OVRlap: Perceiving Multiple Locations Simultaneously to Improve Interaction in VR

Jonas Schjerlund
University of Copenhagen
Copenhagen, Denmark
josh@di.ku.dk

Kasper Hornbæk
University of Copenhagen
Copenhagen, Denmark
kash@di.ku.dk

Joanna Bergström
University of Copenhagen
Copenhagen, Denmark
joanna@di.ku.dk

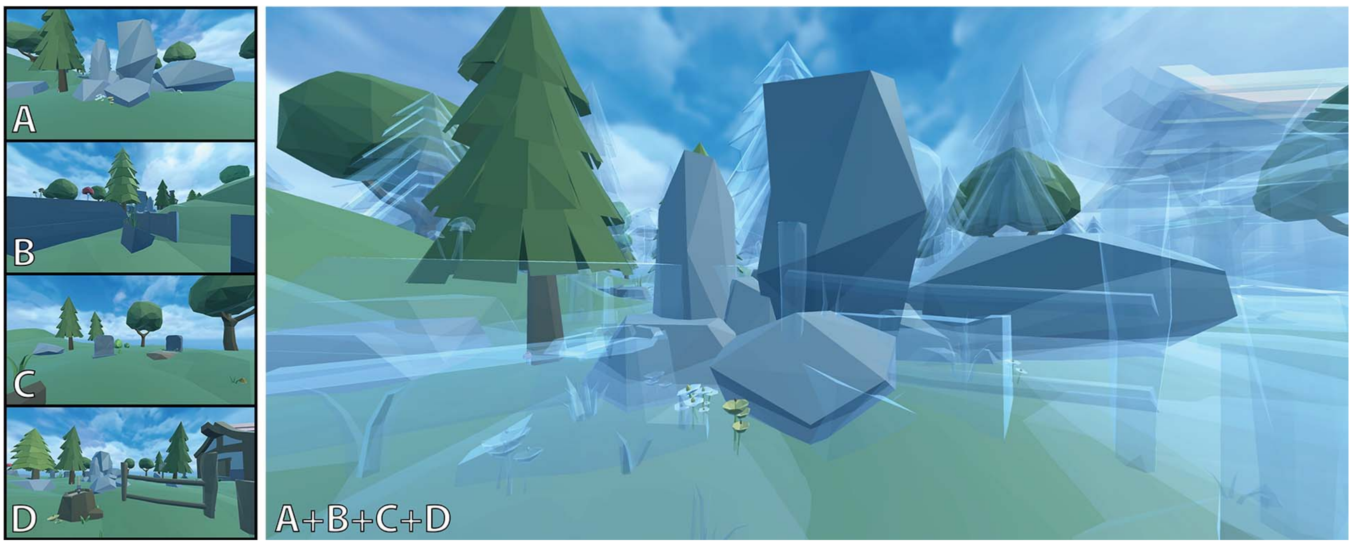

Figure 1: The *OVRlap* technique allows the user to perceive and spatially orient themselves simultaneously across multiple distinct and distant locations from a first-person perspective. It accomplishes this by overlaying multiple viewpoints on top of one another, with one active and opaque viewpoint (here, A) and multiple transparent and passive viewpoints (here, B, C, and D). The user can only interact in the active viewpoint.

ABSTRACT

We introduce *OVRlap*, a VR interaction technique that lets the user perceive multiple places at the same time from a first-person perspective. *OVRlap* achieves this by overlapping viewpoints. At any time, only one viewpoint is active, meaning that the user may interact with objects therein. Objects seen from the active viewpoint are opaque, whereas objects seen from passive viewpoints are transparent. This allows users to perceive multiple locations at once and easily switch to the one in which they want to interact. We compare *OVRlap* and a single-viewpoint technique in a study where 20 participants complete object-collection and monitoring tasks. We find that in both tasks, participants are significantly faster and move their head significantly less with *OVRlap*. We propose how the technique might be improved through automated switching of the active viewpoint and intelligent viewpoint rendering.

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04…$15.00
https://doi.org/10.1145/3491102.3501873

CCS CONCEPTS

• **Human-centered computing** → *User studies*; **Interaction techniques**; **Virtual reality**.

KEYWORDS

virtual reality, interaction techniques, large environments, user studies

ACM Reference Format:
Jonas Schjerlund, Kasper Hornbæk, and Joanna Bergström. 2022. OVRlap: Perceiving Multiple Locations Simultaneously to Improve Interaction in VR. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3491102.3501873

1 INTRODUCTION

Being in multiple places at the same time is a common theme in fiction. Examples range from classic literature (e.g., doppelgängers as in Dostoevsky's *The Double* from 1846) to modern cinema (e.g., Dr. Strange's multiplied selves in 2018's *Avengers: Infinity War*). Conceptually, this allows us to see and engage with the world from multiple locations simultaneously, something normally impossible. We propose that this impossibility can be realized in virtual reality (VR). This could aid in interactions that are difficult or complex in VR, such as navigating and acting in distinct and distant locations in larger virtual environments.

We introduce *OVRlap*, a VR interaction technique that allows the user to perceive multiple places at the same time by showing them the overlapping viewpoints of multiple avatars from a first-person perspective (Figure 1). Only one overlapping viewpoint is active at a time, indicated by its objects being opaque. The user can directly interact with the environment only from this viewpoint. The passive viewpoints are visible as semi-transparent shapes, spatially overlaid on top of the active viewpoint, allowing the user to always perceive objects at all viewpoint locations from a first-person perspective. The active viewpoint can be switched.

The *OVRlap* technique offers two novel benefits over existing techniques for interacting in multiple places in a large virtual environment. First, it allows the user to maintain a first-person perspective while perceiving and spatially orienting themselves in multiple distinct and distant locations. The preservation of the first-person perspective and the spatial relationship between the user and the environment stand in contrast to traditional techniques for perceiving multiple locations in a large environment, such as world-in-miniature [42] which distort the user's perspective, or portals [23, 33] which distort space. Second, *OVRlap* allows a user to spatially orient themselves and begin to act and preemptively reach for objects in a location before arriving there. For example, the user can spot the relative location of an object on a table at a distant location and move their hand into position to interact with it, while at the same time switching through the available viewpoints to reach the location where the table is located. In traditional VR, using a technique such as teleportation [1, 6], the user must first navigate to the distant location before being able to accurately determine the relative location of the object and start moving their hand to interact with it. Together, we expect that these benefits will offer a significantly improved sense of spatial perception of multiple locations. This should be reflected in an improved ability to perform common VR interaction tasks that take place in large environments, such as locating or manipulating objects in distinct and distant locations.

We evaluate an implementation of *OVRlap* in an empirical study that examines the efficiency and usability of the technique, as well as its effects on presence and body ownership. Participants complete an object collection task and a monitoring task that involve interactions in four distinct locations in a VR environment. The tasks are completed using both *OVRlap* and a cycling technique where participants switch between the locations one at a time, as seen in traditional VR. We find that *OVRlap* significantly outperforms the cycling technique in completion time for either task. It similarly offers a significant reduction in physical motion of the user's head. This suggests that the technique provides a significantly improved ability to perceive the four locations at once. However, participants switch the active viewpoint more times than is required to complete their task. The technique also offers comparable levels of presence, body ownership, usability to the control condition. Based on these findings, we discuss a number of future directions for the *OVRlap* concept and outline some of the concept's limitations.

Our work introduces the concept of perceiving multiple distinct and distant locations at the same time from a first-person perspective through the use of overlapping viewpoints. We demonstrate that it is possible to meaningfully implement this concept in VR through the *OVRlap* technique, providing a novel form of perception across multiple locations that grants significant benefits in completion time and reduction of physical motion for object collection and monitoring tasks. We also demonstrate that it enables novel forms of interaction in areas that VR techniques have traditionally struggled with, such as navigation and interaction across long distances in larger environments. Altogether, we show that the impossible concept of perceiving multiple locations at once is not only possible in VR, but that it can offer benefits that are useful and enjoyable for the user.

2 RELATED WORK

We expect *OVRlap* to be useful for interacting in large VR environments by enabling an improved ability to see and spatially orient oneself in multiple distinct locations at once. We first discuss related techniques that allow navigation to distant location and then ones that map the control of the body to reach far distances for the purpose of interaction. We then discuss related techniques that join multiple locations or viewpoints together.

2.1 Moving to Distance for Interaction

There are a number of interaction techniques that allow the user to relocate themselves in a large virtual environment without needing to move physically. Such techniques often function by the user indicating a target location and then being transported there instantly [1]. A desktop example of this is *Navidget*, where a camera in a 3D space can be positioned by, for example, circling objects that the camera should center on [17]. In VR, two such examples are teleportation [6], where the user can indicate a target location in the environment that they are instantly transported to, or map dragging [8], where the user instead indicates their target location on a 2D map of the environment. The map metaphor is similar to the world-in-miniature concept (WIM) [42], where users can indicate a target location on a miniature 3D representation of the virtual environment with their hands [13], or even by stepping on it [26]. The teleportation concept has also been extended to context-appropriate metaphors [45], for example by moving a user through a library through rotating bookshelves or through an outdoor environment by having a bird carry them. However, because these teleportation-based approaches apply an instant or constant transition, which can be disorienting. Therefore, the users often experience difficulties in re-orienting themselves after the transition [6].

Instead of an instant or constant speed, previous work has suggested how the users can glide through an environment towards a target location at variable speeds [6] or follow a self-defined route [7]. Such gliding through the environment can help preserve the user's spatial orientation [5, 7]. However, this is slower than instant transportation, and moving virtually but not physically often induces VR sickness [41].

OVRlap aims to overcome these pitfalls by overlaying views from distinct locations, and allowing the users to instantly change the location in which they are interacting. This way, the users can maintain a first-person viewpoint to the environment and thereby continuously orient themselves with the environment across each location.

2.2 Interaction at a Distance

Another approach to interaction at distant locations are techniques that allow reaching there. One such body of work changes how the physical body is mapped to the virtual one. A classic example of this is the *Go-go* technique [34]. Go-go allows the user to move their hand to distant locations through a non-linear mapping function. The function maps the virtual hand location 1-to-1 with the physical hand near the body but when reaching the physical hand out, the virtual hand reaches exponentially further. One physical body or body part can also be mapped to multiple virtual ones. For example, it has been shown that users can experience body ownership over extra fingers [20], hands [10, 15], arms [16], or even whole bodies [19]. Another, more functional perspective has explored how a one-to-many mapping might improve interaction in VR. An example of this is *Ninja Hands*, in which a single physical hand is mapped to numerous distributed virtual hands for improved target selection at far distances [38]. It has also been demonstrated that in a balloon-popping task, users could be trained to use a third arm controlled by wrist rotations to be more efficient than with two arms [44]. In these approaches, the user's perspective is not manipulated, but the control often is, such as by breaking the natural 1-to-1 mapping between the user's body and their avatar to allow them to move beyond the reach of their physical space [32, 38]. As the perspective is not manipulated, a drawback here is that the use of the techniques is limited to distances where the users can still well see their objects of interest. A the control is manipulated, the use is limited to distances at which the mapping still allows interaction with precision.

Another class of techniques for interacting at a distance is to warp the space around the user to bring the desired locations within reach. Two key examples of this approach are the world-in-miniature [42], where the user can indicate objects and locations on a miniature model of the virtual environment; and the portal metaphor [23], where two distinct areas are spatially connected through a portal interface. A similar idea is to merge the concept of moving the user and allowing them to point by temporarily moving them to their pointing destination and back again [31]. Integrating portals with teleportation is also useful for redirecting users back towards the centre of their calibrated physical space [30]. This metaphor has also been extended to three dimensions, where spherical or cuboid segments of the environment can be linked, carried around, and manipulated [24, 33]. All of these techniques, however, require users to learn complex perspective- and space-warping metaphors.

OVRlap technique aims to overcome these limitations by maintaining a 1-to-1 mapping between user's physical and virtual body, while retaining a near-distance view to the interaction space.

2.3 Multiplexing Distinct Views

The concept of multiplexing distinct locations or viewpoints has a small body of previous work. In 2D desktop interfaces, translucent layers have been used to layer windows [18], or zoomed and un-zoomed views of content such as maps [29], such that these different views share the same screen space. Saraiji et al. [37] layered multiple distinct telepresence viewpoints on top of each with the user's gaze used to disambiguate them and decide which one should be in focus. However, this work concerns only layering views in a window format. [21] multiplexes views in a first-person perspective, but so that a user sees other user's viewpoints side-by-side to their own. This approach has also been used in data visualization, where geospatial data are visualized in multiplexed, side-by-side, or stacked windows in front of the user [40]. The approach of Armstrong et al. [2] and Fan et al. [14] relates more closely to OVRlap. In that approach a 360° video of the user's physical environment was halved and the two halves were spatially overlaid and blended. However, these works do not concern interacting in the multiplexed locations.

OVRlap builds on such ideas on multiplexing viewpoints, but importantly, also allows interacting in those. It does so by allowing the user to passively see the many viewpoints from a classic first-person perspective, as well as changing between those to interact.

3 THE OVRLAP TECHNIQUE

The *OVRlap* technique allows a user to perceive multiple places at the same time from a first-person perspective. It does this by spatially overlapping viewpoints from these distinct locations on top of each other. It further features the concept of an active viewpoint, in which the user can interact, and one or more passive viewpoints. The key components of the technique are consequently (A) the functionality and representation of the overlapping viewpoints and (B) the mechanism for placing, removing, and switching viewpoints.

3.1 Function and Representation of Active and Passive Viewpoints

The key idea in *OVRlap* is to render viewpoints layered and fully overlapping on top of each other (Figure 2). The main benefit of this idea is that it allows a user to perceive visual elements and spatial relationships in multiple distinct and distant locations from a first-person perspective (Figure 4). To our knowledge, *OVRlap* is the first work to leverage the first-person viewpoint for multiple distinct locations with a focus on interaction. This allows a user to look around and spatially orient themselves in these multiple locations in the same manner as they would in traditional VR despite acting across distinct and distant locations.

In *OVRlap*, there is an *active viewpoint* and a number of *passive viewpoints*. Active and passive viewpoints are distinguished by their function and their representation.

Functionally, the main difference between the active and the passive viewpoints is that the user can only interact with objects in the active viewpoint. Though interactive objects in passive viewpoints are still visible, they cannot be interacted with. For example, the user can spot a ghostly apple on a table from an passive viewpoint, but while they can move their hand to the apple, it cannot be picked up and physics-related events will not be invoked by

Figure 2: The *OVRlap* concept. The user wants to perceive both a forest (1) and a cemetery (2). In traditional VR, the user would have to navigate between these locations in turn to perceive them. With the *OVRlap* technique, the two locations are instead overlaid on top of each other and the user can perceive both at the same time from a first-person viewpoint. The user can interact in the active and opaque viewpoint (here, the forest) while spatially orienting themselves in one or several passive and transparent viewpoints (here, the cemetery).

Figure 3: Motion in *OVRlap*. When the user moves in *OVRlap*, this is equivalent to moving in the forest (1) and the cemetery (2) simultaneously and synchronously. The user only collides with objects from the active viewpoint. This synchronous motion allows the user to reach for an object from a passive viewpoint before and while switching to make it active.

the user's touch. This allows the technique to function in complex scenarios, where there are multiple viewpoints that have many relevant and interactable objects. If objects in all viewpoints could be interacted with, this would cause issues with disambiguation when multiple objects in different viewpoints occupy the same space relative to the user. However, while the user can only interact in the active viewpoint, a key concept of *OVRlap* is that the user moves simultaneously and synchronously in all viewpoints at all times. This allows novel interactions such as preemptively reaching for an object in a passive viewpoint before and while switching to it (Figure 3).

The active and passive viewpoints are represented differently. Objects in the active viewpoint are opaque, while objects in passive viewpoints are transparent. This distinction serves two purposes. First, by making the active viewpoint opaque and solid, the technique takes advantage of users' existing VR experience. Usually in VR, the environment we act in is opaque, which makes it intuitive to interact with opaque objects. Second, by having passive viewpoints be transparent and ghostly, it is easier to layer multiple passive viewpoints on top of one another while retaining the ability to tell distinct objects apart.

A more advanced representation of passive viewpoints can introduce some notion of ranking, with higher-ranked viewpoints being less translucent than lower ranked ones (as in [29]). Further, it would be possible to have a partial overlap, with specific objects from different viewpoints being brought into focus. Lastly, the opacity of objects could be determined by their function in the VR experience. For example, if a wall in front of the user is not important to the user's interaction, it could have a lower opacity to allow objects that are spatially located behind it to be visible.

3.2 Placing, Removing, and Switching Viewpoints

The *OVRlap* technique features viewpoints from multiple overlapping locations, of which only one is active. Therefore, there has to be a mechanism for placing and removing viewpoints, as well as for switching the active viewpoint.

We propose two hypothetical approaches to placing a viewpoint: As part of the design of the environment, or manually by the user. It is already common for VR designers to designate specific locations as target destinations for navigation techniques. For example, many VR experiences feature designated target spots where teleportation is allowed. By extending this to *OVRlap*, the designer can designate relevant locations within a scene where a viewpoint should be generated for the user.

The other approach is to have the user manually generate and remove these viewpoints. For example, the user navigates an environment through traditional means of navigation, such as teleportation [1, 6]. When they have reached a location they would like to be able to see from or quickly access in the future, they can then generate a viewpoint at this location through a button press. Similarly, users can have the option to remove a viewpoint that they no longer want to use, or to archive it in such a way that it is not be displayed, but can be brought back later (similar to the way portals can be stored in *Poros* [33]). Viewpoint generation can also be automated. This way, the application automatically generates viewpoints at locations where users interact a lot, allowing them to easily perceive and return to the locations that are most relevant for interaction. The process of viewpoint removal may similarly be automated, such that viewpoints that have gone unused for a longer period of time are automatically removed or archived.

Figure 4: Example of a first-person perspective of the *OVRlap* technique. The sphere is the user's hand, which is reaching for a coin. Since the coin is transparent, it is actually located at a passive viewpoint far from the opaque chest at the active viewpoint.

We also propose two hypothetical approaches to switching between viewpoints: Manual and automatic. In a manual switching approach, users can switch by manual input, for example by pressing the thumbstick on a controller left or right. The switching order of the viewpoints may be based by natural features of the environment. For example, in a long hallway, switching can move you to the closest viewpoint in front or behind you. In a large open environment, switching can attempt to cycle through the avatars in a clockwise pattern. A manual approach has the advantage of familiarity, as users are used to navigating longer distances in virtual environments using controller input, such as through teleportation [1, 6].

In an automatic switching approach, an algorithm determines the most likely interaction target for the user. For example, this algorithm can be based on a predictive model of interaction intent. Relevant features for such a model can be derived from the user's hands or gaze. We expand this concept in the discussion based on the results from our study (see Section 6.2).

The biggest question for the technique is whether it is possible for users to perceive multiple overlapping viewpoints simultaneously. We next evaluate that in our user study.

4 STUDY

The purpose of the study is to evaluate how the *OVRlap* technique and its use of overlapping viewpoints affect performance and subjective experience in tasks where a user has to act in multiple distinct and distant locations. We do so by comparing it to a more traditional technique for being in VR, which is seeing a single view at a time. Participants complete an object collection task and a monitoring task using these techniques. This study design follows the guidelines for evaluation and manipulation of objects in VR outlined by Bergström et al. [4].

4.1 Design

The study uses a within-subjects design. The independent variable is interaction technique, with two levels: the *OVRlap* techniques with its overlapping viewpoints, and a *cycling* technique where the participant cycles through the same viewpoints, one at a time. The study consists of two tasks: An object collection task and a monitoring task. The order of both the interaction techniques and the tasks are counter-balanced across participants, so that each interaction technique is used in turn, and the order of the tasks is not changed during the experiment.

In the first task, the participant collects a total of 32 coins (4 coins $\times$ 8 trials). In the second task, the participant selects a total of 50 discs. This is repeated for each technique, for a total of 64 observations in the object collection task, and 100 observations in the monitoring task. These numbers were chosen to reach an average of 12 minutes of usage per technique, based on a pilot study. This duration is intended to be long enough to counteract training effects and allow participants to familiarize themselves with a technique before being prompted to evaluate it in subjective questionnaires. The total duration of the experiment, including setup and questionnaire responses, was approximately one hour.

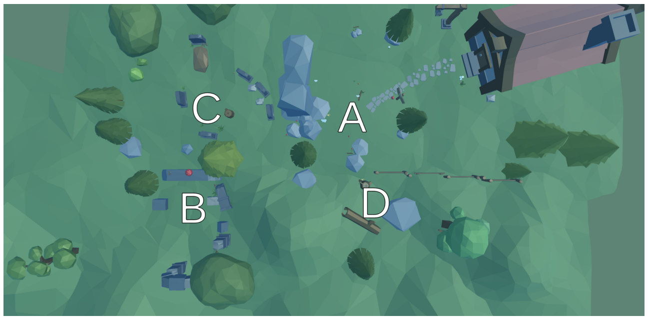

Figure 5: Bird's-eye view of the study environment. It is an outdoor environment with four distinct landmark locations: Outside a house (A), some ruins (B), a cemetery (C), and by a wood chopping station (D). There is a viewpoint at each location.

4.2 Virtual Environment

The participant uses the *OVRlap* and cycling techniques to switch between four viewpoints. The viewpoints are placed at distinct and distant locations in the scene. The scene is an outdoor environment approximately 100 m across containing objects such as a house, trees, rocks, and stumps (Figure 5), all with low-resolution textures and low-polygon meshes.

We choose to evaluate an *OVRlap* implementation of four overlapping viewpoints since this is a similar number of layered or multiplied viewpoints (e.g., three [37] or four [21] for telepresence data and four for geospatial data [40]). The locations of the viewpoints were chosen so that they are near distinct landmarks where interaction could naturally occur, such as where objects could lay. For each task, the forward direction of each avatar is adjusted such that the participant would need to rotate their body to face all the objects. The participant manually switches between viewpoint locations by pressing the joystick left or right.

4.3 Tasks

There are two tasks in the study: object collection and monitoring.

4.3.1 Object collection. In object collection, participants need to pick four coins and place those in a chest. This task is chosen to represent object manipulation wherein the user needs to locate, pick, and move objects from one location to another, distant, location (Figure 6).

A set of four coins are placed on plausible surfaces, such as on tree stumps, rocks, or on the ground. They are all placed within arm's reach from the four viewpoints. The placement of the coins is systematically randomized so that each of the four viewpoints has an equal total number of coins in the reach, but that in each trial, the number of coins at any given viewpoint can vary from none to all four. This randomization is constant across techniques and participants.

At one viewpoint is a chest. The participant must pick up each coin in turn and place it into the chest. The user picks up the coins by touching them with spherical colliders around their avatar's hands (10 cm diameter) and pressing and holding the trigger on the

Figure 6: The object collection task. Left: The user first locates a coin in a random location. Right: The user then brings the coin to the chest in a location and deposits it.

controller. One trial consists of placing four coins into the chest. Only one coin can be picked up and moved at a time. If a coin is not within the chest or picked up again within 5 seconds of being released, it gets removed and logged as an error. This is to avoid edge cases where the participant accidentally throws a coin out of reach. Once all coins are placed, the task is reset and a new trial begins, for a total of eight trials. Between trials, the chest is moved to a fixed random spot at a 90° angle on a circle around one of the avatar, and the placement of each coin is changed, to reset the participant's spatial memory.

4.3.2 Monitoring. In monitoring, the participants need to hit a moving disc. This task is chosen to represent object selection wherein the use needs to locate a moving object and point at it to select it (Figure 7).

In the task, a flat disc (50cm radius) will appear at one of the viewpoints at a time. The disc is vertically oriented, facing the viewpoint. The disc flies along a figure-8 trajectory, with the diameter of each of the two circles in the figure-8 trajectory being 2m. The trajectory also faces the viewpoint. The trajectories are centered two meters away and up from the viewpoint on one of four 90° increments around the viewpoint. These trajectories are designed to ensure a wide variety of potential positions for the discs. The disc moves forward on its trajectory at a speed of 1 m/s.

The participant has a thin ray pointing out in a straight line from their right hand. The participant is tasked with changing their viewpoint to one at the location closest to the disc, pointing the ray at the disc as quickly as possible, and pulling the trigger on their controller. On a success, another disc appears at a random viewpoint and orientation, starting at a random point on the figure-8 trajectory. This randomization is constant across techniques and participants.

4.4 Dependent Measures

We use measures of performance and subjective user experience to evaluate the interaction techniques. As performance measures, we include trial duration, accuracy, and error rate.

In the object collection task, a trial is defined as the time it takes to collect and place all four coins in the chest. In the monitoring task, a trial is defined as the time between selecting one disc and selecting the next.

Figure 7: The monitoring task. Left: The user first locates the target in a random location. Right: The user then points the ray extending from their right hand at the target and presses the trigger on their controller to select it.

We measure accuracy for each task, defined as the number of perspective switches used during a trial. We also measure the error rate for the two tasks, with an error in the monitoring task defined as clicking the button while not pointing at the target disc, and an error in the object collection task defined as releasing a coin without placing it in the chest within 5 seconds and without picking it up again.

We also measure how much the participant moves their head within each trial as an expression of physical motion, which the *OVRlap* technique should ideally reduce due to its improved overview.

To investigate the subjective experience and potential costs of using the technique, we include three measures. The technique introduces visual and conceptual complexity. It overlays multiple viewpoints and the user might see other instances of themselves at a distance based on the relative location of these viewpoints. This might affect both the user's sense of being in a place from the overlapping locations, their sense of body ownership from seeing mirrored selves. The complexity might also affect the user's overall experience of usability. Consequently, we assess presence, body ownership, and usability after each technique. We assess presence using the IPQ presence questionnaire [35], based on recommendations from recent work on measuring presence within VR environments [39]. We measure subjective usability using the System Usability Scale (SUS) [3, 9]. Lastly, to measure body ownership we use the VEQ body ownership questionnaire extended from Roth & Latoshik [36].

4.5 Apparatus

The display used for the study is the Oculus Quest or Oculus Quest 2. The Oculus Quest has a display resolution of 1440×1600 per eye and a 72 hz refresh rate. The Oculus Quest 2 has a display resolution of 1832×1920 per eye and ran at a 72 hz refresh rate. The displays have a horizontal field of view of around 90°, but can vary by a few degrees based on subjective factors such as each user's head shape, display adjustment, and interpupillary distance settings.

The software for the study was implemented using the Unity engine (version 2019.4). The blending effect for the *OVRlap* technique was achieved by generating multiple instances of the environment and providing offsets for each instance's position and rotation to match the user's intended location and forward direction in that instance. To implement viewpoint switching, a non-transparent shader was applied to renderer materials in the active environment instance and a transparent shader was applied to renderer materials in the passive instances. Physics collision events were controlled such that objects in each environment instance could only collide with other objects in that instance.

4.6 Participants

We recruited 20 participants (1 female), age range 19-52, mean = 27.3, SD = 8.32. Due to the COVID-19 pandemic, the study was run remote and recruitment took place on Reddit, within the /r/OculusQuest subreddit, where users were likely to have a compatible VR display. Participants were required to have access to an Oculus Quest or Oculus Quest 2 HMD. 7 participants used the Quest, and 13 participants used the Quest 2. Consequently, participants were likely to have prior experience with VR. Participants were compensated for their time with the equivalent of €20.

4.7 Procedure

In the invitation to the study, each participant was instructed to clear an hour of time, ensure that their display was fully charged, and create an isolated indoor physical environment with no external interruptions and at least 1 meter of cleared space in all directions when standing. They were additionally instructed to create a stationary physical boundary calibration for the display and instructed to remain within it during the study, meaning no locomotion was required beyond turning and leaning. Further, the participant was provided with instructions on how to install the study software on their HMD. At the start of the study, the participant was presented with an informed-consent form and was unable to proceed until agreeing to it.

The participant was then presented with a training task for their first technique. The training task took place in the study's virtual environment. It had four buttons located on tables at such a distance that the participant had to use all of four viewpoints in the environment to accomplish the task by pressing each button. After completing the training task, an overview of the environment was presented, with the locations of the participant's viewpoints highlighted by red markers; this was explained in a text box. Then, the participant was presented with instructions for their first task, and asked to complete this task. After the first task, the participant completed their second task.

Upon completion of the second task, the participant was immediately presented with the questionnaires, which were completed one at a time. The user was moved to a randomly chosen viewpoints (constant between participants), but the rest of the environment was frozen. This process was then repeated for the other technique. Upon completion of the study, the participant was instructed in how to extract the data files from their HMD and where to upload them. After receiving the participant's data, the experimenter provided them with a gift card to the Steam store worth €20.

5 RESULTS

This section describes the outcome of the study. We discard outliers, defined as trials that take longer than 3 minutes for either task. We take these to represent the user taking a break or being distracted,

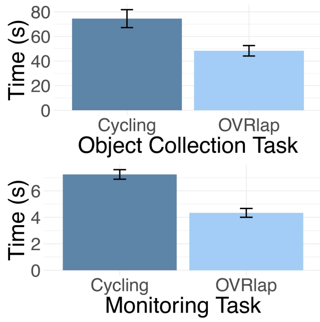

Figure 8: Mean completion time for either task and technique. Error bars represent 95% confidence intervals.

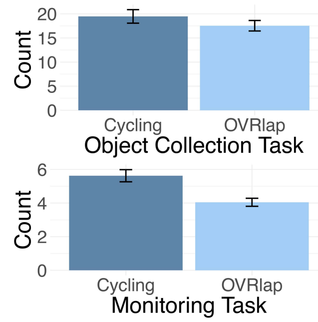

Figure 9: Mean switch count for either task and technique. Error bars represent 95% confidence intervals.

which violates the instructions they were presented with. There was a total of 3 such outliers in the monitoring task, and 2 in the object collection task.

5.1 Completion time

Completion time is the time it took for a participant to complete a trial successfully. In the monitoring task, a trial is the time it takes for a participant to locate the target, point at it, and press the trigger. In the object collection task, a trial is the time it takes to locate a set of four coins and place them in the chest, one at a time. Figure 8 shows the mean completion time for either task. Mean completion times for the monitoring task were 7.24 seconds for the cycling technique, and 4.33 seconds for the *OVRlap* technique. Mean completion times for the object collection task were 74.46 seconds for the cycling technique, and 48.23 seconds for the *OVRlap* technique. We used an aligned rank transform for non-parametric factorial analyses [43] (ART analysis) to evaluate the effects of technique on completion time. The ART analysis showed a significant effect of technique on completion time in both the monitoring task ($F_{1,19} = 28.72, p < .01$) and object collection task ($F_{1,19} = 24.69, p < .01$).

Participants were significantly faster in either task with the *OVRlap* technique, showing the benefit of its improved overview.

5.2 Switching

A switch is the event of changing the active viewpoint. In the study, this is manually invoked by the user pressing the thumbstick on their controller either left or right. The four viewpoints are arranged in a randomized circular list, so that the first element links to the last. Figure 9 shows the mean switch count for either task. The mean switch counts were 5.62 for the cycling technique and 4.02 for the *OVRlap* technique in the monitoring task, and 19.47 for the cycling technique and 17.54 for the *OVRlap* technique in the object collection task. An ART analysis showed a significant effect of technique on switch count in the monitoring task ($F_{1,19} = 43.51, p < .01$), but not in the object collection task ($F_{1,19} = 1.86, p = .24$).

These results indicate that participants switched more than they needed to with *OVRlap*. This is most clear in the monitoring task, the highest number of switches for any given trial should be two: The longest path from one viewpoint to any other viewpoint in a circular list of four elements is two. Even though *OVRlap* has a lower switch count than the cycling technique, it is still higher than it needs to be, indicating that participants either overshoot in their switches or cycle the long way around. Together, these data suggest that the technique can be improved by reducing the number of switches.

5.3 Errors

In the object collection task, there was a total of 63 errors, for an error rate of 5.04%. Of these, 39 happened while using the *OVRlap* technique. In the monitoring task, there was a total of 97 errors, for an error rate of 4.85%. Of these, 49 happened while using the *OVRlap* technique. These data indicate that participants understood the tasks and that the *OVRlap* technique did not significantly increase the error rate.

5.4 Physical Motion

Physical motion is the distance moved by the user's head in each trial. This is sampled at a rate of 10 hz based on the in-engine transform position of the HMD. For reference, rotating around oneself with the HMD horizontally leveled averaged to 0.93 m of motion in the study setup. Figure 10 shows the mean head motion exerted for either task and technique. Mean head motion exerted in the monitoring task was 10.15 m for the cycling technique, and 0.86 m for the *OVRlap* technique. Mean head motion exerted in the object collection task was 107.94 m for the cycling technique,

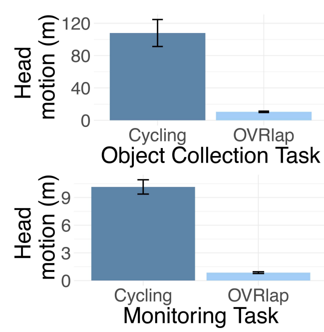

Figure 10: Mean head motion exerted for either task and technique. Error bars represent 95% confidence intervals.

Table 1: Results for presence (IPQ), usability (SUS), and body ownership (VEQ) questionnaires for either condition. SUS scores have a range of 0-100, whereas the other scores are on a 7-point Likert scale.

		Cycling		*OVRlap*	
Questionnaire	**Subscale**	**Mean**	**SD**	**Mean**	**SD**
Presence (IPQ)	General Presence	4.85	1.95	5.2	1.57
	Experienced Realism	3.76	1.94	4.01	1.76
	Involvement	4.56	1.8	4.75	1.66
	Spatial Presence	3.38	1.68	3.11	1.29
Usability (SUS)		66.62	15.75	74	12.81
Body Ownership (VEQ)		4.36	2.06	4.5	2.15

and 10.46 m for the *OVRlap* technique. An ART analysis showed a significant effect of technique on head motion in the monitoring task ($F_{1,19} = 28.77, p < .001$), as well as in the object collection task ($F_{1,19} = 43.12, p < .001$).

These data indicate that with *OVRlap*, participants move as little as they have to. For the monitoring task, this means turning their head around once to spot the target. For the object collection task, this means looking down and around once to spot the coins and the remainder of the motion coming from moving the coins into the chest. This is an immensely significant gain over having to individually check each location for coins and targets.

5.5 Questionnaire Data

The questionnaire results are summarized in Table 1.

We ran pairwise T-tests with Bonferroni correction to identify significant differences between the *OVRlap* and cycling technique conditions. No significant differences were found for general presence ($p = .93$) or any IPQ subscales. While *OVRlap* ranks above the average for the SUS scale (which is 68 on a scale of 0-100) and the cycling condition ranks slightly below it, the difference between them was not significant ($p = .18$). There was similarly no significant difference in the VEQ body ownership scores ($p = .82$).

6 DISCUSSION

The results from the study show that the *OVRlap* concept brings significant benefits in two different tasks. We found a significant reduction in task completion time and physical motion exerted by the user's head during each task. Further, we found comparable levels of presence, body ownership, and usability for the technique. Lastly, users switched the active perspective more times than is required to complete the tasks.

6.1 The Influence of Environment and Viewpoint Count

The study results show that the technique performed well in two tasks in one environment with four viewpoints. Other environments and viewpoint counts might affect the performance of the technique.

The implementation of the technique used in the study will likely translate poorly to certain types of environments. In visually complex environments that strive for realistic graphics and with many virtual objects, it will be more difficult to distinguish overlapping objects from different viewpoints due to visual occlusion. In an environment with lower visual complexity, more viewpoints can be distinguished. A visually complex environment can also support fewer viewpoints. However, even in a environment such as the one used in the study, there is an upper limit to the number of viewpoints that can be generated before the locations become difficult to distinguish (see Figure 11). The representation component of the technique can be adjusted to better accommodate different types of environments.

The representation component of a viewpoint can be changed to intelligently filter the content instead of naïvely overlaying all objects from each viewpoint. Building on data visualization concepts such as sampling lenses [12], a viewpoint could include only those virtual objects that are relevant for interaction, such as objects that can be selected and manipulated. While this would negatively affect the ability to spatially orient oneself in the environment, it would preserve the ability to preemptively orient oneself for interaction.

Another approach to allow for visually complex environments or higher viewpoint counts could be to introduce a rank-order version of the technique. In a rank-order version, the level of opacity for a viewpoint is not static or binary, but changes gradually based on the computed usefulness of the viewpoint (similar to [29]). For example, an algorithm could infer a user's intent to interact in a given viewpoint, similar to what is proposed for automated switching (see Section 6.2). That viewpoint could be given a higher opacity

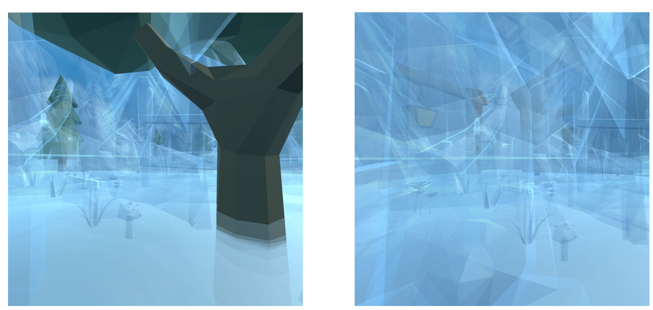

Figure 11: Visual complexity from increasing the number of viewpoints from four to eight (left) or sixteen (right). There is a limit to the number of viewpoints that can meaningfully be overlaid with our implementation of the technique.

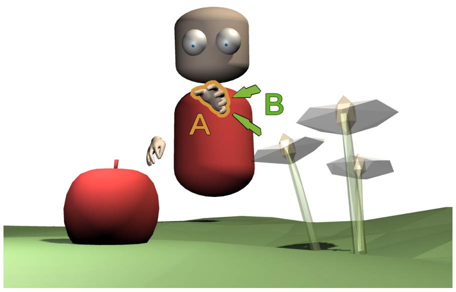

Figure 12: Proposal for automated switching with hands. A hypothetical model could integrate such features as (A) hand state and (B) distance and velocity of the hand relative to each object to determine the user's intent to interact with the apple rather than the flower and make the apple's viewpoint active with no user input.

than others, with only the completely opaque viewpoint retaining the properties of an active viewpoint.

6.2 Improving Viewpoint Switching

In the study, participants switched the active viewpoint by pressing a joystick on their controller either left or right. That switched between the four viewpoints arranged in a randomized circular list. Our data indicate that despite general performance improvements, participants switched more times than they need to accomplish their tasks. Thus, this suggests that improving the method for switching the active perspective might lead to even greater performance improvements.

Based on our switching data, we suggest to reduce the number of switches with automated switching to improve the technique. This would further allow the technique to function without the use of controllers in hand-tracking based VR experiences. A hypothetical approach to such an automated switching model could rank each viewpoint based on the likelihood that it is the one that a user would like to interact with. The viewpoint with the best fit would then be active. Here, we discuss features we propose to be relevant for a hypothetical model for automated switching, as illustrated in Figure 12 and 13.

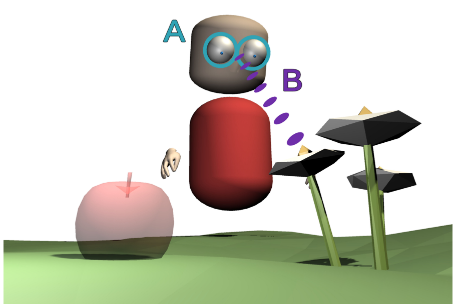

Figure 13: Proposal for automated switching with gaze. A hypothetical model could integrate such features as (A) gaze dynamics and (B) gaze point to determine the user's intent to interact with the flower rather than the apple and make the flower's viewpoint active with no user input.

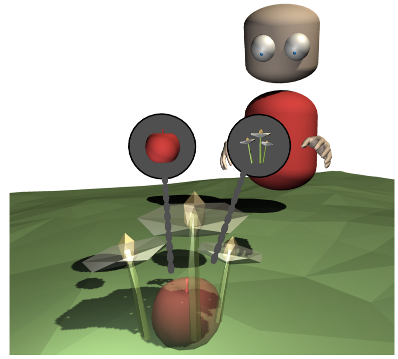

Figure 14: Proposal for overlapping objects. An apple and a flower from two distinct passive viewpoints overlap spatially relative to the user. To disambiguate such an overlap when the objects are used to switch the active perspective, we propose using a preview that instead would show the objects side by side, allowing the user to quickly indicate which they want to interact with.

One set of proposed features for determining interaction intent uses the user's hands. For out-of-reach objects, gestures such as pointing or circling [17] of objects could be used to express the user's intended viewpoint. For objects in the user's reach, the active viewpoint may be based on a hand's proximity to interactable objects in each given environment. Beyond proximity, other relevant hand features are be the direction in which the user is moving their hand, its current velocity, as well as the state of its fingers (e.g., reaching, grasping, closed). In case multiple objects in distinct viewpoints overlap spatially, a simple UI switch that the user can touch might allow users to make the disambiguation of which object they wish to interact with (for example, see Figure 14).

Another set of proposed features uses gaze tracking. Here, the user's gaze tracking point can be used to infer interaction intent (as in [37], where it used to disambiguate between multiple telepresence layers). However, objects that spatially overlap from different viewpoints would be difficult to distinguish from a simple gaze tracking point. For this reason, it would be useful to draw on gaze dynamics such as fixation detection and gaze velocity, which have been shown to be useful in predicting a user's intent to interact [11]. It is, however, unclear how these features translate to two spatially overlapping objects. A similar disambiguation method as for hands could be used, where if fixation is detected for two overlapping objects, the objects are spatially segmented, allowing the user to quickly gaze at the one they wish to interact with (see Figure 14).

6.3 Limitations and Future Work

In addition to the proposed hypothetical improvements to the technique discussed in Section 6.1 and 6.2, this subsection outlines a number of limitations of our work and future directions for the *OVRlap technique*. While we expected the technique to outperform the control condition in completion speed and physical motion, the comparable levels of subjective measures are surprising. We reason this could be due to the sampling of users, who were experienced VR users from a VR forum and might be prone to excitement about novel forms of interacting in VR. It is also possible that the limited sample size for the subjective data could affect the significance of those results.

A conceptual feature of the *OVRlap* technique is that, based on viewpoint placement, users might see additional instances of their avatar-selves at a distance, moving synchronously. It is likely that this might influence the user's sense of body ownership, and it certainly could be used to systematically manipulate components of embodiment [22]. For example, we expect that this could be used to influence the sense of self-location, since it has been shown that it is malleable and can be manipulated towards avatars viewed from a third-person perspective given synchronous visuotactile stimulation [27, 28]. Further, a broader question is that if the user experiences a comparable sense of body ownership with *OVRlap*, as our findings suggest, then which avatar-self or selves do they feel this sense towards? Is it the avatar in the active viewpoint, the ones they see at a distance, or a context-dependent mix of these? A future study could systematically compare how different counts and placements of *OVRlap* selves affects ownership through the visible presence of additional avatars.

Although a main benefit of the *OVRlap* technique is the ability to spatially orient oneself in multiple environments at once, we did not specifically measure its effects on spatial memory. For example, it would be interesting to measure the users' ability to reconstruct the virtual space after using the technique, compared against other techniques that break the first-person perspective to present an overview of the virtual environment, such as WIM [13, 42].

The evaluation of the technique in this paper used pre-determined viewpoint placements. Future work could explore the ability to navigate through placement and manipulation of *OVRlap* viewpoints. One approach could extend WIM [42] by visualizing previews of viewpoint placements on a 3D miniature of the environment (similar to [13]). The technique could also extend teleportation [6], and an *OVRlap* viewpoint could be used to have a first-person preview of the teleportation target. This would allow precision teleportation to distant objects, as the user would be able to look around to orient themselves around their teleportation target before actually teleportation. Further, it would allow teleportation through solid and opaque objects such as walls as rocks, which typically has been impossible as the user would not be able to know where they would end up.

In our study, participants were instructed to remain within a stationary boundary calibration and the tasks were designed to be completed without locomotion. Participants would only turn and lean. It is possible that walking through an environment with many active *OVRlap* viewpoints would influence the sense of VR sickness [25, 41] due to the addition of vection and navigating through numerous objects from multiple viewpoints. While our study did not investigate VR sickness due to its stationary nature, future implementations of *OVRlap* that feature locomotion should consider this.

It is theoretically possible to integrate other senses than the visual into the *OVRlap* concept, which can be explored in future work. For example, using spatialized audio cues from different viewpoint locations can likely add to the ability to distinguish them. A more complex sensory integration is haptics, by adding different strengths of haptic feedback as the user touches active and inactive objects. The strength of the haptic feedback can be tied to the viewpoint's opacity. Similarly, if the viewpoints are ordered by an algorithm, for example as proposed for automatic switching (Figure 13 and 12), haptic stimuli could be a way for a user to "feel" how active a passive object that they are touching is. A haptics integration should investigate how an object that is "there", but not quite – such as a passive *OVRlap* viewpoint – should feel.

7 CONCLUSION

We have introduced *OVRlap*, a technique that allows users to perceive multiple distinct and distant locations at once from a first-person perspective by rendering multiple layered viewpoints on top of each other. To enable meaningful interaction, only one of many viewpoints is active at any time and interaction is only possible in this active viewpoint.

We evaluated the technique in a study where it is compared against a control condition where the user sees only one viewpoint at a time, as in traditional VR. We find that the technique offers a significant reduction in completion time and physical motion exerted by the user's head in an object collection task and a monitoring task in a sample outdoor environment with four distinct locations. Further, the technique affords comparable results for presence, body ownership, and usability measures.

We discussed how these promising initial findings can inform future versions of the technique by exploring the relationship between the visual complexity of the environment and the number of generated viewpoints, and how the switching functionality for the active viewpoint might be improved through automatization. Altogether, our work suggests that a significantly improved overview and interaction space for distinct and distant locations can be achieved by letting a user perceive and spatially orient in such locations at the same time.

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement 648785).

REFERENCES

[1] Majed Al Zayer, Paul MacNeilage, and Eelke Folmer. 2020. Virtual Locomotion: A Survey. *IEEE Transactions on Visualization and Computer Graphics* 26, 6 (jun 2020), 2315–2334. https://doi.org/10.1109/TVCG.2018.2887379

[2] Mark Armstrong, Keitaro Tsuchiya, Feng Liang, Kai Kunze, and Yun Suen Pai. 2020. Multiplex Vision: Understanding Information Transfer and F-Formation With Extended 2-Way FOV. In *26th ACM Symposium on Virtual Reality Software and Technology* (Virtual Event, Canada) *(VRST '20)*. Association for Computing Machinery, New York, NY, USA, Article 33, 9 pages. https://doi.org/10.1145/3385956.3418954

[3] Aaron Bangor, Philip T. Kortum, and James T. Miller. 2008. An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction* 24, 6 (2008), 574–594. https://doi.org/10.1080/10447310802205776

[4] Joanna Bergström, Tor-Salve Dalsgaard, Jason Alexander, and Kasper Hornbæk. 2021. How to Evaluate Object Selection and Manipulation in VR? Guidelines from 20 Years of Studies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 533, 20 pages. https://doi.org/10.1145/3411764.3445193

[5] Jiwan Bhandari, Paul MacNeilage, and Eelke Folmer. 2018. Teleportation without spatial disorientation using optical flow cues. *Proceedings - Graphics Interface* 2018-May (2018), 153–158. https://doi.org/10.20380/GI2018.20

[6] D.A. Bowman, D. Koller, and L.F. Hodges. 1997. Travel in immersive virtual environments: an evaluation of viewpoint motion control techniques. In *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*. IEEE, Albuquerque, NM, USA, 45–52. https://doi.org/10.1109/VRAIS.1997.583043

[7] Doug A. Bowman, Elizabeth T. Davis, Larry F. Hodges, and Albert N. Badre. 1999. Maintaining Spatial Orientation during Travel in an Immersive Virtual Environment. *Presence: Teleoperators and Virtual Environments* 8, 6 (dec 1999), 618–631. https://doi.org/10.1162/105474699566521

[8] Doug A. Bowman, Donald B. Johnson, and Larry F. Hodges. 2001. Testbed evaluation of virtual environment interaction techniques. *Presence: Teleoperators and Virtual Environments* 10, 1 (2001), 75–95. https://doi.org/10.1162/105474601750182333

[9] John Brooke. 1996. *SUS – a quick and dirty usability scale*. Taylor & Francis, London, 189–194.

[10] Wen Yeo Chen, Hsu Chia Huang, Yen Tung Lee, and Caleb Liang. 2018. Body ownership and the four-hand illusion. *Scientific Reports* 8, 1 (2018), 1–17. https://doi.org/10.1038/s41598-018-19662-x

[11] Brendan David-John, Candace Peacock, Ting Zhang, T. Scott Murdison, Hrvoje Benko, and Tanya R. Jonker. 2021. Towards Gaze-Based Prediction of the Intent to Interact in Virtual Reality. In *ACM Symposium on Eye Tracking Research and Applications* (Virtual Event, Germany) *(ETRA '21 Short Papers)*. Association for Computing Machinery, New York, NY, USA, Article 2, 7 pages. https://doi.org/10.1145/3448018.3458008

[12] Geoffrey Ellis, Enrico Bertini, and Alan Dix. 2005. The Sampling Lens: Making Sense of Saturated Visualisations. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (Portland, OR, USA) *(CHI EA '05)*. Association for Computing Machinery, New York, NY, USA, 1351–1354. https://doi.org/10.1145/1056808.1056914

[13] Carmine Elvezio, Mengu Sukan, Steven Feiner, and Barbara Tversky. 2017. Travel in large-scale head-worn VR: Pre-oriented teleportation with WIMs and previews. In *2017 IEEE Virtual Reality (VR)*. IEEE, Los Angeles, CA, USA, 475–476. https://doi.org/10.1109/VR.2017.7892386

[14] Kevin Fan, Jochen Huber, Suranga Nanayakkara, and Masahiko Inami. 2014. SpiderVision: Extending the Human Field of View for Augmented Awareness. In *Proceedings of the 5th Augmented Human International Conference* (Kobe, Japan) *(AH '14)*. Association for Computing Machinery, New York, NY, USA, Article 49, 8 pages. https://doi.org/10.1145/2582051.2582100

[15] A. Folegatti, A. Farnè, R. Salemme, and F. de Vignemont. 2012. The Rubber Hand Illusion: Two's a company, but three's a crowd. *Consciousness and Cognition* 21, 2 (jun 2012), 799–812. https://doi.org/10.1016/j.concog.2012.02.008

[16] Arvid Guterstam, Valeria I. Petkova, and H. Henrik Ehrsson. 2011. The Illusion of Owning a Third Arm. *PLoS ONE* 6, 2 (feb 2011), e17208. https://doi.org/10.1371/journal.pone.0017208

[17] Martin Hachet, Fabrice Decle, Sebastian Knodel, and Pascal Guitton. 2008. Navidget for Easy 3D Camera Positioning from 2D Inputs. In *2008 IEEE Symposium on 3D User Interfaces*. IEEE, Reno, NV, USA, 83–89. https://doi.org/10.1109/3DUI.2008.4476596

[18] Beverly L. Harrison, Hiroshi Ishii, Kim J. Vicente, and William A. S. Buxton. 1995. Transparent Layered User Interfaces: An Evaluation of a Display Design to Enhance Focused and Divided Attention. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '95)*. ACM Press/Addison-Wesley Publishing Co., USA, 317–324. https://doi.org/10.1145/223904.223945

[19] Lukas Heydrich, Trevor J. Dodds, Jane E. Aspell, Bruno Herbelin, Heinrich H. Bülthoff, Betty J. Mohler, and Olaf Blanke. 2013. Visual capture and the experience of having two bodies – Evidence from two different virtual reality techniques. *Frontiers in Psychology* 4, DEC (2013), 1–15. https://doi.org/10.3389/fpsyg.2013.00946

[20] Ludovic Hoyet, Ferran Argelaguet, Corentin Nicole, and Anatole Lécuyer. 2016. "Wow! I Have Six Fingers!": Would You Accept Structural Changes of Your Hand in VR? *Frontiers in Robotics and AI* 3, MAY (may 2016), 1–12. https://doi.org/10.3389/frobt.2016.00027

[21] Shunichi Kasahara, Mitsuhito Ando, Kiyoshi Suganuma, and Jun Rekimoto. 2016. *Parallel Eyes: Exploring Human Capability and Behaviors with Paralleled First Person View Sharing*. Association for Computing Machinery, New York, NY, USA, 1561–1572. https://doi.org/10.1145/2858036.2858495

[22] Konstantina Kilteni, Raphaela Groten, and Mel Slater. 2013. The Sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments* 22, 1 (2013), 373–387.

[23] Ioannis Kotziampasis, Nathan Sidwell, and Alan Chalmers. 2003. Portals: increasing visibility in virtual worlds. In *Proceedings of the 18th spring conference on Computer graphics - SCCG '03*, Vol. 1. ACM Press, New York, New York, USA, 257. https://doi.org/10.1145/984952.984995

[24] André Kunert, Alexander Kulik, Stephan Beck, and Bernd Froehlich. 2014. Photoportals: Shared References in Space and Time. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Baltimore, Maryland, USA) *(CSCW '14)*. Association for Computing Machinery, New York, NY, USA, 1388–1399. https://doi.org/10.1145/2531602.2531727

[25] Eike Langbehn, Paul Lubos, and Frank Steinicke. 2018. Evaluation of Locomotion Techniques for Room-Scale VR: Joystick, Teleportation, and Redirected Walking. In *Proceedings of the Virtual Reality International Conference - Laval Virtual* (Laval, France) *(VRIC '18)*. Association for Computing Machinery, New York, NY, USA, Article 4, 9 pages. https://doi.org/10.1145/3234253.3234291

[26] Joseph J. LaViola, Daniel Acevedo Feliz, Daniel F. Keefe, and Robert C. Zeleznik. 2001. Hands-Free Multi-Scale Navigation in Virtual Environments. In *Proceedings of the 2001 Symposium on Interactive 3D Graphics (I3D '01)*. Association for Computing Machinery, New York, NY, USA, 9–15. https://doi.org/10.1145/364338.364339

[27] Bigna Lenggenhager, Michael Mouthon, and Olaf Blanke. 2009. Spatial aspects of bodily self-consciousness. *Consciousness and Cognition* 18, 1 (mar 2009), 110–117. https://doi.org/10.1016/j.concog.2008.11.003

[28] Bigna Lenggenhager, Tej Tadi, Thomas Metzinger, and Olaf Blanke. 2007. Video ergo sum: Manipulating bodily self-consciousness. *Science* 317, 5841 (2007), 1096–1099. https://doi.org/10.1126/science.1143439

[29] Henry Lieberman. 1994. Powers of Ten Thousand: Navigating in Large Information Spaces. In *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology* (Marina del Rey, California, USA) *(UIST '94)*. Association for Computing Machinery, New York, NY, USA, 15–16. https://doi.org/10.1145/192426.192434

[30] James Liu, Hirav Parekh, Majed Al-Zayer, and Eelke Folmer. 2018. Increasing Walking in VR Using Redirected Teleportation. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. Association for Computing Machinery, New York, NY, USA, 521–529. https://doi.org/10.1145/3242587.3242601

[31] Daniel Mendes, Daniel Medeiros, Maurício Sousa, Eduardo Cordeiro, Alfredo Ferreira, and Joaquim Jorge. 2017. Using iterative refinement for out-of-reach selection in VR. *Proceedings - SCCG 2017: 33rd Spring Conference on Computer Graphics* 2017-May (2017), 237–238. https://doi.org/10.1145/3154353.3154371

[32] Niels Christian Nilsson, Stefania Serafin, and Rolf Nordahl. 2016. Walking in Place Through Virtual Worlds. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Masaaki Kurosu (Ed.). Lecture Notes in Computer Science, Vol. 9732. Springer International Publishing, Cham, 37–48. https://doi.org/10.1007/978-3-319-39516-6_4

[33] Henning Pohl, Klemen Lilija, Jess McIntosh, and Kasper Hornbæk. 2021. Poros: Configurable Proxies for Distant Interactions in VR. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 532, 12 pages. https://doi.org/10.1145/3411764.3445685

[34] Ivan Poupyrev, Mark Billinghurst, Suzanne Weghorst, and Tadao Ichikawa. 1996. The go-go interaction technique. In *Proceedings of the 9th annual ACM symposium on User interface software and technology - UIST '96*. ACM Press, Seattle, WA, USA, 79–80. https://doi.org/10.1145/237091.237102

[35] Holger Regenbrecht and Thomas Schubert. 2002. Real and Illusory Interactions Enhance Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 11, 4 (aug 2002), 425–434. https://doi.org/10.1162/105474602760204318

[36] Daniel Roth and Marc Erich Latoschik. 2020. Construction of the Virtual Embodiment Questionnaire (VEQ). *IEEE Transactions on Visualization and Computer Graphics* 26, 12 (dec 2020), 3546–3556. https://doi.org/10.1109/TVCG.2020.3023603

[37] MHD Yamen Saraiji, Shota Sugimoto, Charith Lasantha Fernando, Kouta Minamizawa, and Susumu Tachi. 2016. Layered Telepresence: Simultaneous Multi Presence Experience Using Eye Gaze Based Perceptual Awareness Blending. In *ACM SIGGRAPH 2016 Emerging Technologies* (Anaheim, California) *(SIGGRAPH '16)*. Association for Computing Machinery, New York, NY, USA, Article 14, 2 pages. https://doi.org/10.1145/2929464.2929467

[38] Jonas Schjerlund, Kasper Hornbæk, and Joanna Bergström. 2021. Ninja Hands: Using Many Hands to Improve Target Selection in VR. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 130, 14 pages. https://doi.org/10.1145/3411764.3445759

[39] Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. 2019. Using Presence Questionnaires in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300590

[40] Maxim Spur, Vincent Tourre, Erwan David, Guillaume Moreau, and Patrick Le Callet. 2020. Exploring multiple and coordinated views for multilayered geospatial data in virtual reality. *Information (Switzerland)* 11, 9 (2020), 1–31. https://doi.org/10.3390/info11090425

[41] Frank Steinicke, Gerd Bruder, Jason Jerald, Harald Frenz, and Markus Lappe. 2010. Estimation of detection thresholds for redirected walking techniques. *IEEE Transactions on Visualization and Computer Graphics* 16, 1 (2010), 17–27. https://doi.org/10.1109/TVCG.2009.62

[42] Richard Stoakley, Matthew J Conway, and Randy Pausch. 1995. Virtual reality on a WIM. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '95*. ACM Press, New York, New York, USA, 265–272. https://doi.org/10.1145/223904.223938

[43] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[44] Andrea Stevenson Won, Jeremy Bailenson, Jimmy Lee, and Jaron Lanier. 2015. Homuncular Flexibility in Virtual Reality. *Journal of Computer-Mediated Communication* 20, 3 (may 2015), 241–259. https://doi.org/10.1111/jcc4.12107

[45] Run Yu, Wallace Lages, Mahdi Nabiyouni, Brandon Ray, Navyaram Kondur, Vikram Chandrashekar, and Doug Bowman. 2017. Bookshelf and Bird: Enabling real walking in large VR spaces. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, Los Angeles, CA, USA, 116–119. https://doi.org/10.1109/3DUI.2017.7893327