of immersion. Furthermore, many scenes cannot be disoccluded by a top view, when there are multiple horizontal occluding layers. Finally, many tasks require visualizing the scene with a horizontal view direction, such as recognizing a particular person, and they cannot be performed well relying solely on the top view.

X-ray methods render the occluders semi-transparently and expose the ROIs to the user directly [2, 16, 31]. Optimized opacity visualization has also been applied to render feature lines of dense 3D line fields [12]. The X-ray method achieves transparency by rendering image layers at different depths and then blending all the layers together. Although X-rays method weaken depth cues, a forte of HMD visualization, they provide a convenient way for users to view areas behind occluders. The X-ray occlusion management method is more suitable for scenes where occluders and ROIs, possibly dynamic ROIs, are predefined. For example, in games, the user has the option of a "special vision mode" to render specific occluders transparently to reveal known ROIs, such as a person behind a wall [29]. Unlike the top-view method, X-ray visualization remains first-person, which sustains the sense of immersion. One limitation of X-ray methods is that they do not scale well with the number of occluding layers, as the transparent visualization becomes hard to understand. Furthermore, when the ROIs are not known a priori, the visualization could inadvertently show the ROI transparently and the user has to adjust the depth of the transparency effect.

The cutaway method removes the occluders in front of the ROIs [6, 21], which comes at the cost of not providing context for the ROIs, which makes it difficult to understand where the ROIs are actually located with respect to the user's current position. Explosion methods [5,15,20] divide the scene into sections that are moved away from the ROIs to reduce occlusion. Explosion methods improve exploration efficiency but this comes at the cost of significantly perturbing the scene geometry, at topological level, which results in visualization discontinuities that hinder user spacial perception and virtual scene immersion.

The benefit of multiple perspectives has been noted early on, for example in the Super Mario 64 game which allows the player to select a second player's view [42]. Unlike a top view, the additional first-person view allows the user to preview what they would see if they were at a different location, without the cost of actual navigation for assuming that view. These earlier approaches would not integrate the multiple perspectives into a single, non-redundant and continuous image, so the user was left with examining the individual perspectives one at the time.

Researchers in vision, visualization, and graphics have developed non-traditional camera models that overcome the limitations of the traditional pinhole camera, including the line of sight limitation. The push-broom camera [13], the multi-center of projection [26] and the general linear camera models [48] have rays that do not pass through a common point, and they provide disocclusion capability. Multiperspective cameras are constructed to integrate multiple viewpoints, which mitigates occlusions and increases the information bandwidth of the resulting image. The occlusion camera [23] relaxes the single viewpoint constraint with piecewise linear rays that converge towards the occluder center to reveal samples that are not visible from the original viewpoint. The graph camera [25] generates a continuous and non-redundant multiperspective image by combin-

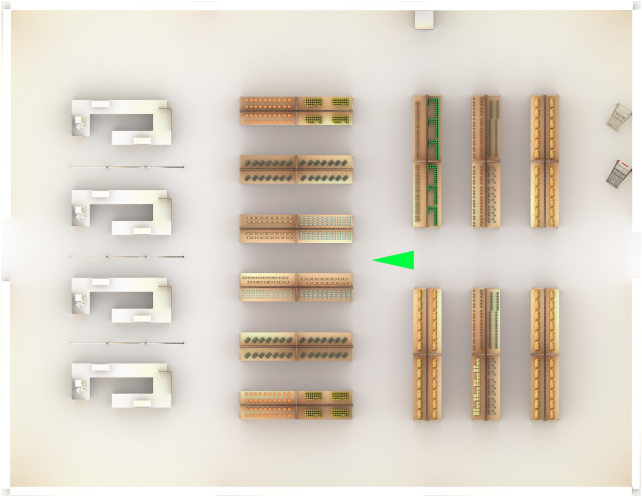

Figure 2: Top-view visualization for Supermarket scene.

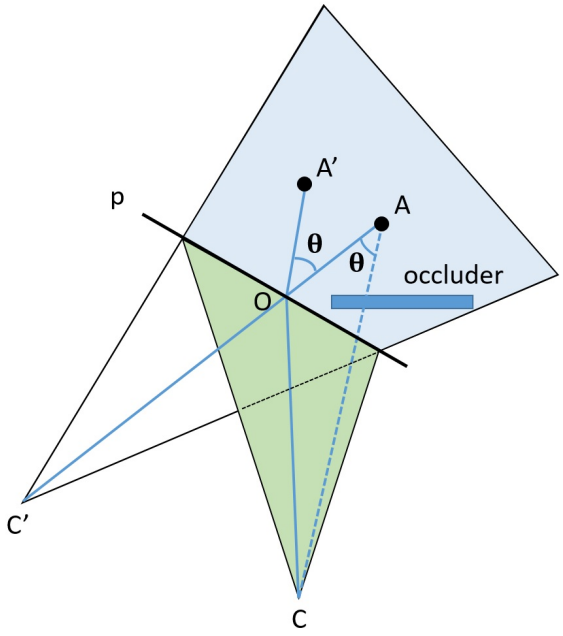

Figure 3: Construction of graph camera underlying our multiperspective visualization.

ing multiple conventional cameras in an acyclical graph. The graph camera has been extended to achieve view frustum splitting while preserving visualization continuity [44]. The graph camera has been used to overcome occlusions in HMD visualization for VR and AR, where it has proven to increase scene exploration efficiency [45]. Like the X-ray approaches, the multiperspective approach lets the user see the scene with a horizontal view direction, which is beneficial when the ROIs distinguishing features are not apparent in a top view. Unlike the X-ray approach, multiperspective visualization does not blend multiple scene regions together; instead, the mapping of the scene to the image is rearranged to accommodate all regions without overlap, which comes at the cost of global distortions of the user's natural view of the scene.

Since occlusion management also helps with the problem of mapping the larger virtual space to the more constrained physical space hosting the application, we briefly discuss other prior solutions of the same problem. Local walking systems based on treadmills [33] and inertial sensing hardware [39] were used to address the limitations of physical space size. To accommodate the difference between the physical space and the much larger virtual space, several techniques have been developed, including redirected walking [27, 28], variable scene scaling [19], variable translation gain [47], pose resetting [43], and non-linear mapping from physical to virtual space [37]. The challenge is to design a mapping between the real and virtual space that is both effective at bridging substantial size differences between the two, and that is not noticeable by the user [34, 35]. The latest redirection methods aim to dissimulate the incongruence between the real and the virtual space by applying the manipulation in moments when the user is the least perceptive. For example, the virtual scene is rotated abruptly when the user blinks [18] or changes fixation point [36], approaches that require eye tracking.

3 Occlusion Management Methods

In this section we describe the three occlusion management methods used in our user study: top view (Section 3.1), X-ray (Section 3.2), and multiperspective visualization (Section 3.3).

3.1 Top View Visualization

We render the top view with an orthographic projection along the vertical direction. The user position and orientation is indicated with a green triangle (Fig. 1(b), Fig. 2). The top view provides a comprehensive visualization of the scene. When the ROIs are easily identifiable from above, the user can build a mental model of their location, relative to the user's current location. However, when the ROIs cannot easily be identified from above, or when they are hidden underneath horizontal occluders, the top view could be ineffective. An additional potential challenge of top-view visualization stems from the limited resolution of the HMD. Showing the top-view overlaid on a corner of the first-person view might not provide sufficient detail. Switching between the first person view and the top view breaks the sense of immersion, and complicates the interface.

3.2 X-ray Visualization

X-ray visualization renders occluders semi-transparently, with the goal of exposing occluded ROIs. We implement X-ray visualization with a depth peeling approach [9] that blends the k closest occluding layers. In our experiments, the user controls selects a k between 1 and 2 (Fig. 1(c)). A single occluding layer is blended with the opaque layer behind it with weights 40%, 60%, respectively. For two occluding layers the weights we use are 16%, 24%, and 60%.

3.3 Multiperspective Visualization

Multiperspective visualization removes the single viewpoint constraint of the traditional camera. The resulting multiperspective image integrates continuously and non-redundantly samples from multiple viewpoints, alleviating occlusions. Our multiperspective visualization is based on the graph camera [25]. The graph camera is built starting from a conventional camera frustum to which bending, splitting, and merging operations are applied recursively. For our experiment, we limit the graph camera to a single bending operation, which has the ability to reveal ROIs that are visible from a secondary viewpoint, without a significant increase in visualization complexity. A multiperspective image accommodates multiple scene regions simultaneously, which comes at the cost of a smaller number of pixels per region, and HMD resolution limitations restrict the number of regions that can be shown effectively, at the same time.

Our graph camera construction is illustrated in Fig. 3. Given a pinhole camera that models the primary user view with viewpoint

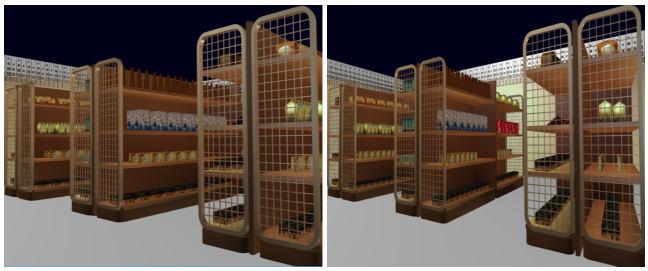

Figure 4: Conventional (left) and multiperspective (right) visualization for the Supermarket scene. The multiperspective visualization lets the user inspect, from the current position, the far part of the shelf, to which the user does not have direct line of sight.

C, a secondary viewpoint C′, and a viewpoint transition plane p, the resulting camera has two sub-frusta, with two-segment piece-wise linear rays. The camera allows the user located at C to see the red target to which there is no line of sight. We render the scene with the graph camera by rotating geometry vertices in the secondary sub-frustum (blue) around the center O(x_0, y_0) of the part of p inside the frustum. For example, a point A is rotated in the plane ACC′ with Eqs. 1 and 2, where (x_1, y_1) are the 2D coordinates of A in the plane ACC′, and the angle θ is determined by the angle between AC and AC′. Alternatively, one could use the original graph camera projection which first computes the intersection between C′A and P, and then uses this intersection to define the ray from C on which A′ is located [25].

$$x_2 = (x_1 - x_0)\cos\theta - (y_1 - y_0)\sin\theta + x_0 \quad (1)$$

$$y_2 = (x_1 - x_0)\sin\theta + (y_1 - y_0)\cos\theta + y_0 \quad (2)$$

In our experiments, the hinge planes p are predefined for each scene. When the disocclusion effect is activated, the user slides the intermediate viewpoint from C to C′ by translating their head laterally. This way the disocclusion effect is deployed gradually and intuitively. Furthermore, the part of the scene closest to the user, i.e. the geometry located inside the green primary frustum in Fig. 3, is rendered from the user perspective, which anchors the user and alleviates disorientation [46]. In Fig. 4 the multiperspective visualization reveals the entire shelf, from the current position of the user, without the requirement that the user walk up to the isle. In Fig. 1(d) the user can examine the part of the garage hidden by the wall, again, without any physical locomotion.

4 Experimental Design

We have conducted a user study to compare occlusion management based on top-view, X-ray, and multiperspective visualization with conventional visualization. The comparison metrics are task completion time, total viewpoint translation, total view direction rotation, spatial awareness, depth perception, and simulator sickness. We recruited 32 participants from our university's graduate student population. They were between the ages of 22 and 26, 6 participants were women, and 10 participants had prior experience with VR applications. We randomly assigned 8 participants to each of the four conditions: top view, X-ray, multiperspective, and conventional visualization.

4.1 Experimental Setup

The experiment was conducted in an empty room, using an HTC Vive Pro System with a wireless hand-held controller. The HMD is connected to a desktop PC (Intel i7 processor, 32GB RAM, and NVIDIA 1080 graphics card). The tracked area was 4m×4m. The virtual environment is rendered for each eye at 60fps. The hand controller is rendered so participants can see the direction in which the controller is pointing. We used two scenes: the Supermarket (2.7M triangles, 13m×14m Fig. 2) and the Garage (5.8M triangles, 56m×38m Fig. 1). To make up for the difference in size between the physical and the virtual space, participants could transfer directly from one location of the virtual space to another using a teleportation mechanism triggered with the controller. We used an approach based on the *friendly teleportation* method [4]: when a participant approaches a boundary of the physical space, a warning is issued, and the participant can choose to trigger a forward jump of 1m.

4.2 Tasks

Each participant performed two tasks. In the first task (VR1) participants were asked to purchase 5 specific items in the Supermarket. The items to be purchased were provided as a list that participants could invoke at any time using the controller. To purchase an item the user has to find the item, to move such that the item is within arm's reach, and then to touch the pad on the controller. Once an item is purchased, a check mark appears by the item in the list. For participants in a group that uses an occlusion management technique, the top-view, X-ray, or multiperspective visualization is turned on by pressing a button on the controller. If the item is disoccluded, the participant still has to go close to the item to purchase it, but occlusion management has the potential to avoid unnecessary navigation around an occluder that does not hide the item of interest. In the second task (VR2), participants were asked to find three persons in the Garage, one at the time. The person that is the target of the search is displayed for 10s at the beginning, and then the participant can invoke their image at any time using the controller. As for VR1, the participants in the experimental groups can activate a disocclusion effect.

4.3 Spatial Awareness Test

We tested spatial awareness, outside of the two tasks above, using a pointing test [22], as shown in Fig. 5. Initially, the participant is at position O, and only the green target is present. For the conventional visualization group, the participant is guided to position A, where they can see the target, and then they are guided to position C, where they are asked to point to the location of the target. Since the target is occluded, the participant must rely on spatial awareness. The angle between the correct direction to the target and the pointing direction is a measure of spatial awareness error. The procedure is repeated for the red target. For the occlusion management groups, the participant starts at O, they are guided to B, where they trigger their disocclusion effect, and then they are guided to C, where they cannot use the disocclusion effect, and, again, they have to point to the target based on spatial awareness.

4.4 Depth Perception Test

We tested depth perception, also outside of the two primary tasks, by asking the participant to estimate the distance between a red and a green ball [1]. The participant starts out and remains at position O throughout the depth perception test (Fig. 6). The red ball is always present, at the same position, 2m away from the participant's position O. A green ball appears, and the participant is asked to estimate the distance between the red ball and the green ball, in multiples of 2m. The green ball appears at one of five positions. For conventional visualization, the occluder is not present. For the occlusion management conditions, when the green ball appears in positions 4 or 5, the occluder prevents seeing the ball directly, and the participant has to rely on their disocclusion effect to estimate distance. Any incorrect answer, no matter how far off, is counted as one depth perception error. The procedure is repeated 5 times, with the green ball location chosen randomly from the five options.

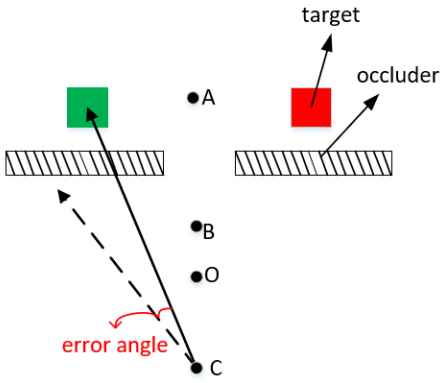

Figure 5: Spatial awareness estimation based on pointing direction accuracy. The participant has to point from memory to the occluded red and green targets.

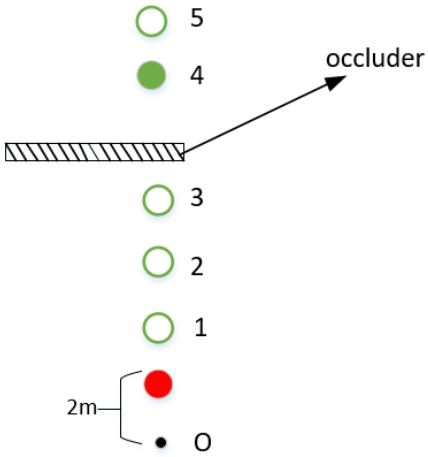

Figure 6: Overview of depth perception test. The participant has to estimate the distance between the red ball and the green ball. Participants in the experimental groups have to rely on their disocclusion effect when the green ball appears in positions 4 or 5.

4.5 Simulator Sickness Assessment

It has been long known that Wearing a VR HMD can cause simulator sickness [14]. We asked participants to self-assess simulator sickness using the standard Simulator Sickness Questionnaire (SSQ) [17]. Participants completed the SSQ before and after the experiment. The questionnaire contains a total of 16 questions, covering fatigue, headache, eye strain, sweating, nausea, and so on. There are four possible answers to each question: none, slight, moderate, and severe, which are converted to the numerical scores of 0, 1, 2, and 3.

4.6 Experimental Procedure

We followed the following experimental procedure.

step 1 The participant completes the SSQ before the experiment.

step 2 The experimental procedures are explained to the participant by an experimenter. This includes a description of how to wear the HMD, and a description of the function of the three controller buttons.

step 3 Under experimenter guidance, the participant practices the interaction with the Supermarket scene used in the shopping task, including the use of the HMD, teleportation, selection of items on shelves, and display of the shopping list. Participants in the experimental conditions also practice triggering their disocclusion effect. The participant is introduced to the first task. This practice/tutorial session lasts 5 minutes.

step 4 The participant performs the first task.

step 5 After 2 minutes of rest, the participant conducts the spatial awareness test.

step 6 After 2 minutes of rest, the participant practices interaction with the Garage scene used in the person finding task. The participant is introduced to the second task. This second practice/tutorial session lasts 5 minutes.

step 7 After 2 minutes of rest, the participant performs the depth perception experiment.

step 8 The participant completes the SSQ after the experiment.

The 2 minutes of rest separate the various parts of the experiment, for the comfort of the participant. For example, the spatial awareness test was run with no break between the exposure and test, i.e. the participant indicates the direction to the target as soon as the participant returns from position A. If a participant feels uncomfortable they can stop the experiment at any time. Out of the 32 participants, one participant in the top-view group complained of dizziness, his experiment was stopped, and his data was discarded.

Table 1: Task completion times, in seconds.

Task	Group	Avg± std.dev.	Red.	p	Cohen's d	Effect size
VR1	CG	530 ± 194				
	TG	523 ± 253	1%	0.95	0.03	VSmall
	XG	488 ± 154	8%	0.64	0.24	Small
	MG	434 ± 130	18%	0.26	0.58	Medium
VR2	CG	616 ± 224				
	TG	562 ± 212	9%	0.64	0.25	Small
	XG	406 ± 178	34%	0.06	1.0	Large
	MG	404 ± 151	34%	0.04	1.1	Large

VR1/2 : VR task 1/2; CG: conventional group; TG: top-view group; XG: X-ray group; MG: multiperspective group.

5 Results and Discussion

We measured and compared the VR scene exploration performance of our top-view (TG), X-ray (XG), multiperspective (MG) and conventional visualization groups (CG), using task performance metrics (Section 5.1), and perceptual metrics (Section 5.2). We analyze the data using the t-test p value, and Cohen's effect size d [7,32,49]. Effect sizes are qualified as "very small", "small", "medium", "large", "very large", and "huge" according to the d thresholds of 0.2, 0.5, 0.8, 1.2, and 2.0.

5.1 Task Performance Metrics

Table 1 gives the task completion times, Table 2 gives the total viewpoint translation, and Table 3 gives the total view direction rotation, for each of the four groups, and compares each of the experimental groups to the control group, for each of the two tasks. Column 4 of the tables, with heading "Red." gives the relative reduction brought by the occlusion management technique, for each metric. Overall, the occlusion management techniques brought more benefits for task VR2 than task VR1. For the Supermarket scene, the occlusion pattern is complex and the occlusion management techniques are harder to use effectively. Furthermore, the Supermarket is a relatively smaller scene, and the efficiency gains brought by avoiding

Table 2: Total viewpoint translation, in meters.

Task	Group	Avg± std.dev.	Red.	p	Cohen's d	Effect size
VR1	CG	198 ± 72				
	TG	194 ± 85	2%	0.93	0.04	VSmall
	XG	194 ± 50	2%	0.89	0.07	VSmall
	MG	142 ± 28	29%	0.06	1.0	Large
VR2	CG	617 ± 248				
	TG	445 ± 259	28%	0.21	0.68	Medium
	XG	329 ± 193	47%	0.02	1.3	VLarge
	MG	244 ± 75	60%	< 0.01	2.0	Huge

VR1/2 : VR task 1/2; CG: conventional group; TG: top-view group; XG: X-ray group; MG: multiperspective group.

Table 3: Total view direction rotation, as a multiple of full 360° rotations.

Task	Group	Avg± std.dev.	Red.	p	Cohen's d	Effect size
VR1	CG	57 ± 21				
	TG	47 ± 22	17%	0.42	0.43	Small
	XG	54 ± 14	5%	0.75	0.17	VSmall
	MG	37 ± 8.5	35%	0.03	1.2	VLarge
VR2	CG	79 ± 26				
	TG	49 ± 17	38%	0.02	1.4	VLarge
	XG	49 ± 24	37%	0.03	1.2	VLarge
	MG	37 ± 12	53%	< 0.01	2.1	Huge

VR1/2 : VR task 1/2; CG: conventional group; TG: top-view group; XG: X-ray group; MG: multiperspective group.

unnecessary navigation are less substantial. For the Garage scene, there are fewer occlusion layers, and the occlusion layers are farther apart, which simplifies removing occlusions effectively. Furthermore the Garage scene had 12 times the area of the Supermarket scene, and avoiding a long viewpoint translation just to find that the region beyond the occluder is empty translates to bigger savings.

The benefit brought by top-view visualization was small, and almost always the smallest among the three occlusion management techniques. In the Supermarket, it is impossible to identify using the top view items that are not located on the top shelf. In the Garage, the top view did not provide sufficient detail to identify the persons seen. The only exception is for the view direction rotation metric, where TG outperformed XG. The better performance for this metric is likely due to the fact that the top view allows the participant to gain a global understanding of the scene, without having to pan the view direction as much as the other three "ground-anchored" methods require. Whereas both of our scenes and our tasks are challenging for the top view method, the method could prove useful when the top view provides sufficient resolution and disocclusion to perform the task, such as, for example, locating a uniquely colored car in a 2.5D urban environment.

The X-ray visualization brought the second-most benefit from the three occlusion management techniques. In the Garage scene, X-ray visualization is quite effective, bringing significant improvements along the three metrics. The X-ray visualization works well in VR2 because of the small number of occluding layers, and because it provides a conventional, horizontal sampling of the scene that allows identifying the person visualized. X-ray visualization did not work well in the Supermarket scene. A specific challenge is that the depth peeling of the scene could end up breaking items apart, making them hard to identify. The X-ray method would benefit from a pre-processing of the scene that pre-defines the depth layers to optimize the search task. These results are in agreement with prior work which found that X-ray visualization is effective when the occluders are pre-defined.

Table 4: Pointing angle error, in degrees, used to quantify spatial awareness.

Task	Group	Avg± std.dev.	Red.	p	Cohen's d	Effect size
SA	CG	4.9 ± 2.7				
	TG	14 ± 16	-175%	0.15	0.77	Medium
	XG	4.5 ± 2.8	9%	0.75	0.16	VSmall
	MG	6.1 ± 4.8	-24%	0.56	0.30	Small

SA : spatial awareness; CG: conventional group; TG: top-view group; XG: X-ray group; MG: multiperspective group.

The multiperspective visualization brought the largest benefit, for all metrics, and for all tasks, and the p value was at most 0.26. The two scenes are favorable to multiperspective visualization since the occluders are large, inducing depth discontinuities that are leveraged by the multiperspective visualization constructors. Unlike the top-view method, the multiperspective visualization does not aim to show everything in a single image, hence the resolution issues are alleviated. The MG group too had a harder time with VR1 than with VR2. The complexity of the Supermarket made it that the items disoccluded would be represented with too few pixels to accelerate the search significantly. Since our tasks require that the participant not only see, but actually get close to the item or person sought, the disocclusion technique has not only to reveal the target, but also to provide clues for the user to reach it once the disocclusion effect is retracted. The multiperspective visualization provides context sampled horizontally, which easily leads to the target, whereas X-ray might not show the context clearly enough, and top-view might not show context in sufficient detail.

Figure 7 shows the VR2 trajectory of a participant in each of the four groups.The CG trajectory shows confusion, as the same region of the scene is visited multiple times, whereas some regions remain unexplored. The XG and MG trajectories focus on the main alley with occasional visits to the isles to collect the items seen from the main alley.

One concern is that the small number of participants might imply that the assumption of normality and homogeneity is not met. To address this concern, we checked for differences between the four different methods using a Friedman test [10]. For VR1, for task completion time $\chi^2(3) = 1.5$ and $p = 0.69$, and for total view-point translation $\chi^2(3) = 6.6$ and $p = 0.09$, which indicates that there are no significant differences between the four methods. However, there is a significant difference in the total view direction rotation ($\chi^2(3) = 9.0$, $p = 0.03$) between four methods. To further analyze the total view direction data, we ran Tukey's test [40] in a post hoc analysis, which shows that the difference is due to the difference between conventional navigation and multiperspective visualization ($p = 0.04$). For VR2, for task completion time $\chi^2(3) = 7.6$ and $p = 0.05$, which is borderline significant. For total view-point translation $\chi^2(3) = 10.7$ and $p = 0.01$, and for total view direction rotation $\chi^2(3) = 11.1$ and $p = 0.01$, which means that there is a significant difference between our four methods. The Tukey post hoc analysis reveals that there is a difference in total view-point translation between conventional and multiperspective visualization ($p = 0.04$), and between conventional and X-ray visualization ($p = 0.02$). The post hoc analysis also reveals that the difference in total view direction rotation is due to the difference between conventional and multiperspective visualization ($p = 0.01$).

5.2 Perceptual Metrics

Table 4 gives the pointing angle errors for the four groups. No difference was significant. Only X-ray visualization improved over conventional visualization, but the improvement is small, indicating

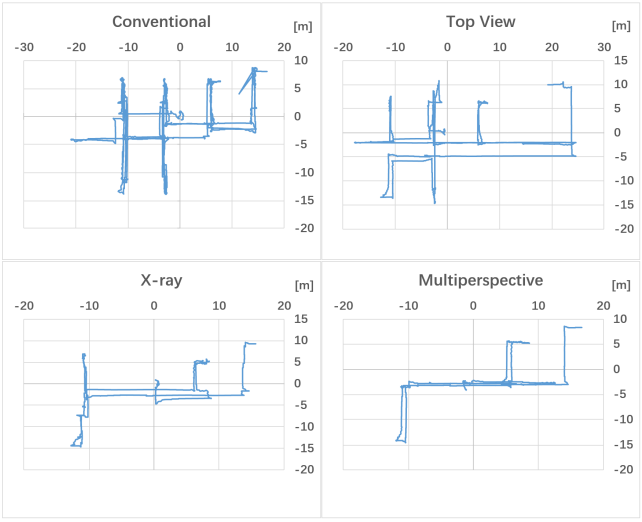

Figure 7: VR2 trajectories for CG (top left), TG (top right), XG (bottom left) and MG (bottom right) participants.

Table 5: Distance estimation errors, in times, used to quantify depth perception.

Task	Group	Avg± std.dev.	Red.	p	Cohen's d	Effect size
DP	CG	2.0±1.3				
	TG	1.6±1.4	22%	0.55	0.32	Small
	XG	1.6±1.3	19%	0.58	0.28	Small
	MG	2.6±1.1	-32%	0.32	0.52	Medium

DP: depth perception; CG: conventional group; TG: top-view group; XG: X-ray group; MG: multiperspective group.

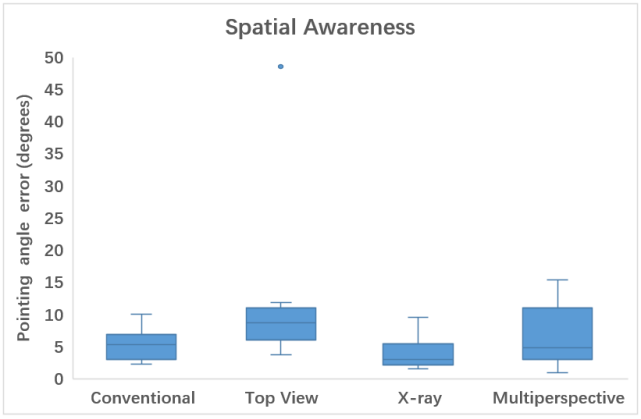

Figure 8: Boxplot of pointing angle errors for the four visualization methods. Dots show outliers that are beyond 1.5 times the interquartile range, which is shown with whiskers.

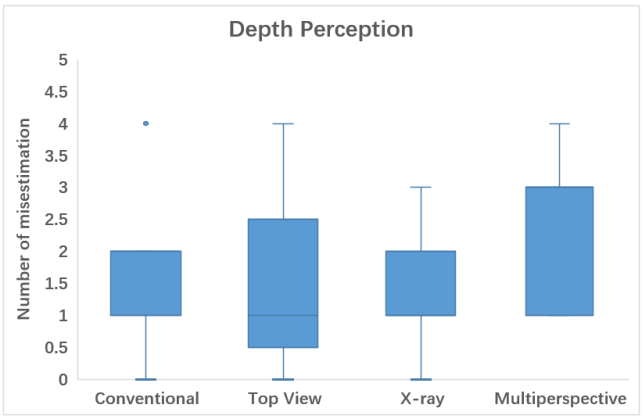

Figure 9: Boxplot of distance mis-estimating results in the depth perception test with the conventional method and three occlusion management methods.

that participants would remember the location of the target once the occluder became opaque again. Multiperspective visualization performed worse than conventional visualization, but not significantly. The multiperspective visualization used relies on a hinge plane to rotate the occluded target into view, which prevents the participant from learning the correct location of the target. Multiperspective visualization techniques that show a target where it would be seen in the absence of occluders have been developed [44], which could be used once the target is found and the visualization switches from browsing to tracking mode, placing the target at its correct position in the image. Top-view visualization performed the worst, as switching to the 2D view with a vertical view direction to a 3D view, with a horizontal confused participants, with one completely losing orientation, see the outlier in Fig. 8. Overall, the angle errors are small for all groups, and future tests should probably test spatial awareness on more challenging tasks.

Table 5 gives the number of distance estimation errors for the four groups. No difference was significant. The worst performer was multiperspective visualization (see also Fig. 9), which is likely due to moving the green ball to disocclude it, and, as explained above, a technique that shows the green ball in its correct position relative to the user might help.

Regarding simulator sickness, the mean Total Severity (TS) score for the pre-experiment SSQ was 4.7 (SD = 5.9) for the CG participants, while the mean TS score for the post-experiment SSQ was 13.6 (SD = 6.3). TS values were calculated using the formula specified by Kennedy et al. [17]. In top view, the mean TS for the pre-experiment SSQ was 1.07 (SD = 1.8) and the post-experiment was 20.8 (SD = 19.9). In X-ray, the mean TS for the pre-experiment SSQ was 8.9 (SD = 10.8) and the post-experiment was 12.6 (SD = 10.9). In multiperspective, the mean TS for the pre-experiment SSQ was 3.3 (SD = 5.5) and the post-experiment was 15.9 (SD = 12.3). While all values are well below the threshold of 70 above which simulator sickness is considered to be present [8], the increase from pre-exposure to post-exposure does indicate that simulator sickness may be an important factor in this context.

Like for the task performance metrics, we analyzed the perceptual metric data with a Friedman test. For spatial awareness $\chi^2(3) = 5.1$ and $p = 0.17$, for depth perception $\chi^2(3) = 4.2$ and $p = 0.24$, for the pre-experiment SSQ $\chi^2(3) = 2.7$ and $p = 0.44$, and for the post-experiment SSQ $\chi^2(3) = 0.6$ and $p = 0.91$, which shows that there are no significant differences between the four conditions.

5.3 Limitations

One limitation of this work is that the virtual scenes were much larger than the tracked physical space, requiring occasional teleportation, which interferes with the VR scene exploration efficiency that our experiments attempt to measure. This is a fundamental limitation of VR, and future work will continue to devise techniques for hiding this limitation, or for more effective teleportation. Occlusion management techniques can help with reducing the need for teleportation, as a participant can save an unnecessary teleportation when they can see from the current location that the teleportation

destination is of no interest.

Another limitation of this work is that the participant had to learn several controller button functions, and the lack of familiarity with the interface might have affected a participant's ability to benefit from the occlusion management techniques. Furthermore, participants experienced with computer gaming and its controllers might find the interface more intuitive than others, creating a substantial difference between participants. Occlusion management techniques complicate interface design as they add degrees of freedom to the camera used to navigate the virtual scene. Future work should pursue intuitive interface constructs that allow the user to deploy and retract the disocclusion effect naturally.

Finally, our participant population was recruited narrowly, from the graduate student population of our graphics, visualization, and vision laboratory, many of whom had VR exposure and even experience, and future studies should expand this participant base. Furthermore, our studies tested occlusion management techniques in short exposures to the VR environment, in relatively passive tasks. Future studies should test these techniques in extensive exposures to very dynamic VR environments, such as VR games, where high cognitive load and cybersickness are more likely.

6 Conclusions and Future Work

We have presented a user study that compares three occlusion management techniques to conventional visualization, in the context of VR scene exploration. The comparison was based both on task performance and perceptual metrics. The tasks chosen benefit from visualization with a horizontal view direction, which is needed to tell apart the various items on the shelves of the supermarket and the persons in the garage. Of the four methods, three methods excel at providing horizontal visualization, i.e. multiperspective, X-ray, and conventional, and one method is ill-suited for such a visualization, i.e. top-view visualization. Of course, there are tasks where horizontal visualization should be sacrificed in favor of the comprehensive visualization provided by the top view, e.g. counting the number of people in the garage, or finding a lost child in the isles of the grocery store. Whereas our paper indicates benefits in the case of two tasks of practical importance, future work is needed to illuminate and quantify the exact benefits of each method on additional tasks.

Our paper explores the benefit of occlusion management techniques, previously developed for non-immersive visualization, in the context of HMD VR visualization. All disocclusion techniques benefit from high output image resolution, and that is a challenge in VR. Furthermore, all disocclusion techniques bring to the user a supernatural power, either of seeing the world from up high, through normally opaque walls, or around corners, which could break the sense of immersion. The disocclusion effect should be designed such that the user is convinced of the immersion in the virtual world, which they explore with the benefit of a superpower. Each disocclusion method also has specific VR challenges. Top view visualization has to decide when and how to deploy the top view. The discrepancy between the orientation top view and the world's orientation could be alleviated by displaying the top view horizontally, but requiring the user to look down while wearing an HMD of non-negligible weight is uncomfortable. X-ray visualization has the challenge of letting the user easily specify how many layers should become transparent. In addition to the interface complexity issues mentioned earlier, multiperspective visualization is limited by having to design the disocclusion effect manually, as a preprocess. Future work should explore algorithms for the automatic construction of the camera model underlying the disocclusion effect.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China through Projects 61772051, by the Beijing Natural Science Foundation L182016.

References

[1] C. Armbrüster, M. Wolter, T. Kuhlen, W. Spijkers, and B. Fimm. Depth perception in virtual reality: distance estimations in peri- and extrapersonal space. *Cyberpsychology behavior : the impact of the Internet, multimedia and virtual reality on behavior and society*, 11(1):9, 2008.

[2] B. Avery, C. Sandor, and B. H. Thomas. Improving spatial perception for augmented reality x-ray vision. In *Virtual Reality Conference, 2009. VR 2009. IEEE*, pp. 79–82. IEEE, 2009.

[3] F. Bork, C. Schnelzer, U. Eck, and N. Navab. Towards efficient visual guidance in limited field-of-view head-mounted displays. *IEEE transactions on visualization and computer graphics*, 24(11):2983–2992, 2018.

[4] E. Bozgeyikli, A. Raij, S. Katkoori, and R. Dubey. Point & teleport locomotion technique for virtual reality. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, pp. 205–216. ACM, 2016.

[5] S. Bruckner and M. E. Groller. Exploded views for volume data. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1077–1084, 2006.

[6] M. Burns and A. Finkelstein. Adaptive cutaways for comprehensible rendering of polygonal scenes. *ACM Transactions on Graphics (TOG)*, 27(5):154, 2008.

[7] J. Cohen. Statistical power analysis for the behavioral sciencies. *Technometrics*, 31(4):499–500, 1988.

[8] J. A. Ehrlich and E. M. Kolasinski. A comparison of sickness symptoms between dropout and finishing participants in virtual environment studies. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 42, pp. 1466–1470. SAGE Publications Sage CA: Los Angeles, CA, 1998.

[9] C. Everitt. Interactive order-independent transparency. *White paper, nVIDIA*, 2(6):7, 2001.

[10] M. Friedman. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Publications of the American Statistical Association*, 32(200):675–701, 1939.

[11] U. Gruenefeld, D. Hsiao, and W. Heuten. Eyeseex: Visualization of out-of-view objects on small field-of-view augmented and virtual reality devices. In *Proceedings of the 7th ACM International Symposium on Pervasive Displays*, p. 26. ACM, 2018.

[12] T. Günther, C. Rössl, and H. Theisel. Opacity optimization for 3d line fields. *ACM Transactions on Graphics (TOG)*, 32(4):120, 2013.

[13] R. I. Hartley and R. Gupta. Linear pushbroom cameras. In *European Conference on Computer Vision*, pp. 555–566. Springer, 1994.

[14] L. J. Hettinger, K. S. Berbaum, R. S. Kennedy, W. P. Dunlap, and M. D. Nolan. Vection and simulator sickness. *Mil Psychol*, 2(2):171–181, 1990.

[15] D. Kalkofen, M. Tatzgern, and D. Schmalstieg. Explosion diagrams in augmented reality. In *Virtual Reality Conference, 2009. VR 2009. IEEE*, pp. 71–78. IEEE, 2009.

[16] Y. Kameda, T. Takemasa, and Y. Ohta. Outdoor see-through vision utilizing surveillance cameras. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 151–160. IEEE Computer Society, 2004.

[17] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.

[18] E. Langbehn, F. Steinicke, M. Lappe, G. F. Welch, and G. Bruder. In the blink of an eye: leveraging blink-induced suppression for imperceptible position and orientation redirection in virtual reality. *ACM Transactions on Graphics (TOG)*, 37(4):66, 2018.

[19] J. J. LaViola Jr, D. A. Feliz, D. F. Keefe, and R. C. Zeleznik. Hands-free multi-scale navigation in virtual environments. In *Proceedings of the 2001 symposium on Interactive 3D graphics*, pp. 9–15. ACM, 2001.

[20] W. Li, M. Agrawala, B. Curless, and D. Salesin. Automated generation of interactive 3d exploded view diagrams. In *ACM Transactions on Graphics (TOG)*, vol. 27, p. 101. ACM, 2008.

[21] W. Li, L. Ritter, M. Agrawala, B. Curless, and D. Salesin. Interactive cutaway illustrations of complex 3d models. In *ACM Transactions on*

Graphics (TOG), vol. 26, p. 31. ACM, 2007.

[22] A. Macquarrie and A. Steed. The effect of transition type in multi-view 360° media. *IEEE Transactions on Visualization Computer Graphics*, 24(4):1564–1573, 2018.

[23] C. Mei, V. Popescu, and E. Sacks. The occlusion camera. In *Computer Graphics Forum*, vol. 24, pp. 335–342. Wiley Online Library, 2005.

[24] R. Nakatani, D. Kouno, K. Shimada, and T. Endo. A person identification method using a top-view head image from an overhead camera. *JACIII*, 16(6):696–703, 2012.

[25] V. Popescu, P. Rosen, and N. Adamo-Villani. The graph camera. In *ACM Transactions on Graphics (TOG)*, vol. 28, p. 158. ACM, 2009.

[26] P. Rademacher and G. Bishop. Multiple-center-of-projection images. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pp. 199–206. ACM, 1998.

[27] S. Razzaque, Z. Kohn, and M. C. Whitton. Redirected walking. In *Proceedings of EUROGRAPHICS*, vol. 9, pp. 105–106. Citeseer, 2001.

[28] S. Razzaque, D. Swapp, M. Slater, M. C. Whitton, and A. Steed. Redirected walking in place. In *EGVE*, vol. 2, pp. 123–130, 2002.

[29] Robert Zak. 10 annoying things that are ruining video games. `https://whatculture.com/gaming/10-annoying-things-that-are-ruining-video-games`, 2016.

[30] R. A. Ruddle and S. Lessels. The benefits of using a walking interface to navigate virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 16(1):5, 2009.

[31] C. Sandor, A. Cunningham, A. Dey, and V.-V. Mattila. An augmented reality x-ray system based on visual saliency. In *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on*, pp. 27–36. IEEE, 2010.

[32] S. S. Sawilowsky. New effect size rules of thumb. *Journal of Modern Applied Statistical Methods Jmasm*, 8(2):597–599, 2009.

[33] J. L. Souman, P. R. Giordano, M. Schwaiger, I. Frissen, T. Thümmel, H. Ulbrich, A. D. Luca, H. H. Bülthoff, and M. O. Ernst. Cyberwalk: Enabling unconstrained omnidirectional walking through virtual environments. *ACM Transactions on Applied Perception (TAP)*, 8(4):25, 2011.

[34] F. Steinicke, G. Bruder, J. Jerald, H. Frenz, and M. Lappe. Analyses of human sensitivity to redirected walking. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pp. 149–156. ACM, 2008.

[35] F. Steinicke, G. Bruder, J. Jerald, H. Frenz, and M. Lappe. Estimation of detection thresholds for redirected walking techniques. *IEEE transactions on visualization and computer graphics*, 16(1):17–27, 2010.

[36] Q. Sun, A. Patney, L.-Y. Wei, O. Shapira, J. Lu, P. Asente, S. Zhu, M. McGuire, D. Luebke, and A. Kaufman. Towards virtual reality infinite walking: dynamic saccadic redirection. *ACM Transactions on Graphics (TOG)*, 37(4):67, 2018.

[37] Q. Sun, L.-Y. Wei, and A. Kaufman. Mapping virtual and physical reality. *ACM Transactions on Graphics (TOG)*, 35(4):64, 2016.

[38] T. Teshima, H. Saito, S. Ozawa, K. Yamamoto, and T. Ihara. Vehicle lateral position estimation method based on matching of top-view images. In *null*, pp. 626–629. IEEE, 2006.

[39] S. Tregillus and E. Folmer. Vr-step: Walking-in-place using inertial sensing for hands free navigation in mobile vr environments. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 1250–1255. ACM, 2016.

[40] J. W. Tukey. Comparing individual means in the analysis of variance. *Biometrics*, pp. 99–114, 1949.

[41] M. Usoh, K. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater, and F. P. Brooks Jr. Walking> walking-in-place> flying, in virtual environments. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pp. 359–364. ACM Press/Addison-Wesley Publishing Co., 1999.

[42] Wikipedia contributors. Super mario 64 — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Super_Mario_64&oldid=881286914`, 2019. [Online; accessed 8-February-2019].

[43] B. Williams, G. Narasimham, B. Rump, T. P. McNamara, T. H. Carr, J. Rieser, and B. Bodenheimer. Exploring large virtual environments with an hmd when physical space is limited. In *Proceedings of the 4th symposium on Applied perception in graphics and visualization*, pp. 41–48. ACM, 2007.

[44] M. L. Wu and V. Popescu. Multiperspective focus+context visualization. *IEEE Transactions on Visualization Computer Graphics*, 22(5):1555–1567, 2016.

[45] M. L. Wu and V. Popescu. Efficient vr and ar navigation through multiperspective occlusion management. *IEEE Transactions on Visualization Computer Graphics*, PP(99):1–1, 2017.

[46] M.-L. Wu and V. Popescu. Anchored multiperspective visualization for efficient vr navigation. In *International Conference on Virtual Reality and Augmented Reality*, pp. 240–259. Springer, 2018.

[47] X. Xie, Q. Lin, H. Wu, G. Narasimham, T. P. McNamara, J. Rieser, and B. Bodenheimer. A system for exploring large virtual environments that combines scaled translational gain and interventions. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, pp. 65–72. ACM, 2010.

[48] J. Yu and L. McMillan. General linear cameras. In *European Conference on Computer Vision*, pp. 14–27. Springer, 2004.

[49] D. W. Zimmerman. A note on interpretation of the paired-samples t test. *Journal of Educational Behavioral Statistics*, 22(3):349–360, 1997.