Efficient 3D Object Exploration in a Virtual Environment

Category: Research

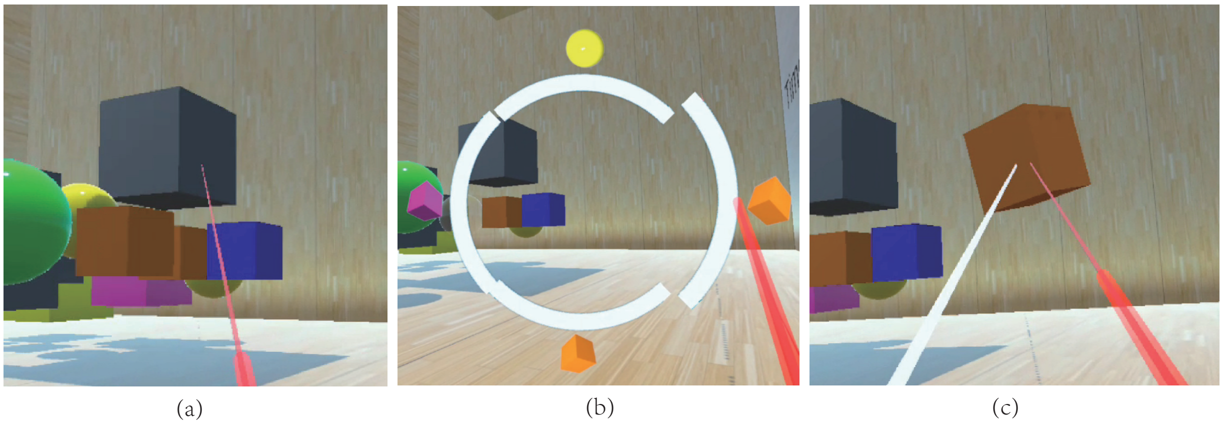


Figure 1: (a) Objects in the virtual environments are overlapped and mutually occluded. The depth axis of the objects are also varied. It's hard to manipulate the objects accurately or efficiently. (b) The proposed object selection method using a radial menu. Users can select the object by pointing at their avatars on the menu or rotating the thumbstick to select the target object. (c) The proposed object translation method is used to translate a selected object in a varied depth task. Users can translate the virtual objects by rotating the thumbstick or specify the target positions by intersecting two rays emitted from two controllers.

Abstract

Interacting with virtual objects is a significant part of VR applications. Various techniques have been proposed to cope with the difficulties in selecting and translating objects in complex virtual environments. However, there are still some limitations. Objects can be hard to interact with in virtual environments when they are mutually occluded or have varied depth (VD) issues. We design two object selection methods to cope with selection tasks with heavy occlusion, and propose two translation methods using thumbstick and cross ray, respectively, to cope with translation tasks with varied depth and reduce motion sickness (MS) problem during interactions. We conduct two studies to evaluate the designed methods by comparing them with the existing methods quantitatively and qualitatively. In the first study, we compare time performance, accuracy and subjective measurements of the proposed techniques with the existing methods. The results show that our techniques significantly outperform the control methods. In the second study, we evaluate the MS values which show that the new designed techniques significantly reduce MS during the interaction in virtual environments.

Keywords: Virtual environment, object selection, object translation, immersive exploration

1 Introduction

Virtual Reality (VR) introduces a new era of immersive experiences and changes the users' interaction with digital environments. The major progress made in this domain has a great contribution in various aspects such as education [41], healthcare [33], construction [44] and game industry [50]. An crucial aspect of this engagement is the ability to interact with objects in the virtual environments, including the ability to select and translate them. As VR applications become increasingly sophisticated, the need for intuitive and effective interaction mechanisms becomes paramount. During interactions with virtual objects, occlusion occurs when objects get hidden due to obstructions from other virtual elements. Ensuring accurate and reliable selection of occluded objects is essential for maintaining a natural and seamless user experience. Moreover, translating virtual objects in situations with varied depth (VD) requires navigating objects at various distances from the user within the VR environment. This task can be intricate due to the complexities of translating physical gestures to virtual motions [46], particularly along the depth axis [2]. Evaluating the efficacy of interaction techniques addressing these challenges is crucial for enhancing user satisfaction, minimizing discomfort, and optimizing the overall VR experience. We delve into a comprehensive evaluation of object selection under occlusion and object manipulation in VD conditions within the realm of VR. We examine current interaction methods, assess their limitations, and propose novel approaches to address these challenges.

A commonly used approach to interact with objects in virtual environments named ray casting. However, in heavy occlusion (HO) scenarios, ray cast interaction becomes harder because the objects are partly or completely blocked. Besides, the small volume of the ray cast activation area makes it difficult to select small and medium-sized objects. Users have to move around the scene to get close to the target object to interact with the object. There are many existing ray casting techniques dealing with these complicated scenarios in virtual environments. Grossman and Balakrishnan proposed the depth ray technique [18]. It operates by selecting the object closest to the cursor along the ray. The spot light technique [31] uses a conic selection volume to improve the performance of ray casting. Andrew Forsberg [14] made an improvement by enabling users to adjust the apex angle of the selection conic volume dynamically. Vanacken proposed the bubble cursor [42] technique to deal with dense 3D environments by automatically resizing the cursor and reach the object closest to the center of it. However, we think the performance of the object selection methods in VR can still be enhanced. In this work, we further explore the use of ray casting in HO (heavy occlusion) environments and propose new selection techniques. Apart from the improvements of the ray selection techniques, the translation of the virtual objects is also a interesting topic, especially in VD environments. When users translate objects in virtual environments, they can experience MS. Moreover, the low efficiency of object translation techniques in VD tasks strengthens the discomfort. In this paper, we seek to explore new techniques for virtual object translation with higher efficiency and lower MS value.

To evaluate the performance of different object selection and object translation methods, we build an experimental prototype using Oculus Quest and conduct two user studies. In the first study, we evaluate the proposed interaction methods and the traditional methods from both objective and subjective perspectives, respectively. The results show that the proposed methods are significantly more efficient and accurate than traditional methods. Furthermore, judging from subjective evaluation results, the proposed methods are more preferred by users. In the second study, we evaluate the MS scale for the participants when using different object translation techniques. The results show the designed methods can significantly reduce MS for users.

2 Related Work

2.1 Interaction Techniques in VR

In the context of evaluating VR interaction methods, it is essential to explore the landscape of VR controllers and their impact on user interactions. The current VR ecosystem provides various controller choices, each designed for distinct modes of interaction in virtual environments. Traditional handheld VR controllers like Oculus Touch and HTC Vive controllers [1] are popular for their ergonomic design and user-friendly button layout.These controllers allow users to grasp, manipulate, and interact with virtual objects in a natural manner. These controllers incorporate haptic feedback [43] mechanisms. These mechanisms further enhance the sense of presence and realism during interactions. Additionally, more recent advancements introduces controller-less interactions through hand tracking approaches [27], and gaze guidance techniques [2], exemplified by devices like the Leap Motion controller [45]. These approaches allow users to directly manipulate virtual objects using their hands. The use of hand gestures can eliminate the need for handheld controllers. However, some researchers argued that the mid-air hand gestures can cause arm fatigue [20] in long-term usage. There are investigations into alternative interaction methods, such as using body gestures [39] or eye-tracking [34]. They also expand the possibilities for moving and interacting with virtual objects. Although there are many new interaction methods, in the development of most games and other applications, it is difficult for developers to ensure that their interaction methods can be adapted to the user's device. Therefore, in most applications, it is a necessary decision to develop a set of convenient, intuitive, and fast handle-based interaction methods.

2.2 Object Translation in a 3D Virtual Environment

/mmA Another issue presented in 3D object exploration is object translation. We summarize some current difficulties of VR object translations in this section.

Translation in depth axis. VR controllers excel at capturing user's motions in horizontal and vertical directions. In VD environments, however, accurately translating virtual objects along the depth axis can be challenging. This is mainly because current methods for moving objects in VR rely on hand and eye tracking technologies [2, 47] and ray casting [4, 37]. The difficulties arise partially due to the limitations of tracking technologies. These methods primarily use sensors and cameras to track the motions. Consequently, when actions involve forward and backward motions, like pushing or pulling virtual objects, the interaction can become less precise and can sometimes be misunderstood. Users may struggle to control depth-based interactions accurately. Moreover, this limitation can hinder the naturalness and realism of object manipulation and affect the immersive VR experience. To address these challenges, new solutions are required, such as improving tracking algorithms, exploring the integration of additional depth-sensing technologies, or developing alternative interaction methods to reduce the need for precise depth-based manipulation. Additionally, these tracking methods relying on sensors and cameras can be dependent on the user's VR devices, making it challenging to ensure consistent support across all user devices. To overcome the difficulties associated with depth-axis interactions, we propose several methods to enable 3D object translation in VR environments using VR controllers.

Motion Sickness (MS). When users select or translate virtual objects by employing hand movements or controller inputs, their operations aren't directly correlated with their physical environment. In such cases, sensory mismatches occur. This discrepancy between visual cues and physical sensations can lead to motion sickness symptoms, such as nausea, disorientation, and discomfort [28]. The translation of objects in virtual environments can inevitably induce motion sickness. In virtual environments, users move constantly in the virtual scenes. However, their physical movements in the real world are limited. Such problem can cause visually induced motion sickness (VIMS) [25] and the sensation of illusory self-motion (Vection) [11, 40]. VIMS is also labeled as cybersickness [35], gaming sickness [36], or simulator sickness [6].

The symptoms of VIMS vary based on hardware [24], gender [16], latency [12] and many other factors. Previous study also demonstrated that vection alone may not cause VIMS. However, vection can be a necessary prerequisite for VIMS when combined with other factors [26]. The most commonly used method to quantify and evaluate VIMS is subjective rating. Kennedy designed the Simulator Sickness Questionnaire(SSQ) [22] to quantify and evaluate subjective symptoms. Behrang Keshavarz and Heiko Hecht introduced Fast Motion Sickness Scale(FMS) [23] to quantify MS more efficiently. Moreover, when users interact with larger virtual objects, they may sense weightlessness or inertia mismatch. This can worsen motion sickness. As we explore and evaluate VR interaction methods, it's vital to consider the possibility of MS and design interaction approaches that reduce discomfort and ensure a smooth user experience.

2.3 Object Selections in a 3D Virtual Environment

It challenging to select objects in a 3D virtual environment, particularly in an HO scenario, because virtual objects are partially or completely hidden [48] from the controller's tracking sensors. Maintaining accurate interaction fidelity then becomes more difficult. Occlusions can result in misinterpretations of user's intentions and lead to unintended actions or interactions. Addressing the complexities of occlusion-related challenges is vital for developing robust and user-centric VR interaction experiences. Many previous techniques were proposed to address the issue. One such solution is gaze-guidance [2]. It requires hand-eye capture components on the device and becomes less efficient when dealing with numerous occluded objects because it is difficult to confirm the object by adjusting the hands. Grossman proposed Flow Ray disambiguation mechanism. When the ray intersects with multiple objects in a virtual scene, all the objects are displayed on a radial menu in order of the depth axis. Kopper proposed the SQUAD technique. In the first step of SQUAD, a spherical ray is used to determine the selectable objects. In the second step, a quad menu is displayed to let users select the target object.

2.4 Traditional Ray Casting Detection Selection Method (RCD)

Ray casting is a commonly used object selection and interaction technique [30]. Ray casting is also widely used in virtual reality (VR) applications. It works by emitting a ray from either the user's viewpoint or the controller's position, and checking the ray's intersection with other objects. In VR applications, the process of selecting objects through ray casting is usually as follows: first, the user holds a controller or wears gloves to simulate real-world hand movements. Then, by tracking the user's gesture and the controller's position, we can obtain the user's viewpoint and interaction position in the virtual environment. Next, we emit a ray. The ray advances along the line of sight from the user's viewpoint or controller position. The ray keeps marching until it hits the objects in the scene. When the hit happens, we can pinpoint the intersected object. In this way, we can identify the object user selects. The selection of objects in VR through ray casting has some advantages [38]. Firstly, it provides an intuitive and natural way for users to interact directly with objects in the virtual environment through gestures and controller positions. Secondly, ray casting has a fast response speed and enables users to interact more smoothly. In addition, ray casting also has good adaptability for complex object shapes and scenes, and can accurately select objects in all shapes and scales. However, sometimes it's hard to intersect the ray with the farther and smaller virtual objects [21]. [] To solve this, previous researchers proposed Bubble Mechanism [32] to identify and select the target closest to the ray. This technique significantly improves performance and preference of ray-casting. However, ray casting also has some drawbacks. Ray casting can lack realism in complex scenes. When the ray only considers the first object it intersects with, it can make interaction troublesome and inaccurate in HO environment. Consequently, users can be limited by occlusion when selecting objects, especially in complex scenes [8].

In summary, traditional ray casting, as a common object selection technique, plays an important role in VR object explorations. It provides an intuitive and real-time interaction method, suitable for various scenarios. However, it also has some limitations, such as limited visibility, lack of accuracy, and susceptibility to occlusion.

3 Exploring Interaction with Radial Menu in VR

~~**Implications from Fitts's Law.** Fitts's law [13] implicates that the difficulty level of selection and translation tasks are decided by two key parameters: distance and width of the target object. The equation 1 implicates that in virtual environments, to make the selection task easy, the target should be as big and as close as possible. Many efficient and high-performance object manipulation mechanisms derive from this law.~~

Evaluation of pointing tasks. Fitts's law [13] implicates that the difficulty level of selection and translation tasks are decided by two key parameters: distance and visual width of the target object.

$$M_t = a + b \cdot \log_2\left(\frac{D}{W} + 1\right) \tag{1}$$

The equation 1 implicates that in virtual environments, to make the selection task easy, the target should be visually as big and as close as possible. Many efficient and high-performance object manipulation mechanisms derive from this law. Another way to evaluate the difficulties of interaction tasks in 3d virtual environments is the predictive model proposed by Kopper [29].

$$ID_{ANGULAR} = \log_2\left(\frac{\alpha}{w^k} + 1\right) \tag{2}$$

The $ID_{ANGULAR}$ value is the index of difficulty. α is the angular amplitude of the wrist movement. The W value is the angular width of the target. The k is a constant power factor.

Exploring the Design of Radial Menu in HO Selection Tasks. In previous researches, James Boritz et al. proved that all items in a radial menu have almost the same difficulty value according to Fitts's law, because they have the same distance from the prime pixel [3]. For right-handed users, selecting the left-most menu item was more difficult than the right-most one. No differences were found for transitions from upper to lower functions and vice versa. The efficiency of radial menu in virtual object manipulation was also proved by the previous researches [10, 15]. When selecting objects in HO environments, target objects are usually partially occluded by other objects in the scene. As is shown in Figure 2, to select the correct occluded objects, users need to shift a certain angle in physical space or an immersive space to avoid the obstruction. Such troublesome movement can make the interaction less efficient. To solve this issue, previous researchers used the method of repositioning the objects into grids [48]. However, selections on linear displayed of menus are less efficient than radial menus [7, 9]. The usage of radial menus in virtual object selection under HO scenarios is yet to be explored. There are also interesting methods exploring the way to display the menus. For example, pop-up menus can reduce time consumed by users to select the target objects compared with fixed menus [9, 49] since users can continue interaction from the current hand position and doesn't have to move to a different preset area.

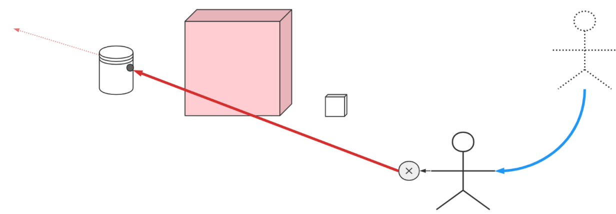


Figure 2: Users need to shift a certain angle in physical space or an immersive space to select the occluded object.

4 Interaction Design

4.1 Conic Ray Activation Area

When the target position gets farther, ~~it becomes harder to encapsulate target objects in the activation area of the ray since the D element (distance between activation area and user) gets larger, according to the equation (1). In order to keep M_t (movement time) constant, we need to keep the value of D/W constant. We can do this by changing the W element(width of activation area) linearly along with D value.~~ it becomes harder to encapsulate target objects in the activation area of the ray since the W element (the visual size of the object) gets smaller, according to the equation (1). Namely, we should resize the activation area dynamically [17]. We build a conic volume in object selections via VR controllers. Specifically, we design an activation area of the controller ray into a cone. The apex of the cone is located on the user's controller. Its axis direction is parallel to the direction indicated by the handle. Along the axis direction, the radius of the cone increases linearly. The radius can be interpreted as the width of the activation area (W in equation 1). The W element gets larger linearly along the depth axis, keeping the M_t value near to constant, making it easier and more efficient to detect objects with smaller colliders or farther from the user.

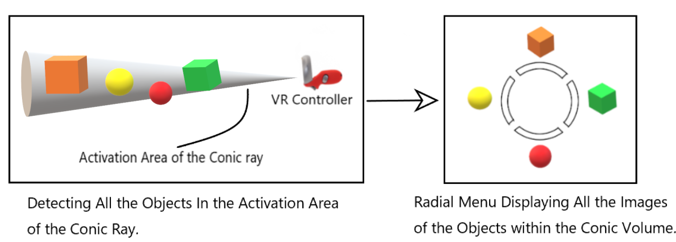


Figure 3: We combine conic ray design into a radial menu. The grey cone represents the conic volume of the ray. The ray is emitted from the VR controller. Using the proposed conic ray activation area, all the objects within the range of the conic volume are currently considered.

In this paper, we develop two object exploration schemes, i.e., object selection and object translation. Regarding the two schemes, we further design two interaction methods for users to choose, and then compare them with the corresponding traditional methods (RCD [38] for selection and HOMER [5] for translation), respectively. Based on the usage of radial menus, we also provide two different ways to interact with the displayed menu.

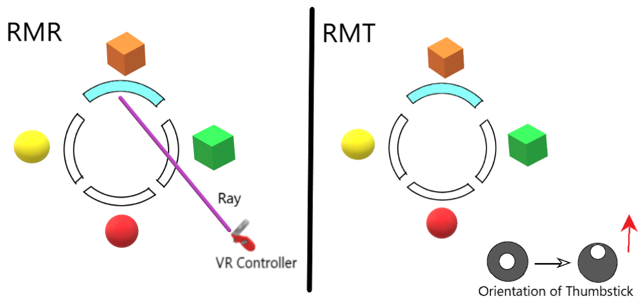


Figure 4: The demonstrations of the proposed object selection methods RMR and RMT. They both use a radial menu as their basic design. In RMR, users can emit a ray to the corresponding avatar images to select the target object on the menu, while in RMT method, they can rotate the thumbstick to the corresponding angle to select the target object on the menu. When users select a candidate avatar on the menu, its corresponding object in the 3D space will be highlighted, other objects will be rendered to be semi-transparent. When user release the trigger button, the candidate object is then selected as the final target object.

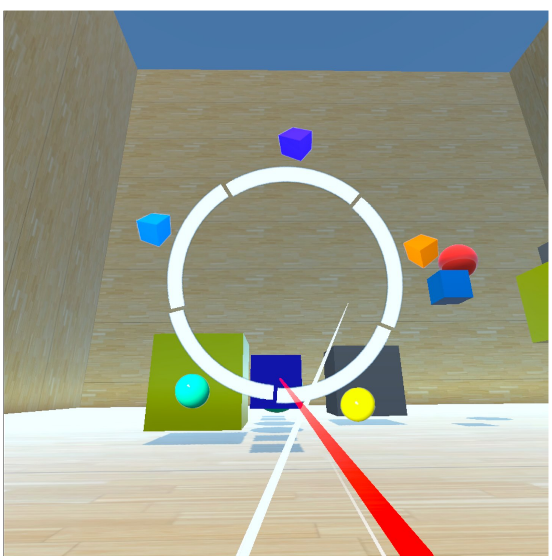


Figure 5: The demonstration of radial menu design in a virtual scene. Only the objects within the conic ray's activation area will be considered as the current candidates and will then be visualized on the radial menu according to their distances to the apex of the cone.

4.2 Radial Menu with Ray Selection (RMR)

The RMR selection process consists of two steps. In the first step, the conic ray will be passed through a set of mutually occluded objects via one of the handles on the controller. The usage of the conic ray distinguishes us from the Flower Ray [18] technique. A radial menu then pops up. All the objects located within the ray's conic activation area are displayed on the radial menu, as is shown in Figure 3. In this step, the trigger button of the controller is clicked to emit the conic ray. The second step is to select and confirm the target object on the menu item by the ray, which is emitted by pressing the same trigger button. Only the target object in the 3D environment will be highlighted in this step, and all the other objects will be rendered to be semi-transparent. The target object is selected when the trigger button is released. The process is shown in Figure 4. This way of interacting with radial menus is similar to the Flower Ray technique [18].

4.3 Radial Menu with Thumbstick Selection (RMT)

Different from the Flower Ray technique, the RMT method displays a new way to interact with the radial menus. The RMT method also applies the radial menu and conic ray casting. The first step of RMT is the same as the first step of RMR. In the second step of RMT, users rotates the thumbstick on the controller with a thumb to select the target object, as is shown in Figure 4. When the thumbstick is rotated to an certain angle, the candidate at the same angle on the radial menu is highlighted. Users can confirm their selection by pressing the trigger button, and they are allowed to cancel the selection by pressing the grip button if they think the selected object is wrong.

The objects with similar distance values (from the viewer) will be placed to the same menu item in RMR and RMT. Then users can shift across the candidate objects (with similar distance values) in an identical menu item bidirectionally via rotating the thumbstick button of the controller.

4.4 Translation by Thumbstick (TB)

After the object is selected, users can move the target object to a given position by the proposed approaches. We also design two object translation mechanisms in this paper. A simple and straightforward method is to use four different buttons on the control to perform directional moving. Two from the left controller and the other two from the right controller. For example, users can rotate the thumbstick on the left controller handle to move the target around horizontally, and use the "A", "B" Button on the right controller to move the object vertically. Users can map four directions to different buttons of the two controllers.

4.5 Translation by Cross Conic Ray (CCR)

Regarding object translation, we also propose another object translation method named CCR to further improve the efficiency compared with TB. In CCR, two rays are emitted from two controllers. The target position can be specified by the intersection of the two rays. Finally, the selected target object (via RMR or RMT) will be directly moved to the target position.

The CCR method also uses the conic ray casting, as shown in Figure 6, we reshape the first ray's activation area into a cone. The apex of the cone is located on the user's controller. Its axis direction is parallel to the direction indicated by the controller. Along the axis direction, the cross-sectional radius of the cone increases linearly. The radius of the cross section can be interpreted as the width of the activation area. When the target intersection point gets farther, the width of the activation area gets larger linearly. The goal of this design is to make M_t value to be constant, according to equation 1. As a result, the translation of the objects doesn't become harder even when the target position is getting farther.

5 User Study 1

It is worth mentioning that in our experiments, we design several tasks for participants to complete. To show the participants whether their current task is completed or not, all the candidate target positions will be highlighted by rendering a semi-transparent cube initially. After the position is confirmed by the participant, the highlighted cube will be rendered to be opaque. In the VR applications, actually, the target positions can be any positions in the VR environment, instead of being restricted by the given positions where the cubes are drawn.

Table 1: The objective metrics of the tasks in Study 1. The first column of the table shows the tasks in the study. The second column shows the objective metrics in each task. The last column shows the detailed descriptions.

Task	Metric	Description
Selection in 3D	Attempt Selection (AS)	Total count of participant's attempts to select the objects. In the task, participants may try multiple times before making one successful selection. AS data is collected by counting the number of times the objects are selected and held.
	Correct Selection (CS)	CS represents the number of successful selections participants complete in the task. When the participant selects the correct object, CS increases by one.
	Accuracy	Accuracy is calculated by dividing CS by AS. The higher the accuracy is, the less irrelevant selections users have to make to select the correct occluded object.
Translation in 3D	Translation Time	The time consumed to translate all the target objects to the target positions. This represents how fast the participant is to move an object to the target position successfully.
Combination tasks in 3D	Total Time	The time consumed to select all the target objects and translate them to the target positions. This represents how fast the participant is to select and move an object to the target position successfully using a specified method.

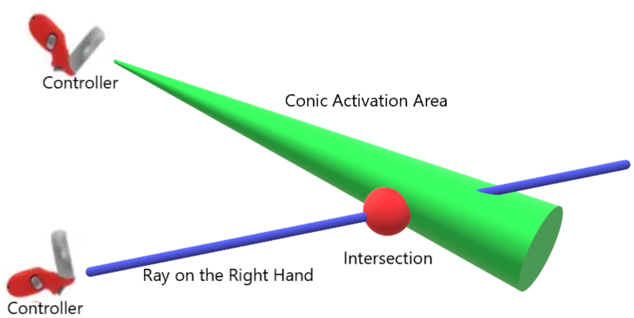


Figure 6: We use the design based on conic ray intersections in the proposed CCR method. The green volume emitted from one of the controllers (e.g., the left controller) represents the conic activation area. The blue line represents the ray emitted from the other controller (e.g., the right controller). The red sphere represents the position of the intersection between the two rays. The target object will be translated to the intersection (in red) in the CCR object translation method.

~~In Study 1, we evaluate and compare the proposed four methods with the corresponding existing methods (i.e., RCD [38] for selection and HOMER [5] for translation).~~ The overall goal of Study 1 is to compare the proposed four methods with the corresponding existing methods (i.e., RCD [38] for selection and HOMER [5] for translation), and explore the tradeoff between them. We evaluate their accuracy and efficiency, including two object selection methods and two object translation methods, respectively. We design three tasks study object selection and translation, one for object selection, one for object translation and the last for their "hybrid" (selection + translation).

In the first task, an HO environment is designed to evaluate the accuracy and efficiency. Participants need to select a target object from a set of occlusions and move it to the target position. We record the task finishing time and the accuracy during the task. The accuracy is calculated by dividing the correct selections by attempt selections. In the second task, participants are required to move virtual objects to different target positions. Similarly, the task finishing time is recorded. In the last task (the task "hybrid"), participants are required to select a specified object, and move it to the target position. Participants have to combine the selection methods with the translation methods. The detailed task procedures of Study 1 are described in Section 5.2.

Table 2: Subjective metrics of the selection tasks and translation tasks in Study 1.

Metric	Description
NASA-TLX [19]	21 point likert scales to measure the participants' mental demand, physical demand, temporal demand, effort, performance and frustration level for all techniques.
Preference Ranking	7 point likert scales to measure participants' personal preference across all techniques.
Open Questions	Open questions about intuitiveness, efficiency level of all techniques and personal suggestions and feedback.

5.1 Participants and Devices

We recruit 22 students in the university (12 males and 10 female) from 18 to 25 years old (mean=20). Twelve of the participants have the experience of using XR applications. We conduct our experiments using Oculus Quest. The software was implemented in C# with Unity.

5.2 Task Design and Experiment Procedure

Participants are first given a detailed introduction to the existing selection method RCD [38] and two proposed selection methods. We design warm-up experiments to ensure participants familiarize with the tasks, respectively. Participants can interact with objects freely in this trial until they fully master each technique. Then, we start the formal task-driven experiments. In the first task, participants need to complete three sub-tasks on object selection in HO environment.

In each sub-task, participants should finish six successful object selections.

After the participants complete all three object selection sub-tasks, we give them a detailed introduction about all the three translation methods. Participants can then master them in the following warm up trial. In the object translation tasks, one of the most challenging problem is the VD issue. The participants are also required to complete six successful translations in each sub-task.

In the end, participants are required to complete the "hybrid" task. We design a scenario with both HO and VD issues. Participants are required to select the correct objects and place them into the given target positions. They have to complete six sub-tasks using the object selection and object translation methods. The individual time each participant consumes to complete six sub-tasks will be recorded. We use the metrics displayed in Table 4.5 to evaluate the objective performance of three different selection methods and three different translation methods. We then use the subjective metrics displayed in Table 4.5 to evaluate the subjective workload preferences of participants.

We utilize the Latin Square method to ensure balanced sequencing of the techniques. Once a task is completed, participants are required to complete the corresponding questions in the questionnaire to gather subjective assessments. They then take a break for 2-5 minutes before proceeding the subsequent tasks. The entire experiment takes approximately 60 minutes. In each experiment, the sub-tasks are displayed in the UI text in front of the participants. Every time the participant finish a sub-task, the virtual environment is refreshed and a new sub-task is assigned.

5.3 Results

5.3.1 Results of Objective Measurements

The results of objective measures are shown in Figure 7. We conduct repeated-measures ANOVA ($a = 0.05$) and post hoc pairwise tests to analyze the data from the study. We first analyze objective measures for object selection.

Results on the total selection time demonstrate that participants spend significantly more time when using the control group method RCD [38] than the proposed RMR and RMT (RCD - RMR, $p < 0.001$; RCD - RMT, $p < 0.001$) in HO task. The results prove that the proposed object selection methods are more efficient when selecting objects in HO environments. Selection accuracy of RCD method is significantly lower (RCD - RMT, $p < 0.001$, RCD - RMR, $p < 0.001$) than the proposed two methods. Both RMT and RMR are more than twice as accurate as RCD [38] (mean = 0.23 for RCD, mean = 0.58 for RMT, mean = 0.56 for RMR). The larger accuracy level indicates that participants make less irrelevant selections.

The objective measures for object translation tasks in VD condition can be also achieved. Results of object translation time show that the control method HOMER takes significantly more time to complete the correct movement in VD tasks (HOMER - TB, $p=0.0013$, HOMER - CCR, $p<0.001$). We then analyse the results of "hybrid" tasks in the study to judge whether the total completion time is significantly different across different combinations of selection and translation techniques. The results show that the completion time of both RMT + CCR and RMR + CCR are significantly lower than RMT + TB ($p < 0.001$, $p=0.0034$, respectively). Besides, the completion time of RMR + CCR is also significantly lower than RMR + TB ($p = 0.0231$).

5.3.2 Results of Subjective Measures

Repeated-measures ANOVA on the NASA-TLX questionnaire demonstrates that three selection techniques have a significant main effect on mental demand ($p = 0.0380$), physical demand ($p < 0.001$), temporal demand ($p < 0.001$), effort ($p < 0.001$) and frustration ($p = 0.0367$). It is interesting to see that mental demand of the RMT method is significantly higher than RCD method ($p = 0.0107$). The physical demand of RCD method is significantly higher than other methods (RCD - RMT, $p = 0.0019$; RCD - RMR, $p < 0.001$). The temporal demand of RCD selection method is significantly higher than other methods (RCD - RMT, $p=0.0061$; RCD - RMR, $p<0.001$). The performance of RMR is significantly better than RCD ($p = 0.0211$). The effort value of RCD is significantly higher than RMT ($p = 0.01290$). The frustration of RCD is significantly higher than RMR ($p = 0.0160$). It is also interesting to see that the frustration level of the proposed RMR method is significantly lower than the proposed RMT method ($p < 0.001$).

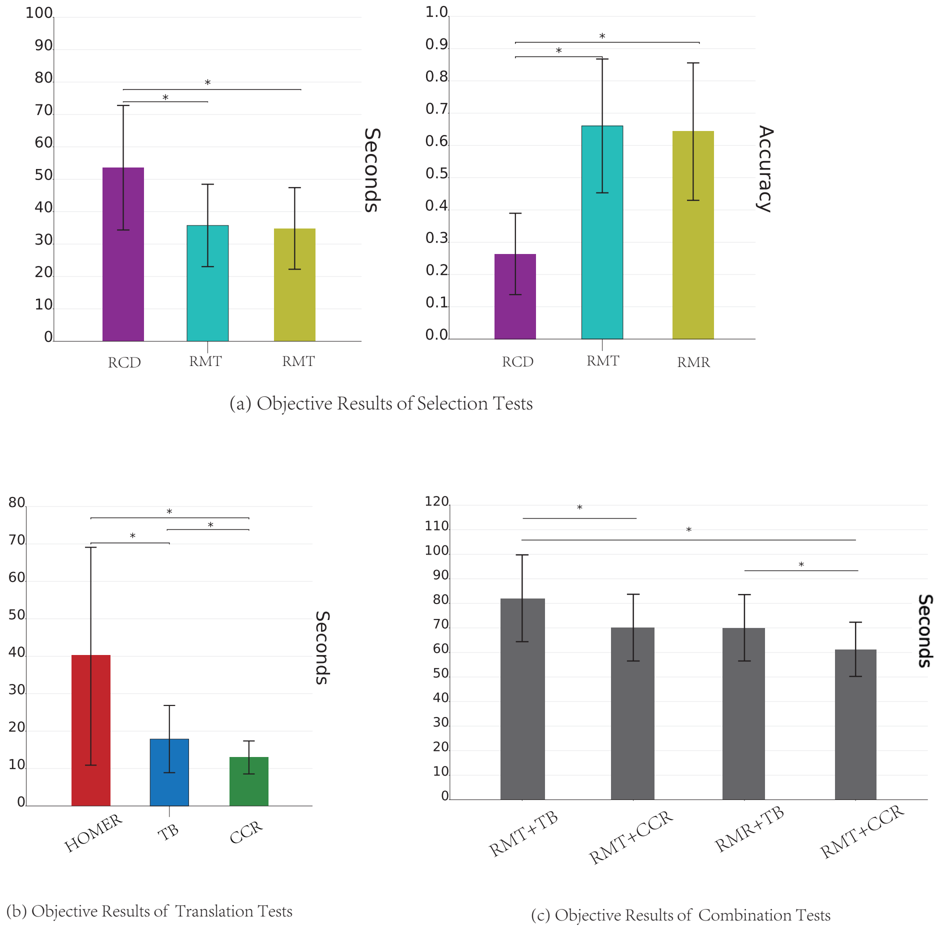
Regarding the results of object translation evaluation, the mental workload of the proposed CCR technique is significantly higher than the other two techniques, i.e., the existing method HOMER [5] and the proposed straight-forward method TB (HOMER - CCR, $p<0.001$; TB - CCR, $p < 0.001$), and temporal demand of the CCR technique is significantly lower than other techniques ($p < 0.001$ for all). The physical demand of the traditional method is significantly higher than the proposed two methods (HOMER - TB, $p = 0.0021$; HOMER - CCR, $p < 0.001$). The same trend continues for the workloads of performance (HOMER - TB, $p = 0.0189$; HOMER - CCR, $p < 0.001$), effort (HOMER - TB, $p < 0.001$; HOMER - CCR, $p < 0.001$), and the frustration (HOMER - TB, $p < 0.001$; HOMER - CCR, $p < 0.001$).

Furthermore, we calculate the total score of NASA-TLX for all of the methods. The final score indicates the total workload scale. As is shown in Figure 10, the total scores of the control group method RCD for object selection are significantly higher than the proposed two methods (RCD - RMT, $p = 0.00912$; RCD - RMR, $p < 0.001$). The same trend continues for the object translation methods (HOMER - TB, $p = 0.0012$, HOMER - CCR, TR - CCR, $p = 0.0020$). This implies that the overall workloads of the proposed methods are significantly lower than that of HOMER. Regarding preference, as is shown in Figure 12, participants prefer both RMR and RMT than RCD method (RCD - RMT, $P < 0.001$, RCD - RMR, $P < 0.001$). For the two proposed methods, the preference of RMR is significantly higher than RMT ($p = 0.0029$), because they think ray-based design (RMR) in the final step of object selections is more efficient and intuitive than thumbstick click (RMT).

As for the object translation tasks, participants prefer both TB and CCR than the traditional HOMER method (HOMER - TB, $p < 0.001$, HOMER - CCR, $p < 0.001$). We question the participants' subjective feeling of intuitiveness and efficiency of different techniques in open question section of the questionnaire. Though the results show that most participants think that the HOMER are the most intuitive, none of them think the HOMER method is more efficient. We then analyse participants' comments from the open question. Some participants ($N = 2$) mentioned that the movement speed in TB translation method should be adjustable according to user's preference. Some others ($N = 4$) also mentioned that more feedback should be displayed when interacting with menus. One participant mentioned that the size of radial menu should vary according to the quantity of objects on it, so that the menu doesn't look too crowded when there are too many objects displayed. Two participants commented that the CCR method is the most efficient, but a little harder to master.

5.4 Summaries of Key Findings in Study 1

Study 1 demonstrates that both from objective and subjective perspectives, the proposed techniques outperforms the existing techniques in terms of efficiency, accuracy, and user preference. When using the existing methods, participants have to select and move obstructions away before selecting the target. However, for the proposed RMT and RMR, participants can select the target directly on the radial menu. The menu allows users to disregard occlusion and select objects directly. This eliminates numerous ineffective or irrelevant selections, and makes the process more efficient and more accurate. With this advantage, both the RMR and the RMT



(a) Objective Results of Selection Tests

(b) Objective Results of Translation Tests

(c) Objective Results of Combination Tests

Figure 7: Objective results of Study 1. Bar charts represent the objective performance of different techniques. Error bars represent the standard error. (a) The time performance and accuracy of three object selection techniques. (b) The time performance of three translation techniques. (c) Time performance of the combination of two new selection techniques and two new translation techniques.

method yield more efficient outcomes compared to traditional selection method RCD. Moreover, the proposed methods are harder to master since they have a higher mental workload.

As for translation methods, the proposed two translation methods significantly outperform the traditional translation method HOMER [5]. The CCR method is the most efficient, but has higher mental workload and is harder to master. That is to say, both the existing object selection and translation methods are easier to learn and use, but less efficient. In the "hybrid" task, we find that the RMR + CCR is the most efficient combination method to select and move a virtual object.

6 User Study 2

In Study 2, we quantify and evaluate the VIMS scale for three object translation techniques. We recruit 10 students in the university (7 males and 3 females) from 19 to 24 years old. About a half (four) of the participants have the experience of using XR applications. The immersive device in Study 2 also is Oculus Quest.

Participants are required to translate the virtual objects to the target positions in dense virtual environment with occlusion and varied depth. The target positions are marked in the scene by translucent cubes. The tasks are given in order by three translation methods. During the tasks, participants are required to use the given object translation method to translate objects into the given target positions. Each task is given 60 seconds to complete (different from Study 1). After participants complete one task, they fill in the corresponding section in the questionnaire. They then have a rest for 5 to 10 minutes until their VIMS during last task completely disappears. The entire experiment takes approximately 30 minutes.

We exploit **SSQ [22]** to evaluate the VIMS symptoms for the participants (different from Study 1). In SSQ, 4 point likert scales are used to measure 16 symptoms correlated with simulator sickness. The Nausea (N), Oculomotor (O) and Disorientation (D) scores and Total Score (TS) are then calculated. The results of SSQ are shown in Figure 13. We conduct repeated-measures ANOVA and post hoc pairwise tests to judge whether the sickness symptoms of the different methods are significantly different. The results show that in the SSQ, both the subscales and the total score of traditional methods are significantly higher than the proposed TB method and the proposed CCR method in all four indices ($p < 0.01$ for all). Using the proposed object translation methods, users don't have to constantly move around the scene to translate the objects to the target positions. This reduce the Vection caused by constant movements in the virtual environment, and further result in less MS.

7 Limitations and Future Work

Despite our best effort, some limitations still exists in the proposed methods.

- Some participants reported that the lack of haptic feedback makes them confused.

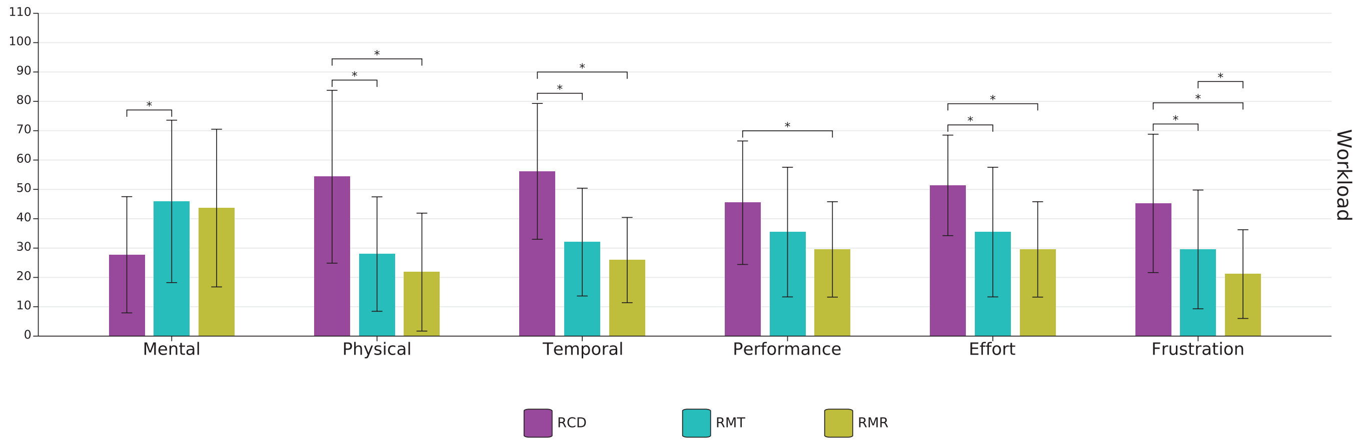


Figure 8: Subjective subscales of NASA-TLX for three different selection methods. The error bars show the standard errors. The statistical significance is labeled with * (p < 0.05).

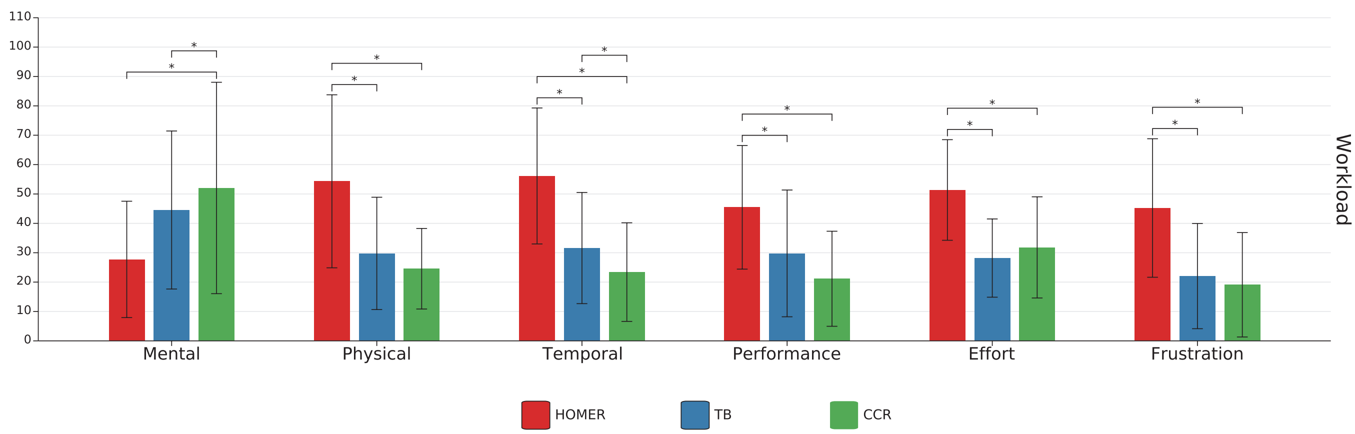
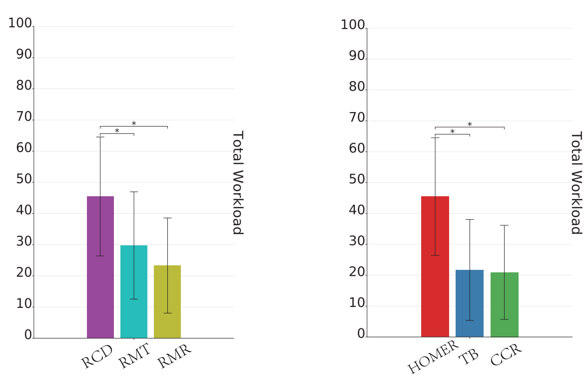


Figure 9: Subjective subscales of NASA-TLX for three different translation methods. The error bars stand for standard errors. The statistical significance is labeled with * (p < 0.05).



(a) Selection (b) Translation

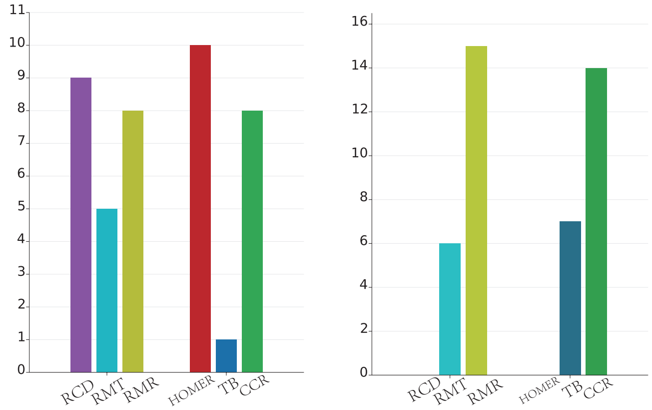
Figure 10: The NASA-TLX total workload scores for different techniques. Error bars represent the standard errors. The statistical significance is labeled with * (p < 0.05)

- Some others mentioned that the some parameters in the application (speed of movement, volume of the activation area for selections) need to be set adjustable by the users.
- Several participants who are susceptible to MS also reported that they still have VIMS even when they used the newly designed object translation methods. Some other factors of VIMS still need future explorations.
- It should also be noticed that we conduct our experiments using Oculus Quest. There are many types of VR headsets in the current market, and different hardware can lead to different results.

In the future, we will improve the design based on the suggestions from participants. Moreover, it should also be noted that we use the Oculus Quest headset in our experiment. However, there are various VR hardware devices in the current market, and the detailed designs of the hardware (such as the controllers) might be different to some extent.

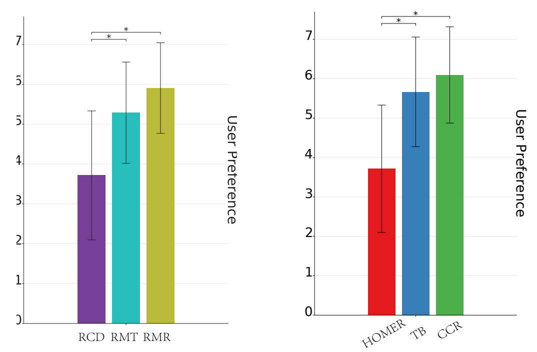
8 Conclusion

We propose object selection methods in HO virtual environments and object translation methods in VD environments in this paper, respectively. We use radial menu and conic ray volume in the HO tasks, and then propose two methods named RMR and RMT to



(a) Most Intuitive (b) Most Efficient

Figure 11: (a) The number of people agree that the corresponding technique is the most intuitive one. (b) The number of people agree that the corresponding technique is the most efficient one.



(a) Selection (b) Translation

Figure 12: The subjective rankings of user preferences. The error bars represent the standard errors. The statistical significance is labeled with * (p < 0.05)

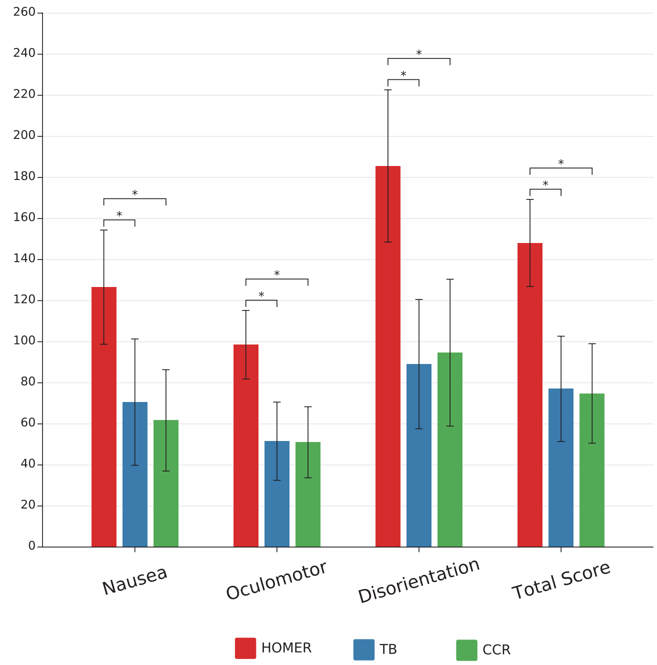


Figure 13: The results of Study 2. The bar charts show three subscales of SSQ for three different translation methods and the total scores for simulator sickness. The error bars represent the standard errors. The statistical significance is labeled with * (p < 0.05)

select objects in HO environment. Then we delve into the challenges of object translation in VD environments. Two object translation methods named TB and CCR are introduced.

We conduct a user study to evaluate the overall performance of the proposed methods and the corresponding traditional methods (Study 1). Results show that the proposed methods are more efficient and more preferred by most participants. Specifically, we evaluate them in terms of accuracy, efficiency and subjective rankings.

Finally, we evaluate the MS values of the three different object translation methods (Study 2). The results show that the proposed two methods significantly reduce MS values for participants compared with the existing method.

References

[1] V. Angelov, E. Petkov, G. Shipkovenski, and T. Kalushkov. Modern virtual reality headsets. In 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1–5, 2020. doi: 10.1109/HORA49412.2020.9152604

[2] Y. Bao, J. Wang, Z. Wang, and F. Lu. Exploring 3d interaction with gaze guidance in augmented reality. In 2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR), pp. 22–32, 2023. doi: 10.1109/VR55154.2023.00018

[3] J. Boritz, K. S. Booth, and W. B. Cowan. Fitts"law studies of directional mouse movement. 1991.

[4] D. A. Bowman and L. F. Hodges. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In Proceedings of the 1997 Symposium on Interactive 3D Graphics, I3D '97, p. 35–ff. Association for Computing Machinery, New York, NY, USA, 1997. doi: 10.1145/253284.253301

[5] D. A. Bowman and L. F. Hodges. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In Proceedings of the 1997 symposium on Interactive 3D graphics, pp. 35–ff, 1997.

[6] J. O. Brooks, R. R. Goodenough, M. C. Crisler, N. D. Klein, R. L. Alley, B. L. Koon, W. C. Logan Jr, J. H. Ogle, R. A. Tyrrell, and R. F. Wills. Simulator sickness during driving simulation studies. Accident analysis & prevention, 42(3):788–796, 2010.

[7] J. Callahan, D. Hopkins, M. Weiser, and B. Shneiderman. An empirical comparison of pie vs. linear menus. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '88, p. 95–100. Association for Computing Machinery, New York, NY, USA, 1988. doi: 10.1145/57167.57182

[8] J. Cashion, C. Wingrave, and J. J. LaViola Jr. Dense and dynamic 3d selection for game-based virtual environments. IEEE Transactions on Visualization and Computer Graphics, 18(4):634–642, 2012. doi: 10.1109/TVCG.2012.40

[9] K. Das and C. W. Borst. An evaluation of menu properties and pointing techniques in a projection-based vr environment. In 2010 IEEE Symposium on 3D User Interfaces (3DUI), pp. 47–50, 2010. doi: 10.1109/3DUI.2010.5444721

[10] M. M. Davis, J. L. Gabbard, D. A. Bowman, and D. Gracanin. Depth-based 3d gesture multi-level radial menu for virtual object manipulation. In 2016 IEEE Virtual Reality (VR), pp. 169–170, 2016. doi: 10.1109/VR.2016.7504707

[11] J. Dichgans and T. Brandt. Optokinetic motion sickness and pseudo-coriolis effects induced by moving visual stimuli. Acta oto-laryngologica, 76(1-6):339–348, 1973.

[12] P. DiZio and J. R. Lackner. Circumventing side effects of immersive

virtual environments. Advances in human factors/ergonomics, 21:893–896, 1997.

[13] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. Journal Information Journal TOC, 1954.

[14] A. Forsberg, K. Herndon, and R. Zeleznik. Aperture based selection for immersive virtual environments. In Proceedings of the 9th annual ACM symposium on User interface software and technology, pp. 95–96, 1996.

[15] S. Gebhardt, S. Pick, F. Leithold, B. Hentschel, and T. Kuhlen. Extended pie menus for immersive virtual environments. IEEE Transactions on Visualization and Computer Graphics, 19(4):644–651, 2013. doi: 10.1109/TVCG.2013.31

[16] J. F. Golding. Motion sickness susceptibility. Autonomic Neuroscience, 129(1):67–76, 2006. Nausea and Vomiting: An Interdisciplinary Approach. doi: 10.1016/j.autneu.2006.07.019

[17] T. Grossman and R. Balakrishnan. The bubble cursor: Enhancing target acquisition by dynamic resizing of the cursor's activation area. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05, p. 281–290. Association for Computing Machinery, New York, NY, USA, 2005. doi: 10.1145/1054972.1055012

[18] T. Grossman and R. Balakrishnan. The design and evaluation of selection techniques for 3d volumetric displays. In Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology, UIST '06, p. 3–12. Association for Computing Machinery, New York, NY, USA, 2006. doi: 10.1145/1166253.1166257

[19] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50(9):904–908, 2006. doi: 10.1177/154193120605000909

[20] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani. Consumed endurance: A metric to quantify arm fatigue of mid-air interactions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14, p. 1063–1072. Association for Computing Machinery, New York, NY, USA, 2014. doi: 10.1145/2556288.2557130

[21] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell. A survey of design issues in spatial input. In Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology, UIST '94, p. 213–222. Association for Computing Machinery, New York, NY, USA, 1994. doi: 10.1145/192426.192501

[22] 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.

[23] B. Keshavarz and H. Hecht. Validating an efficient method to quantify motion sickness. Human Factors, 53(4):415–426, 2011. PMID: 21901938. doi: 10.1177/0018720811403736

[24] B. Keshavarz, H. Hecht, and L. Zschutschke. Intra-visual conflict in visually induced motion sickness. Displays, 32(4):181–188, 2011.

[25] B. Keshavarz, L. J. Hettinger, R. S. Kennedy, and J. L. Campos. Demonstrating the potential for dynamic auditory stimulation to contribute to motion sickness. PloS one, 9(7):e101016, 2014.

[26] B. Keshavarz, B. E. Riecke, L. J. Hettinger, and J. L. Campos. Vection and visually induced motion sickness: how are they related? Frontiers in Psychology, 6, 2015. doi: 10.3389/fpsyg.2015.00472

[27] C. Khundam, V. Vorachart, P. Preeyawongsakul, W. Hosap, and F. Noël. A comparative study of interaction time and usability of using controllers and hand tracking in virtual reality training. Informatics, 8(3), 2021. doi: 10.3390/informatics8030060

[28] E. M. Kolasinski. Simulator sickness in virtual environments. 1995.

[29] R. Kopper, D. A. Bowman, M. G. Silva, and R. P. McMahan. A human motor behavior model for distal pointing tasks. International Journal of Human-Computer Studies, 68(10):603–615, 2010. doi: 10.1016/j.ijhcs.2010.05.001

[30] S. Lee, J. Seo, G. J. Kim, and C.-M. Park. Evaluation of pointing techniques for ray casting selection in virtual environments. In Z. Pan and J. Shi, eds., Third International Conference on Virtual Reality and Its Application in Industry, vol. 4756, pp. 38 – 44. International Society for Optics and Photonics, SPIE, 2003. doi: 10.1117/12.497665

[31] J. Liang and M. Green. Jdcad: A highly interactive 3d modeling system. Computers & Graphics, 18(4):499–506, 1994. doi: 10.1016/0097-8493(94)90062-0

[32] Y. Lu, C. Yu, and Y. Shi. Investigating bubble mechanism for ray-casting to improve 3d target acquisition in virtual reality. In 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 35–43, 2020. doi: 10.1109/VR46266.2020.00021

[33] O. López Chávez, L.-F. Rodríguez, and J. O. Gutierrez-Garcia. A comparative case study of 2d, 3d and immersive-virtual-reality applications for healthcare education. International Journal of Medical Informatics, 141:104226, 2020. doi: 10.1016/j.ijmedinf.2020.104226

[34] P. Majaranta and A. Bulling. Eye Tracking and Eye-Based Human–Computer Interaction, pp. 39–65. Springer London, London, 2014. doi: 10.1007/978-1-4471-6392-3_3

[35] M. E. McCauley and T. J. Sharkey. Cybersickness: Perception of self-motion in virtual environments. Presence: Teleoperators & Virtual Environments, 1(3):311–318, 1992.

[36] O. Merhi, E. Faugloire, M. Flanagan, and T. A. Stoffregen. Motion sickness, console video games, and head-mounted displays. Human factors, 49(5):920–934, 2007.

[37] M. R. Mine. Virtual environment interaction techniques. Technical report, USA, 1995.

[38] T. Nukarinen, J. Kangas, J. Rantala, O. Koskinen, and R. Raisamo. Evaluating ray casting and two gaze-based pointing techniques for object selection in virtual reality. In Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, VRST '18. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3281505.3283382

[39] F. R. Ortega, K. Tarre, M. Kress, A. S. Williams, A. B. Barreto, and N. D. Rishe. Selection and manipulation whole-body gesture elicitation study in virtual reality. In 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 1723–1728, 2019. doi: 10.1109/VR.2019.8798105

[40] S. Palmisano, R. S. Allison, M. M. Schira, and R. J. Barry. Future challenges for vection research: definitions, functional significance, measures, and neural bases. Frontiers in psychology, 6:193, 2015.

[41] P. Smutny, M. Babiuch, and P. Foltynek. A review of the virtual reality applications in education and training. In 2019 20th International Carpathian Control Conference (ICCC), pp. 1–4, 2019. doi: 10.1109/CarpathianCC.2019.8765930

[42] L. Vanacken, T. Grossman, and K. Coninx. Multimodal selection techniques for dense and occluded 3d virtual environments. International Journal of Human-Computer Studies, 67(3):237–255, 2009. Current trends in 3D user interface research. doi: 10.1016/j.ijhcs.2008.09.001

[43] C. Våpenstad, E. F. Hofstad, T. Langø, R. Mårvik, and M. K. Chmarra. Perceiving haptic feedback in virtual reality simulators. Surgical Endoscopy, 27(7):2391–2397, 07 2013. doi: 10.1007/s00464-012-2745-y

[44] P. Wang, P. Wu, J. Wang, H.-L. Chi, and X. Wang. A critical review of the use of virtual reality in construction engineering education and training. International Journal of Environmental Research and Public Health, 15(6), 2018. doi: 10.3390/ijerph15061204

[45] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler. Analysis of the accuracy and robustness of the leap motion controller. Sensors, 13(5):6380–6393, 2013. doi: 10.3390/s130506380

[46] M. Weise, R. Zender, and U. Lucke. How can i grab that? i-com, 19(2):67–85, 2020. doi: doi:10.1515/icom-2020-0011

[47] D. Yu, X. Lu, R. Shi, H.-N. Liang, T. Dingler, E. Velloso, and J. Goncalves. Gaze-supported 3d object manipulation in virtual reality. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, CHI '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411764.3445343

[48] D. Yu, Q. Zhou, J. Newn, T. Dingler, E. Velloso, and J. Goncalves. Fully-occluded target selection in virtual reality. IEEE Transactions on Visualization and Computer Graphics, 26(12):3402–3413, 2020. doi: 10.1109/TVCG.2020.3023606

[49] H. Zhao. Theories in computer human interaction fitts ' law : Modeling movement time in hci. 2010.

[50] M. Zyda. From visual simulation to virtual reality to games. Computer, 38(9):25–32, 2005. doi: 10.1109/MC.2005.297