A Novel Bare-Handed Manipulation Technique for Distant Objects in Virtual Reality

by

Di (Bill) Zhao

Master of Digital Media, Simon Fraser University, 2009

B.A, Communication University of China, 2006

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
School of Interactive Arts and Technology
Faculty of Communication, Art and Technology

SIMON FRASER UNIVERSITY
Spring 2024

Declaration of Committee

Name:	**Di (Bill) Zhao**
Degree:	**Master of Science**
Title:	**A Novel Bare-Handed Manipulation Technique for Distant Objects in Virtual Reality**
Committee:	**Chair** **Kate Hennessy** Associate Professor, Interactive Arts and Technology
	Wolfgang Stuerzlinger Supervisor Professor, Interactive Arts and Technology
	Chris Shaw Committee Member Professor, Interactive Arts and Technology
	Brian Fisher Examiner Professor, Interactive Arts and Technology

Ethics Statement

Abstract

Advancements in bare-hand tracking technology, featured in many modern virtual reality (VR) headsets, have enabled hand interaction for proximate 3D object manipulation. However, manipulating distant 3D objects in VR with bare hands remains a challenge and my work explores the potential of two-handed interaction in this context. Previous work highlighted the ease of using the non-dominant wrist as an interactive surface for interaction, by pointing at it with the dominant hand's index finger. My research expands this paradigm by extending the interactive zone to the entire non-dominant hand, so that the user can interact with their index finger on the non-dominant hand’s fingers and palm to adeptly and precisely manipulate a distant 3D object's position and rotation. Through a user study, I compare my method with HOMER in terms of efficiency, accuracy, and usability. The results demonstrate a notable accuracy improvement (at least 75% better), especially in executing complex tasks, such as hanging a painting on a slanted wall, albeit at the expense of efficiency (about 47% slower) when compared to HOMER. While HOMER is faster, it also exhibits more variation in time, and at the average HOMER end time, my new technique is still substantially more accurate than HOMER. Further, my method received more favorable ratings on the System Usability Scale (SUS), underscoring the user-centric and intuitive design of my approach. Furthermore, the NASA Task Load Index (NASA-TLX) indicates that my technique not only improves task completion rate and user satisfaction but also reduces cognitive, physical, and temporal demands on users. In addition, I discuss insights into error prevention, the integration of constraints, and the facilitation of consistent interaction. My findings lay a solid foundation for the application of my new method in the dynamically advancing field of VR systems operated without controllers.

Keywords: Hand tracking; Virtual reality; Distant object manipulation

Acknowledgements

I extend my deepest gratitude to my supervisor, Dr. Wolfgang, for his invaluable guidance, expertise, and steadfast support throughout my research journey. His consistent encouragement and insightful feedback were instrumental in the development of this thesis. I am profoundly thankful for the opportunity to have him as my supervisor. Additionally, I express my sincere appreciation to my classmates, colleagues, and the volunteers who participated in my user study. Their assistance and constructive feedback were crucial in refining my technique and attaining significant results.

Table of Contents

List of Figures

List of Acronyms

ART	Aligned Rank Transform
ANOVA	Analysis of Variance
AR	Augmented Reality
BMSR	Bimanual Metaphor with Scaled Replica
CAD	Computer-Aided Design
DOF	Degrees of Freedom
DH’s	Dominant Hand's
DWBM	Distant Widget-based Method
HMD	Head-mounted Display
HOMER	Hand-centered Object Manipulation Extending Ray-casting
NASA-TLX	NASA Task Load Index
NDH’s	Non-dominant Hand's
NFWBM	Near-field Widget-based Method
RQ	Research Questions
RM	Repeated Measures
SUS	System Usability Scale
THFP	Two-Hand Fingertip-Palm Technique
THRED	Two-Handed Refining Editor
UI	User Interface
UMSR	Unimanual Metaphor with Scaled Replica
VR	Virtual Reality
VE	Virtual Environment
WIM	World in Miniature

Chapter 1.

Introduction

The landscape of virtual reality (VR) technology has evolved substantially, with real-time hand tracking making notable strides. Today's consumer-level VR headsets offer reasonably accurate bare-hand tracking via the headset's inside-out cameras [4]. This innovation allows users to perform gestures such as pointing and pinching, enabling them to manipulate virtual objects and navigate in a virtual environment (VEs) without the need for physical controllers. This innovation has significantly improved the user experience [75], [14] and has proven useful for tasks requiring more natural body postures [27], [34]. However, despite these advancements, current interaction methods do not fully utilize two-handed interaction, especially when the two hands touch each other. More specifically, the subject of bare-handed manipulation techniques for manipulating distant 3D objects in VR, i.e., those beyond arm's reach, has largely been neglected.

Bowman et al. critically evaluated various techniques for object manipulation in VR in their 1997 survey [12]. They described the go-go technique as having a limited range and affording only imprecise grabbing, while raycasting, although it was particularly good for grabbing/selecting, was challenging for rotation and could not move objects towards or away from the user. Arm extension made it possible to reach remote objects, but did not enable the user to rotate remote objects. They suggested using ray-casting for object-grabbing and hand-centered manipulation. Mendes et al. also suggested a hybrid of touch and mid-air techniques, using mid-air interaction for reaching objects and touch for precision manipulation [45]. However, these hybrid techniques, such as HOMER (Hand-centered Object Manipulation Extending Ray-casting) [12], still have limitations, including the difficulty of specifying the position and orientation of objects that are beyond arm's reach, i.e., in the distance.

The conventional method of interacting with distant objects in VR relies on raycasting [12]. This approach projects a line from the controller into 3D space to enable users to select and manipulate objects [7], [12]. While this method extends the user's reach, it typically requires hand-arm coordination: holding a controller with one hand to grab the object and moving/stretching the arm to position the object. If an object must be moved a considerable distance beyond what the arm's length can project, users may have to engage in repetitive pointing and dragging motions, especially in the absence of additional buttons that allow for "reeling" objects closer or farther. On the other hand, if such buttons are available, their presence, while helpful, could still result in finger fatigue due to frequent use. Consequently, all such repetitive actions can cause user fatigue and extend the time required for operations. With increasing distance, raycasting also amplifies hand rotations, affording less precision and embodiment at farther distances [60]. Yet, since this technique affords easy object selection [80], most current VR systems, such as Meta Quest, HTC Vive, and Windows Mixed Reality, offer raycasting with minor variations on the basic mechanism.

Oculus enhanced its built-in ray-casting pointing method by incorporating a non-linear mapping to reach and manipulate distant objects, mitigating the disadvantages of the go-go technique [12]. However, this non-linear mapping can make it challenging for users to perceive movement interval changes and thus move objects precisely. Sun and Stuerzlinger [70] developed and tested a new 3D positioning technique called extended sliding, which improved the accuracy and efficiency of raycasting in 3D positioning relative to the go-go technique. Although this technique leverages sliding objects on a surface and is most effective in a top-down view, it still inspired me to consider using a physical “surface” (the hand of the user) during interaction when manipulating a remote object. Both Liang et al. and Büschel et al. investigated interaction techniques that use the multi-touch screen of a smartphone as such a surface to perform gestures [15], [40]. Liang’s findings indicated users tended to avoid complex multi-touch gestures but also identified the potential to leverage surface input, and Büschel et al. showed that spatial input using a mobile phone is superior to mid-air gestures for 3D pan and zoom tasks.

In their survey of virtual 3D object manipulation techniques, Mendes et al. reviewed techniques for translating, rotating, and scaling objects, using different input devices such as a mouse, multi-touch surfaces, and mid-air gestures [45]. Although they mentioned that remapping gestures, using two hands, and using virtual widgets to constrain degrees of freedom could improve precision, they did not specifically explore the opportunity for two-handed interaction with touch-based input.

Recent research has underscored the potential for more intricate and efficient two-hand interactions in VR. Wagner et al. [76] presented a two-handed interaction method, which indirectly recognizes the user's non-dominant wrist position through the controller held in the corresponding hand, and the dominant finger through the controller held in the dominant hand. Similar to interacting with a smartwatch, this method affords interaction with the index finger of the dominant hand on the other wrist surface. This method could potentially be used to control the movements of distant 3D objects, but they didn't explore this opportunity, nor investigated bare-hand variants. According to Mine et al.'s [46] observations, body-relative interaction strategies that incorporate proprioceptive feedback work better than visual-only techniques. Body-relative interaction gives users a tangible real-world frame of reference, a more direct and precise sense of control, and "eyes off" engagement. Mine et al. also identified that a user can take advantage of proprioception during body-relative interactions. Pei et al. established the initial feasibility and presented a qualitative assessment of hand-centric interactions for AR/VR [55]. Using hand tracking and heuristics, they presented a large set of 28 hand interface designs to allow interaction with 3D objects within arm's reach without handheld controllers. Their results indicated that hand interfaces provide realism and tactile feedback, support proprioception, and can benefit precision. However, gesture sets may not generalize across different hand sizes, anatomical mobility constraints, and cultural contexts. The direct mapping approach also makes manipulating out-of-reach objects difficult. Learning such gesture sets will also significantly increase users' cognitive load. Leveraging the use of a physical surface and proprioception, I propose to extend this work by considering the entire palm and fingers of the (typically non-dominant) hand as (one or more) surface(s) to interact on, through simple gesture actions from the other (typically dominant) hand.

When considering classifications for 3D interaction proposed by various works [32], [38], [45], [53], the novelty of my approach becomes more apparent. Based on user needs, these classifications divide 3D interaction into four fundamental tasks: navigation, selection, manipulation, and system control. My focus is on manipulation tasks. For 3D object manipulation, Goh et al. reviewed touch-based, mid-air gesture-based, and device-based interaction techniques in handheld mobile AR [33]. They categorized touch-based techniques as those that allow manipulating virtual objects through multi-touch input on the device's screen. Mid-air gesture techniques utilize hand/finger tracking to enable interaction with virtual objects in mid-air, and device-based techniques use the mobile device itself as a controller, with its movement mapped to object manipulation. They also identified the utility of separating translation and rotation, which offers more intuitive (and precise) object manipulation. For translation, they found device-based and touch techniques to be faster and more accurate than mid-air gestures. For rotation, touch and mid-air gestures face challenges like occlusion, instability, and difficulty with large rotations. My solution combines the benefits of both touch-based and widget-based techniques; instead of using an additional device, I use the palm and fingers as the “device” and enable a touch experience on it.

Further, my technique's design draws inspiration from Shaw and Green's 1994 pioneering work, where they introduced THRED (Two-Handed Refining Editor), a novel Computer-Aided Design (CAD) system for creating freeform polygonal surfaces through two-handed interaction techniques [66]. Their system employs a pair of 3D position and orientation trackers, equipped with buttons for each hand, enabling users to construct and modify 3D surfaces consisting of rectangles that can be subdivided for greater detail. The left hand determines the geometric context and other settings, such as constraint modes and the level of selectable refinement, as well as manipulating the overall scene. Meanwhile, the right hand focuses on selecting control points, and the reshaping and refining of selected areas. This research highlighted the fluidity and intuitiveness of two-handed interactions, allowing users to effectively convey their design ideas and construct intricate 3D surfaces with ease. In a similar vein, my technique embraces this philosophy by manipulating virtual handles with the dominant hand for selection and manipulation, while the non-dominant hand specifies the mode and restricts operational capabilities, aiming to improve manipulation accuracy and reduce cognitive and physical demands.

Chapter 2.

Related Work

According to my literature search, the predominant focus of existing technologies for distant object manipulation in VR rests on leveraging techniques such as arm-extension, ray-casting virtual pointers, worlds in miniature (WIM), various widgets, or a blend of these strategies. Further, the implementations of these technologies frequently rely on the use of controllers, with bare-hand techniques receiving comparatively less attention. When free-hand methods are employed, they tend to concentrate on gestures and direct manipulation. Such approaches often fall short in supporting operations on distant objects beyond the arm's reach or are confined to rudimentary movements without the ability to precisely rotate objects. In this section, I provide a concise overview of the diverse range of existing 3D object manipulation techniques for immersive environments that have been presented over the span of nearly three decades.

2.1. Arm-extension

Poupyrev et al. [60] proposed the go-go interaction technique, which uses a non-linear mapping function to convert the user's head-to-hand distance into a controller distance between the real and virtual hands. Bowman and Hodges [12] presented the stretch go-go technique, which allows the virtual arm to be placed at any distance through a “reeling” metaphor. Instead of using arm motions, the user can then stretch or retract the virtual arm at a constant rate through two buttons on a 3D mouse/controller. Yet, this technique was less useful for precise control and/or controlling small objects, particularly when dealing with significant distances or situations where small physical movements result in substantial virtual hand movement. Poupyrev et al. also quantified the tradeoffs between the speed afforded by ray-based manipulation and the improved embodiment afforded by virtual hand techniques and recommended combining the benefits of both approaches [61].

Tseng et al. [74] introduced the FingerMapper technique, which mapped small-scale finger motions to full-scale virtual arm movements. This technique combines arm extension and finger gestures, allowing users to alter the virtual arm’s direction by moving their index finger or wrist and to control virtual arm movements by bending their index fingers. While the technique offers higher precision with reduced physical motion and fatigue, the authors' user study demonstrated that the use of FingerMapper led to an increase in task completion time, which they attributed to the learning curve. Furthermore, some of the mappings used in FingerMapper were found to place a higher mental demand on users compared to hand-centered and ray-casting techniques. Users may thus need additional time to adjust to the correlation between finger movements and virtual arm actions.

2.2. Ray-casting Virtual Pointer

Liang and Green [41] introduced an early "laser gun" ray-casting technique. Bowman and Hodges [12] extended the ray-casting method with a "fishing reel" metaphor and Balaa et al. [6] provided a comprehensive survey of various ray-casting techniques, evaluating their strengths and limitations. Users were able to manipulate distant items through this "laser pointer" metaphor. However, this method was restricted in that it could not handle movements of objects along the depth axis, which limited arbitrary 3D rotation and translation and also prevented the selection of obscured objects.

Feiner [20] added to the discussion by proposing a bendable pointer that could be used with either hand; the curvature of the pointer was determined by the relative positions of the hands. This method was able to resolve the occlusion issue but did not offer a method for manipulating distant objects.

Lee et al. [39] proposed an image-plane method with head and hand-directed ray-cast pointers, which flattened three-dimensional content into a two-dimensional plane. This approach allows for user interaction with faraway objects, but only in a flat, two-dimensional plane. Also, their head-directed method was found to be difficult to control precisely, especially for far-away objects.

In their spotlight (or flashlight) method, Argelaguet and Andujar expanded on the ray-casting method by substituting the ray with a cone [1]. Through this modification, long-distance selection errors were greatly reduced. Yet, the problem with this approach is that typically several objects are within the cone. To address this, Argelaguet and

Andujar proposed adaptable aperture-based selection cones [1]. Despite the fact that these methods extend the user's reach, they are only effective for translations perpendicular to the ray direction and rotations about the ray axis. It is simple to move an object in an arc, but translations in the ray direction or unrestricted rotations require the user to repeatedly grasp, manipulate, release, and re-grab the object, i.e., require clutching.

Bowman and Hodges [12] presented the HOMER technique, which combines ray-casting for object selection with hand-centered manipulation for object positioning. The interaction procedure begins when the user selects an object using ray-casting. Once selected, the virtual hand moves to the object's position and the object is attached to the virtual hand. The user can then directly manipulate the object using hand motions. When the user releases the object, the virtual hand returns to its natural position. This approach makes it possible to exert six-degree-of-freedom control over remote items. For small or faraway virtual objects, however, HOMER still fell short because users could not see how the object would be posed at a distant location. Additionally, HOMER demands greater hand-arm coordination and repetitive activation of such pointing, dragging and clicking actions could lead to significant arm fatigue.

2.3. WIM and Scaled World

Mine et al. [46] proposed a “head-butt zoom” technique, which allows users to switch between close-up and global views by simply leaning forward or backward. Bellarbi et al. presented a similar approach for manipulating distant objects in augmented reality (AR) through their “zoom-in” technique [9]. The purpose of this method was to make the user feel closer to both real and virtual objects by enlarging the image while maintaining the virtual objects' relative position. Yet, this method was unable to select objects within a crowded area, i.e., when a high level of precision was required. Also, when the “zoom-in” occurs, users might lose the broader context of the object's surroundings. This can make it challenging to understand the object's relation to other elements in the scene.

As presented by Pausch et al. [54], Stoakley et al. [69], Mine [47], Hand [29], and Wingrave et al. [78], the World in Miniature (WIM) technique shows the user a miniature 3D model of the VE in their hand. Even though this technique enables users to manipulate objects through their miniaturized representations in the WIM, making fine-grained, precise changes still remains challenging. If the whole environment is scaled down to WIM size, scene elements might even become too small to select, control, or even perceive. Although the user could interactively select a part of the environment to view in the WIM (such as a room instead of the complete house) to make manipulations at arbitrary resolutions easier, the drawback is that the WIM then needs to divide display space between the small copy and the actual environment.

Pierce and Pausch developed a compelling two-handed interaction strategy known as the “voodoo dolls” method by combining the WIM and image-plane techniques [57]. This method permits users to modify the size of virtual objects by manipulating a voodoo doll whose dimensions correspond to the desired size in the VE. However, this technique relies on two physical 6-DOF devices and cannot be adapted directly to bare-hand interaction.

Bacim et al. proposed several distinctive 3D selection strategies, including quad, discrete, and continuous zoom [5]. Despite gradually refining object-selection capabilities in the VE via these 3D interaction techniques, these techniques failed to support precise or concealed object selection.

2.4. 3D Manipulation Widgets

Nguyen et al. introduced a widget-based technique featuring four manipulation control handles for positioning and rotating objects [72]. This technique uses handle points for the precise positioning and orientation of virtual objects, enabling flexible manipulation across objects of varying sizes. Despite its advantages, the results of their study indicated inferior performance compared to direct manipulation techniques. Users found rotating objects by simultaneously moving two handle points to be confusing. While the barycenter handle facilitates approximate positioning and three handle points allow for precise positioning, users still face challenges in accurately placing objects. Nguyen et al. subsequently developed a widget with seven manipulation handles [50], offering enhanced control over the position and orientation of objects through interconnected handles. This advanced system facilitates more detailed and nuanced manipulation. However, even with that technique, direct manipulation proved more effective for smaller objects. Additionally, the increased complexity of the widget posed challenges, particularly for new users, due to the difficulty in managing seven distinct manipulation handles.

Zhang et al. introduced distant widget-based (DWBM) and near-field widget-based (NFWBM) techniques, which incorporate widget-based manipulation for multi-level degrees of freedom (DOF) separation and motion scaling [81]. DWBM emphasizes direct manipulation through ray casting and widget interaction, whereas NFWBM uses a scaled-down replica of the object for more intricate and precise manipulation, capitalizing on the user's natural hand movements within their immediate space. Nonetheless, users might need time to learn these interaction techniques, particularly to understand the relationship between widgets or replicas and the distant objects they manipulate.

Lee et al. introduced three novel near-field interaction metaphors for precise manipulation of distant objects in virtual reality: a widget-based metaphor, a unimanual metaphor with scaled replica (UMSR), and a bimanual metaphor with scaled replica (BMSR) [38]. These techniques utilize direct manipulation of a scaled replica within arm's reach to control the distant object, aiming to enhance precision and user experience. However, these approaches have several potential limitations. First, the scaled replicas may occlude the distant object, causing confusion or reduced situational awareness. Second, the involvement of the non-dominant hand in BMSR may lead to reduced precision compared to UMSR, as the non-dominant hand typically exhibits less dexterity and motor control than the dominant hand. Third, the scalability of UMSR and BMSR may be limited when dealing with extremely large or complex objects, as the scaled replicas may become too small or detailed to manipulate comfortably within arm's reach, potentially impacting user performance and satisfaction. Finally, these techniques still rely on handheld controllers for interaction, which may limit their applicability and naturalness compared to bare-handed tracking approaches. Incorporating advanced hand tracking technologies could potentially enhance the intuitiveness and immersion of these near-field interaction metaphors in future iterations.

Babu et al. further conducted a comprehensive comparison of techniques for manipulating distant objects in VEs, introducing two innovative approaches: Direct BMSR and Scaled HOMER + Near-field Scaled Replica View (NFSRV) [3]. These methods expand on the existing BMSR and Scaled HOMER techniques, respectively. Direct BMSR enhances the original BMSR approach by presenting near-field replicas of the target object within its context, facilitating more precise manipulation by providing a relative context. Scaled HOMER+NFSRV, on the other hand, builds upon Scaled HOMER by incorporating a near-field replica view of both the target object and its surrounding context, thereby enhancing depth perception and contextual awareness during distant object manipulation. The results of Babu et al.'s study underscore the balance between direct manipulation of distant objects and the utilization of near-field replicas, emphasizing the significance of contextual information in the manipulation process. Despite these advancements, the proposed techniques have potential drawbacks. The context radius size in both Direct BMSR and Scaled HOMER+NFSRV, dictated by the target object's dimensions, may not always yield the most effective or relevant contextual frame. For instance, a small target object might result in an overly limited context radius, excluding pertinent objects from the scene. Conversely, a large target object could lead to disproportionately small context replicas, complicating effective manipulation. Moreover, the techniques do not consider potential collisions or occlusions between the scaled replica and the environment or terrain, potentially causing unexpected outcomes or confusion during object manipulation in complex scenarios.

Overall, while innovative, these widget-based techniques share a common drawback: the presence of multiple widgets and handles clutters the scene, complicating the selection of desired controls. Furthermore, manipulation widgets can obstruct parts of the object being manipulated, hindering visibility.

2.5. Hand-Centered Direct Manipulation

Bettio et al. [10] introduced a markerless hand tracking and gesture recognition technique for VEs, enabling direct 3D interaction with models within the display space. This method supports the translation, rotation, and scaling of 3D models using simple hand gestures. However, the system's effectiveness in manipulating objects at various distances or sizes was not explored. In a related development, Pietroszek and Lee [58] proposed a method for selecting and moving objects using a virtual hand metaphor, allowing users to select and move desired objects upon intersection. Complementing these approaches, LaViola Jr et al. [37] employed direct mapping of the user's real hand's positional and orientational tracking onto a virtual hand model in the VE. These methods were straightforward to learn and intuitive to use, but they could only be used to pick and manipulate nearby items.

Bellarbi et al. [9] innovated by merging their above-mentioned Zoom-In technique with direct virtual hand manipulation to enable the manipulation of distant virtual objects in AR. This hybrid approach modifies the user's perceived distance to a virtual object by zooming the camera view, thereby bringing distant objects within arm's reach to facilitate easier selection and manipulation, without necessitating physical movement of the object or the user's viewpoint. However, this technique hinges on precise calibration and alignment between virtual and real-world content in AR to ensure accurate interaction. Its reliance on camera zoom might also curtail its applicability in scenarios where zooming could obscure important context or details. Moreover, in the context of bare-hand tracking in VE, it necessitates a specific gesture to activate the zoom-in or zoom-out feature, adding a layer of complexity to its implementation.

Yao et al. [80] developed a technique that uses an adjustable virtual pointer for direct manipulation, enabling users to select and control objects at various distances in virtual reality. The technique introduces two methods for adjusting the pointer's orientation: either through direct manipulation using a second controller or by interacting with a designated virtual object called an adjustment node. By allowing users to customize the pointer's orientation, this approach enhances the capability to select distant objects that might otherwise be difficult to interact with using a fixed pointer. However, the technique's reliance on a dual-controller setup may require users to invest time in learning and mastering the system. Specifically, users need to become adept at effectively adjusting the pointer's orientation and utilizing it for various tasks, which could involve a learning curve. Moreover, achieving precise control over the pointer's orientation might be challenging, particularly for tasks demanding fine-grained accuracy, as the small size of the pointer (or potential occlusions) makes it challenging to achieve the best accuracy.

2.6. Other Approaches

To enable precise 3D object manipulation in a VE, both Hayatpur et al. [30] and Gloumeau et al. [24] added constraints to the alignment and manipulation of 3D objects. Hayatpur et al. introduced the concept of “tethering” objects to constraints. When an object is tethered to a constraint, it can be manipulated only in ways specific to the type of constraint it is tethered to. For instance, an object tethered to a plane can only move parallel to the plane or along the plane's normal. An object tethered to a ray can move along the ray, rotate around the ray, or move closer or further away from the ray. An object tethered to a point can move closer or further away from the point or rotate around the point. They also discuss the use of hand gestures for specifying constraint-based manipulation. For example, pointing parallel to a ray will translate an object along the ray, while pointing “around” the ray will lock it to rotation. They also introduce the concept of “constraint snapping” to enable constraints to be placed faster and more accurately.

Gloumeau et al. [24] argued that existing techniques for object manipulation in VR are either inaccurate or slow, so they streamlined the process of adding such constraints and introduced an intuitive method, PinNPivot, which uses pins to constrain various rotations. Their technique also supports 3 DOF translation. Although their methods improve the precision of manipulation for 3D objects, all of them are based on direct manipulation, which means these 3D objects have to be placed within arm's reach. The authors did not delve into the challenges of incorporating constraints for remote 3D entities or fully explore the potential benefits of employing gestural interaction without any intermediary devices.

Many researchers have attempted to utilize a smartphone as a controller to control distant objects in AR/VR [15], [25], [33], [40]. All these methods require an additional device and a spatial mapping between the touch input and the object manipulation. Kari et al. presented HandyCast, a technique for controlling two virtual hands in VR using a single smartphone that enables full-range 3D input and bimanual interaction, even in a constrained physical space [33]. They used a pose-and-touch transfer function to map the smartphone's 6D pose (position plus orientation) and 2D touch locations to the 3D positions of the left and right virtual hands. Although their results show HandyCast enables effective distant 3D object manipulation, their method is still an indirect control method that reduces embodiment because it requires mapping smartphone orientation and touch to control a cursor or virtual object. Also, when wearing a head-mounted display (HMD), the phone screen is not visible during interaction, effectively requiring blind interaction. This makes touch-based input more challenging. Overall, while using a smartphone for VR input is interesting, challenges around tracking, input expressiveness, ergonomics, and haptics remain.

Kim and Lee's integration of a pressure-sensitive, ball-shaped controller with the direct HOMER technique represents a novel approach to distant object manipulation in VEs [36]. This compact, handheld device enables users to select, translate, rotate, and scale objects, offering a multifunctional solution. However, the reliance on a physical prop can restrict natural hand movements and diminish the user's sense of immersion within the virtual space. Furthermore, while their study contributes valuable insights into the comparative performance of the ball-shaped controller versus traditional VR controller techniques, it falls short on thoroughly examining the accuracy and precision of the controller in specific tasks of distant object manipulation. The emphasis of their research on performance comparison means that detailed analysis of its effectiveness for intricate object manipulation tasks remains an area for further investigation.

2.7. Summary of Previous Work

Despite the variety of existing approaches, it is difficult to identify a single input option for picking and manipulating distant objects in 3D that could be used in all kinds of interaction scenarios. I also recognize that freely and accurately manipulating remote objects is a much larger challenge compared to doing so for closer ones, as they cannot be manipulated as simply as an object within arm’s range, e.g., by adding axial constraints [24]. For translating objects, the majority of previous experiments also evaluated only axial movements and did not consider diagonal ones.

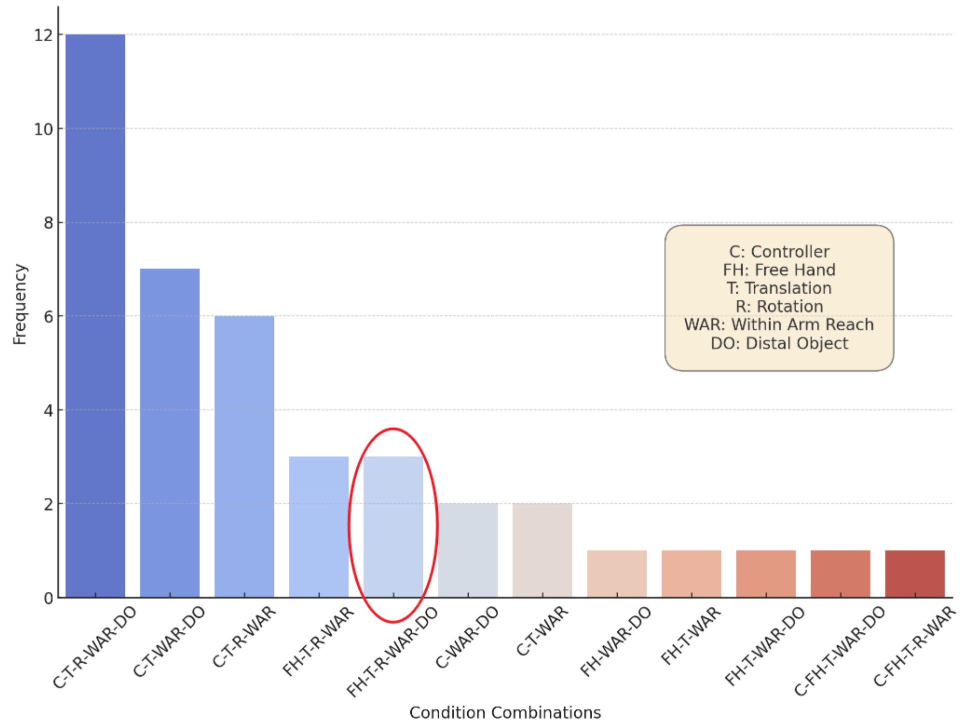

Figure 1. Heatmap visualization of compiled list of references across various conditions.

In addition, in various previous studies, most manipulations were based on controllers or virtual hands and did not directly employ real-time bare-hand tracking for the manipulation of remote objects. I surveyed existing all relevant 3D object manipulation techniques for VE [5], [6], [7], [9], [10], [11], [12], [15], [16], [17], [19], [20], [22], [23], [24], [31], [33], [37], [39], [40], [46], [48], [50], [53], [54], [55], [57], [58], [59], [60], [61], [68], [69], [70], [72], [74], [80], [81] and categorized them by Technique Type (Controller, Free Hand), Control Capability (Translation, Rotation), and Control Distance (Within Arm's Reach, Distal). Then I used a heatmap (Figure 1) to visualize this data, revealing a well-covered foundation in controller-based methods but a notable gap in more naturalistic interactions, particularly those combining Free Hand techniques, Rotation capability, and engagement with Distal Objects. This gap, highlighted by lighter shades in the heatmap, points to significant innovation potential in these less explored areas.

The work presented here addresses these challenges, taking advantage of raycasting for object selection, and then focuses on a novel technique to manipulate remote 3D objects. Some existing techniques allow object manipulation with simultaneous movement and rotation. To reduce the complexity of the technique and the potential interference caused by parallel operation of the two modes, and inspired by Goh et al.'s findings [33], my technique separates rotation and movement into two distinct modes.

Chapter 3.

Choice of Baseline Technique

As mentioned above, upon reviewing the existing literature, I identified a distinct gap in the application of bare-handed tracking for remote object control. Existing studies predominantly focus on direct manipulation for adjusting objects within arm's reach, with remote manipulation often relying on the user moving closer to the object. This approach, however, poses challenges when the object and user are not on the same horizontal plane, particularly if the object is elevated, limiting the feasibility of direct interaction. Alternative strategies, such as employing a magnetic-like attraction to draw objects closer for rotation, are constrained to rotational adjustments and require users to maintain awareness of the context of an object's original spatial positioning and spatial context.

An evaluation of remote manipulation techniques by Pierce et al. compared the Voodoo Dolls method and the HOMER method [57]. Despite their focus on controller-based interactions, their insights suggest that Voodoo Dolls could offer superior precision for tasks requiring fine control, whereas HOMER might be more suited to simpler tasks. However, the precision in a scaled-down environment, as seen with Voodoo Dolls or other world-scaled techniques, may not translate to the accuracy achievable in a full-scale context. Tseng et al.'s approach integrates finger movements, using the bending of fingers to symbolize the arm's extension, similar to the HOMER technique but without using a controller [74]. Also, Oculus Integration SDK's application of HOMER for bare-handed tracking through ray-casting and hand-centred manipulation further underscores the appropriateness of HOMER as a benchmark [51]. Consequently, my research adopts a bare-hand version of HOMER as a comparative benchmark, recognizing its relevance and adaptability for evaluating new methodologies for distal object manipulation.

3.1. HOMER Technique with Bare-hand Tracking

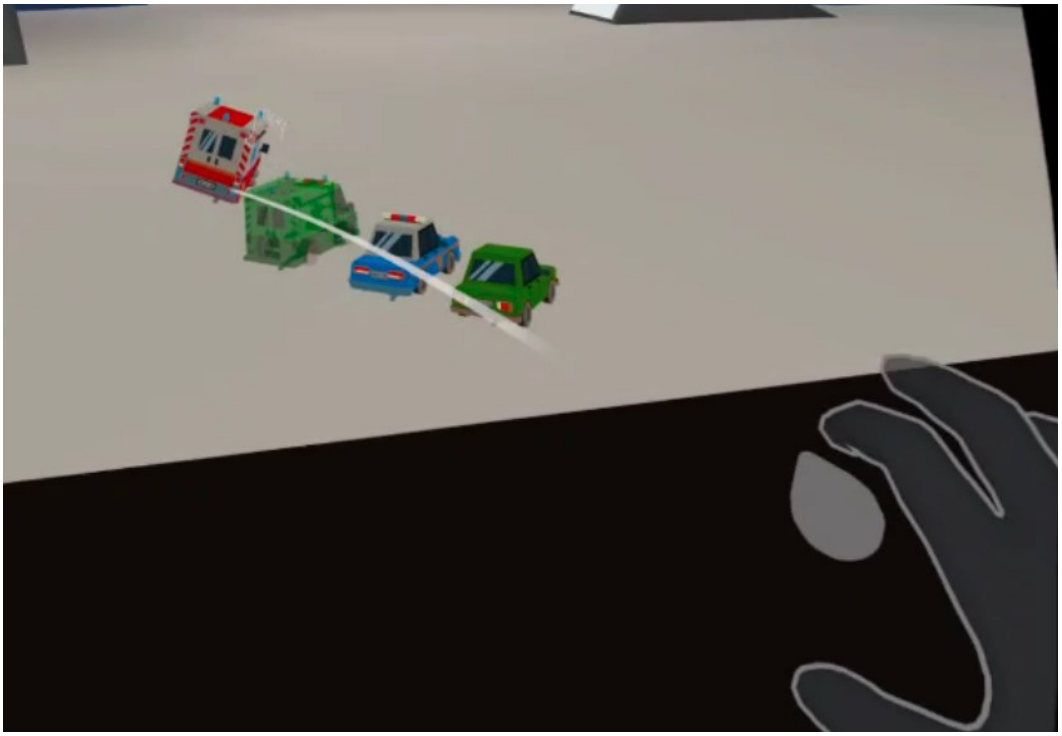

Figure 2. Illustration of bare-hand HOMER technique.

The HOMER technique as implemented by the Oculus Integration SDK employs ray-casting for object grabbing, wherein the user projects a slightly curved ray from the palm towards the scene. Upon selection, rather than attaching the object directly to the ray, a “ghost hand” is moved to the object's position to enhance the precision of interaction. Once grabbed (through a pinch with index finger and thumb), the object becomes tethered to the ghost hand, facilitating direct manipulation. This setup allows the user to mimic the movements of their own hand on the selected object, akin to the arm-extension technique. The technique supports simultaneous movement and rotation through a combination of using the ray-cast curve and wrist orientation, enabling participants to drag objects using an extended, curved ray while concurrently rotating them by twisting their wrists. For object movement, users employ the traditional drag-and-drop method, controlling the ghost hand's displacement of the object through coordinated stretching and contraction of the arm (and hand) (Figure 2).

Chapter 4.

Two-Hand Fingertip-Palm Technique

In this section, I present the features of my new Two-Hand Fingertip-Palm (THFP) technique and the iterations I went through during its design process. To achieve accuracy in finger and palm interactions, I explored various solutions, including using the hand as a touchpad for sliding movements and adding handles on the palm for dragging. Additionally, I experimented with different controller positions, such as placing the handle in the palm center, at the fingertips, on the knuckles, at the back of the hand, and at the finger joints. Through several small-scale internal usability tests, I iteratively improved my technique. During this journey to finalize my technique, I also conducted a pilot test, where I received mostly positive feedback and suggestions for improvement from users, and then correspondingly refined my technique for the final evaluation. Below, I first detail the version of the THFP technique as it was used in the pilot study and then present the features of my final THFP technique.

4.1. Pilot Study Version of THFP

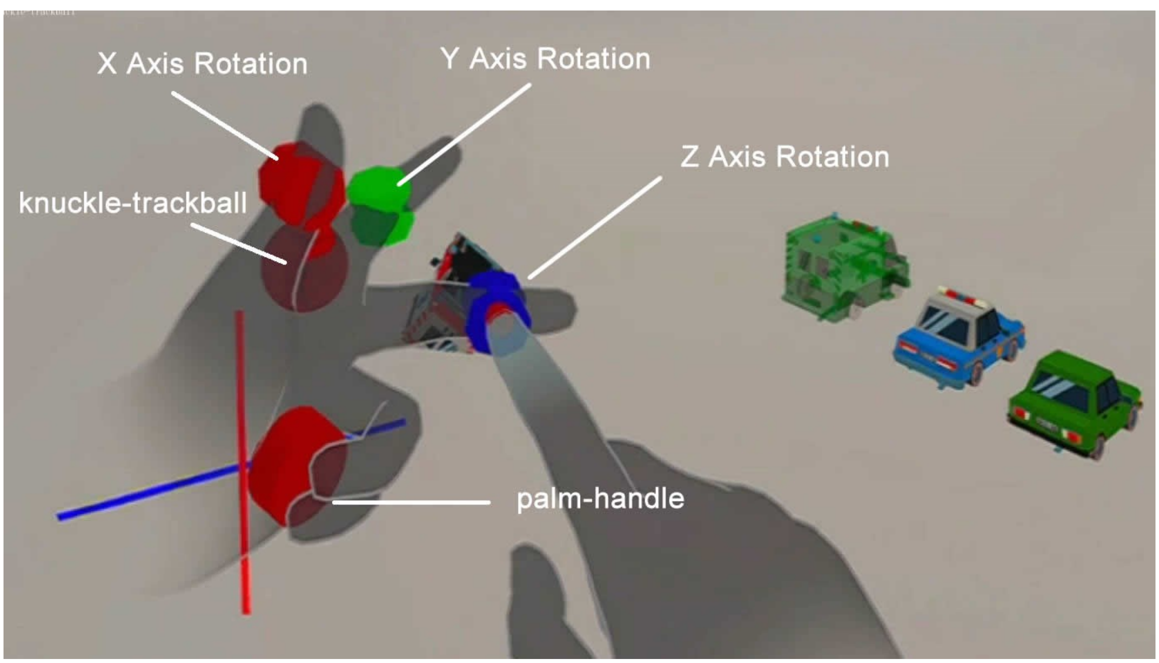

Figure 3. Illustration of my pilot-study version of the THFP technique.

Here, I detail the version of the Two-Hand Fingertip-Palm (THFP) interaction technique used in the pilot test for object movement and rotation. For movement, a virtual sliding handle (palm-handle) on the non-dominant hand's (NDH’s) palm is manipulated with the dominant hand's (DH’s) index fingertip, enabling object movement along the plane of the current palm coordinate system. My THFP technique adopts a virtual mouse wheel metaphor by sliding the DH’s index fingertip along the NDH's fingers. Movement on different fingers controls which rotation axis is affected, while the sliding distance adjusts the rotation speed. The dragging distance is proportional to the object's movement and rotation speed, enabling precise control over the object's displacement. Specifically, movements on the thumb, index, and middle finger control rotation around the X, Y, and Z axes, respectively. Moreover, I incorporate a knuckle-trackball located at the outside of the base knuckle joint of the NDH's index finger for free rotation of distant objects, manipulated by sliding the DH's index fingertip on it.

Chapter 5.

Pilot Study

In this chapter, I provide a detailed description of my pilot study's experimental procedure and discuss the results. This pilot experiment laid a foundation for my final experiment and inspired me to improve my technique. Through this pilot, I was able to identify key areas for enhancement and refine my interaction technique based on the feedback and observations gathered. This ensured that my final technique was both user-friendly and efficient, making it a valuable contribution to the field.

5.1. Research Questions

I expected to evaluate my method’s efficiency in time and accuracy for completing experimental tasks in comparison to the HOMER technique with bare-handed integration for manipulating a distant 3D object outside arm’s reach. I posed the following research questions:

RQ1: Does my new technique's performance in terms of efficiency match or exceed that of the HOMER interaction method for manipulation of distal objects?

RQ2: Does my new technique perform similarly to or better than the HOMER interaction technique in terms of accuracy when moving a distant object?

RQ3: Does my new technique perform as well or better than HOMER in terms of precision when rotating a distant object?

RQ4: Does my new method have a high degree of usability?

5.2. Participants

A diverse group of twelve volunteers from the local university community participated in this study, comprising three males and nine females. Their age distribution was five individuals aged between 18 to 24, four between 25 to 34, and three between 35 to 44. Their engagement with video games varied, with four playing daily, four a few times a week, one about once a week, and three less than once a month. Familiarity with 3D software and VR was also explored: two had never used 3D software before, while on the higher end, two engaged with it daily. Regarding VR, one had never experienced it before, while one used it a few times a week. The majority of participants (eleven) were right-handed, with only one left-handed individual. All participants were briefed prior to the experiment and provided informed consent.

5.3. Apparatus

As introduced above, I used the pilot version of my THFP technique. I compared it to the native HOMER implementation available on a Meta Quest 2 headset, using the Oculus Integration SDK for hand-tracking. My Unity application enables participants to complete the experimental tasks, records task completion time wirelessly in a database, and logs all object movement and rotation data locally frame-by-frame.

5.4. Procedure

After providing consent, participants were given a brief pre-assessment questionnaire encompassing demographic and previous experience questions. They then underwent a succinct tutorial to acquaint them with both methods for manipulating 3D objects in VR. Subsequently, participants engaged in two training phases, one for n each technique to ensure that participants could move and rotate a distant 3D cube successfully.

Thereafter, using one of the two interaction methods in a counterbalanced design, participants undertook six distinct tasks with the tasks presented in randomized order, followed by the six tasks with the other interaction technique (again in randomized order). The system automatically logged all position and rotation changes of the manipulated object to the target object. Once participants aligned the object closely with the target, they indicated task completion by pressing a red button situated on the desk before them, which then presented the next task.

Throughout all tasks, the researcher documented any instances of participants' confusion, inquiries, or remarks. Upon completion of all twelve tasks, participants filled out a System Usability Scale (SUS) survey to evaluate the technique. Lastly, the researcher conducted a semi-structured interview to garner insights into the participants' experiences.

5.5. Experimental Design

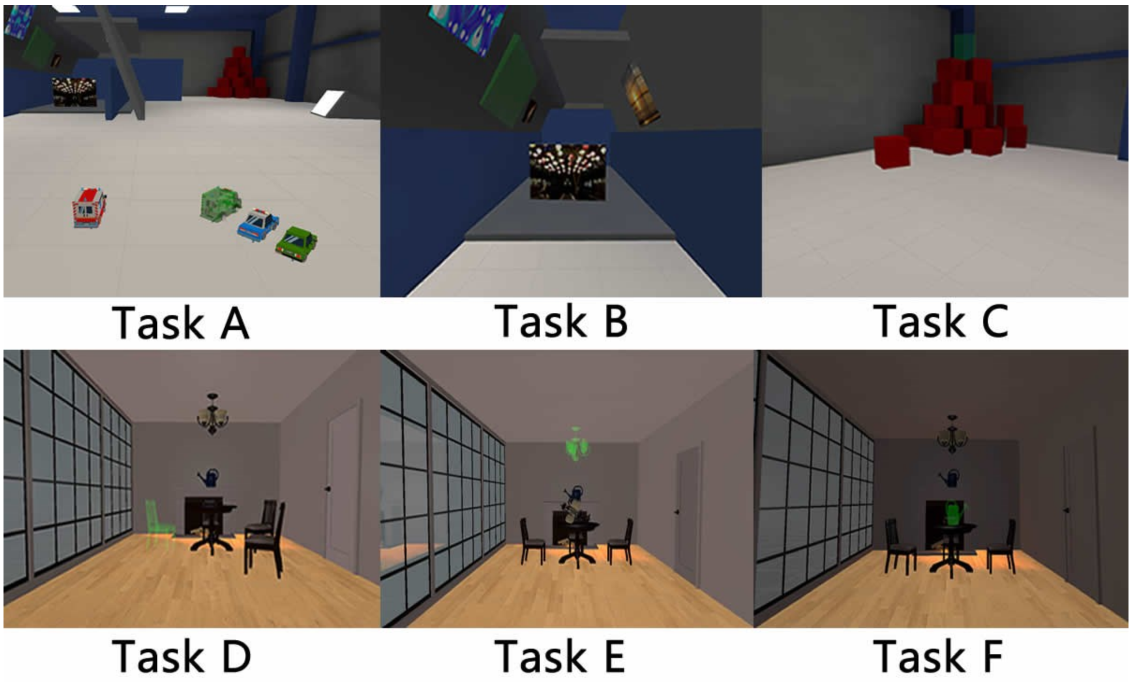

Figure 4. Screenshots of the six tasks used in my pilot experiment.

To thoroughly evaluate the efficiency and accuracy of my interaction technique, I devised six diverse tasks involving various challenges in object manipulation, such as moving and rotating distal objects outside of arm's reach and accurately placing them at specified locations (Figure 4). Collectively, these six tasks aim to gauge the effectiveness and versatility of the technique in handling various object manipulation challenges in a VR environment. Following Penumudi et al.'s recommendation [56], the target locations were between eye height or up to 15° below it.

To investigate research questions 1 through 3 (RQ1 - RQ3), I structured my study as a within-subjects experiment with two independent variables—interaction technique and task type. The two tested interaction methods were the HOMER interaction method and my THFP technique. For addressing RQ4, I employed the System Usability Scale (SUS) survey. My dependent variables comprised task completion time, the positional variance—termed “position difference” and the rotational variance—termed “rotation difference”—between the manipulated object and the target.

5.5.1. Task A

In this task, participants need to move a distant toy car to a designated parking spot on the right, aligning it parallel to another parked car. This spot is visually rotated by 45 degrees, testing rotations around the vertical axis and diagonal movements. The initial and target distances from the participants are both 4 meters.

5.5.2. Task B

Participants are tasked with hanging a large painting (2m x 1.5m) onto the inside wall of a slanted roof to their upper left, with the wall being tilted 30°. This task gauges the technique's ability to handle larger objects' movement and rotation. The initial distance to the painting is 5 meters, with the target location being 6 meters away.

5.5.3. Task C

Participants need to move a box onto a stack of boxes straight ahead, oriented at a specified angle. This assesses the technique's ability to push distant objects upwards and rotate them. The box starts at a 5-meter distance, with the target at 9 meters.

5.5.4. Task D

Participants are to move a chair from the right to the left side of a distant table, rotating it to face the table. This task evaluates the lateral and forward movement and rotation of distant objects. The chair and target points are both 2 meters away from the participants.

5.5.5. Task E

Participants must hang a chandelier from a distant table onto a specified ceiling spot, testing the technique's upward and forward movement as well as the rotational capabilities. The chandelier starts 3 meters away, with the target at 4 meters.

5.5.6. Task F

Participants are required to move a kettle from a distant fireplace to a table in front, rotating it to a specified orientation. This task tests the technique's ability to pull objects closer and rotate them. The kettle starts 5 meters away, with the target at 3 meters.

5.6. Data Analysis

I used JMP 15 to conduct quantitative analysis of the collected data. Shapiro-Wilk tests indicated a non-normal distribution for task completion time, position difference, and rotation difference across various conditions ($p < .0001$).

In quantitative research that utilizes common statistical tests, such as ANOVA (Analysis of Variance), verifying the preconditions for ANOVA, such as the normality of data, is crucial for the validity of the statistical inferences made. For “mild” deviations from normality due to skew, applying a post-hoc transformation is an acknowledged method to maintain the integrity of ANOVA results [28], [71]. Consequently, I applied a log transformation prior to performing Repeated Measures (RM) ANOVA in such cases. If it was not possible to achieve normality post-log-transformation, I employed the Aligned Rank Transform (ART) [79] on the original data before conducting RM ANOVA tests.

An initial review of the logs revealed that upon nearing the target point, participants invested more time fine-tuning either rotation or position, inadvertently increasing pose error during this stage. This prompted me to examine how to enhance my technique's efficiency and accuracy by minimizing these factors. An analysis of my experiment logs uncovered a trend where, upon reaching the point of lowest position and rotation difference, a reverse effect emerged, indicating instances where users' attempts at adjustments led to clearly less optimal positioning (Figure 5). I defined such pose errors as instances where the position value and the rotation value rebounded beyond a specified range (0.5 meters or 5 degrees). Utilizing these definitions, I also filtered the data to analyze situations where the system captured the first optimal result.

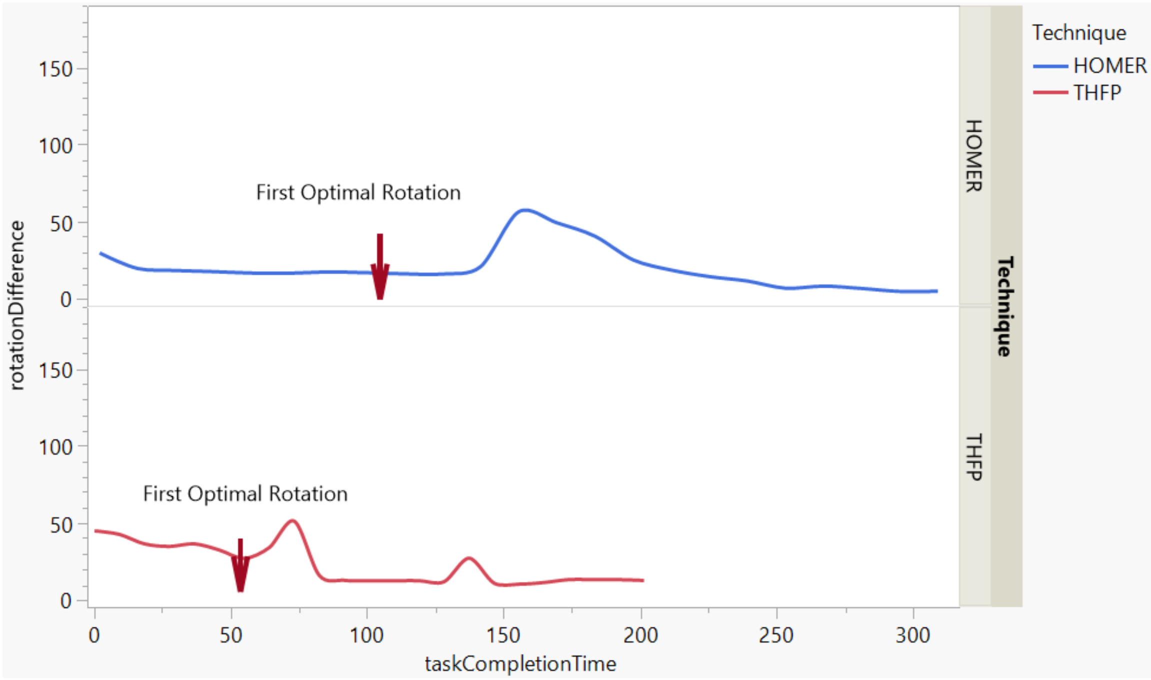

Figure 5. **Example of first optimal rotation of THFP and HOMER interaction technique in task A.**

I further separately analyzed rotation and position, tracking actions from initiation to the point of first optimal positioning prior to the occurrence of a pose error. The results were segmented into two scenarios: “User-triggered task completion” (default) and “System-monitored first optimal completion” (filtered) to facilitate a more nuanced analysis.

During testing, numerous comments regarding my system were recorded. Many participants expressed that they experienced a heightened sense of control, leading to an inclination to spend additional time on the technique to attain precise results. Conversely, some participants found it challenging to ascertain proximity to the target goal due to the object being distant (and thus visual pose differences appearing very small), leading to increased pose errors after reaching an optimal pose. I also noted a variety of strategies, with participants alternating between initiating translation and rotation in different sequences.

5.7. User-Triggered Task Completion Analysis

5.7.1. Task Completion Time

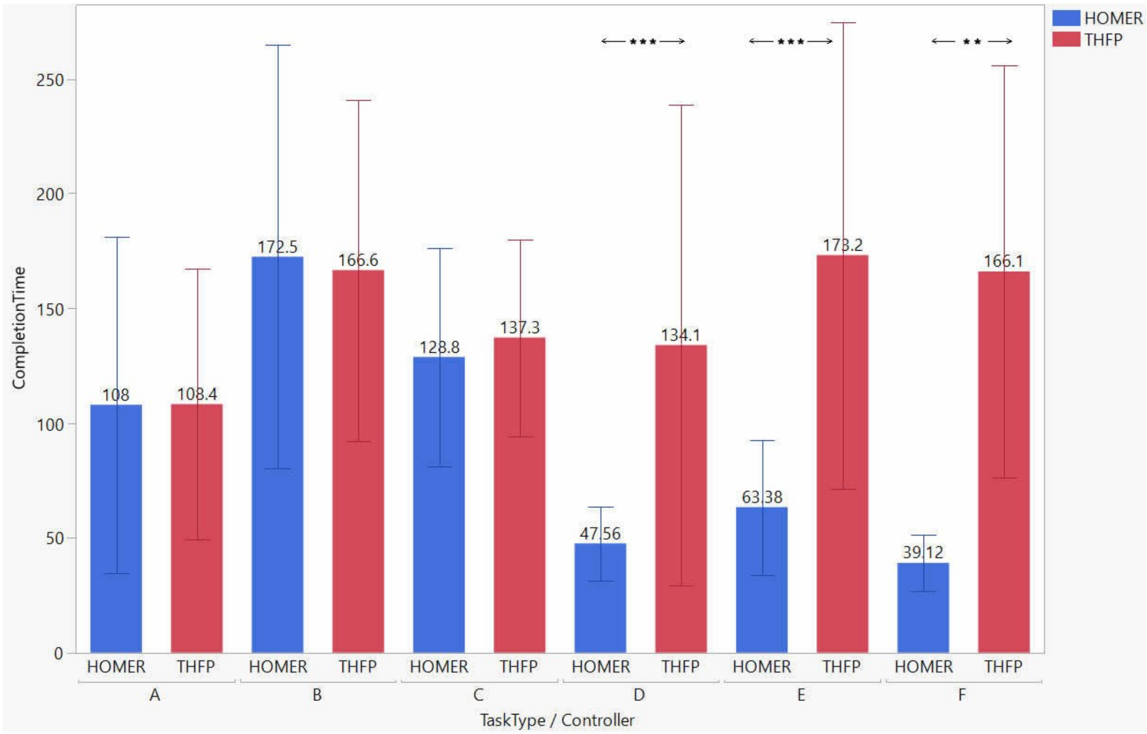

Figure 6. Average task completion time with HOMER vs. my THFP interaction technique for the six tasks.

Note: ***, **, and * in the tables and graphs indicate *p* < .001, *p* < .01, and *p* < .05, respectively.

Levene's test confirmed the homogeneity of variances for interaction techniques, F(1, 142) = 1.87, p = .28, and task types, F(5, 138) = 1.53, p = .18. RM ANOVA revealed significant differences between interaction techniques, F(1, 11) = 83.12, p < .0001, ω^2 = 0.86, and task types, F(5, 55) = 11.83, p < .0001, ω^2 = 0.47. Overall, my technique exhibited a longer mean completion time (M = 147.60, SD = 82.27) compared to the HOMER (M = 93.23, SD = 70.75). The interaction between task type and interaction technique was significant, F(5, 55) = 11.02, p < .0001, ω^2 = 0.45. Post hoc Tukey HSD tests demonstrated that while there were no significant differences between interaction techniques for tasks A, B, and C, my technique was significantly slower on tasks D, E, and F with p-values < .0001, < .0001, and = .0003, respectively (Figure 6).

5.7.2. Rotation Difference Analysis

My technique exhibited a lower mean rotation difference (M = 12.95, SD = 22.29) in comparison to HOMER (M = 16.49, SD = 34.91). Despite this, RM ANOVA with ART-transformed data revealed no significant variance between the methods, $F(1, 11) = 1.4882$, $p = .25$.

5.7.3. Position Difference Analysis

A similar trend was observed in position difference, with my technique showing a lower mean (M = 0.13, SD = 0.12) relative to HOMER (M = 0.16, SD = 0.47). Post-log transformation, a normal distribution was achieved with a single outlier in task A. Levene's test indicated homogeneity of variances, $F(1, 141) = 1.28$, $p = .26$, but RM ANOVA disclosed no significant difference, $F(1, 10.71) = 1.99$, $p = .18$.

5.8. System-Monitored First Optimal Completion Analysis

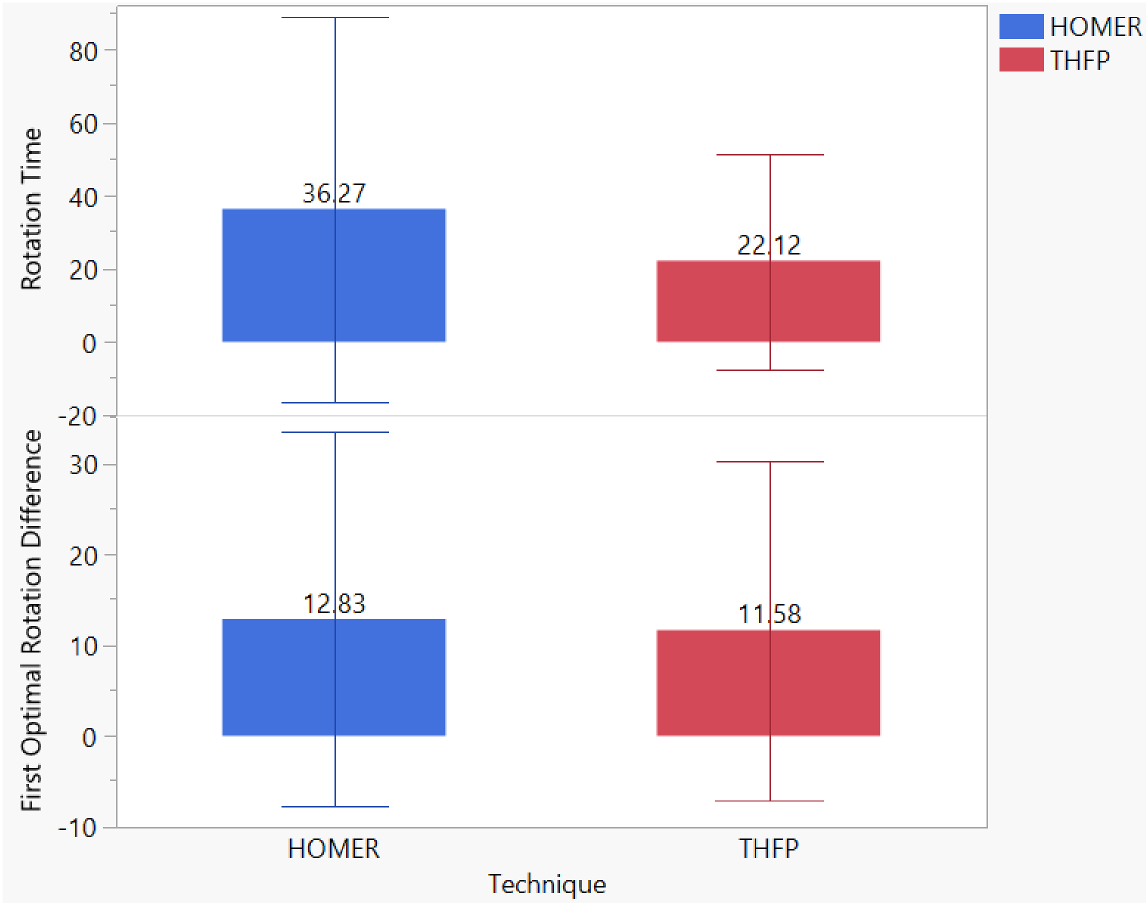

Figure 7. Mean rotation time and first optimal rotation difference vs. interaction technique.

While the user-triggered scenario revealed my technique being slower in certain tasks, a shift to system-monitored first optimal completion portrays a different outcome. The average task completion time for rotation across the six tasks was lower with my technique (M = 22.12, SD = 29.43) compared to the HOMER option (M = 36.27, SD = 52.46). The accuracy of rotation was marginally better in my technique, yet the accuracy of position was superior in HOMER (Figure 7). Given the variance in outcomes across the six tasks, I undertook a task-wise analysis.

5.8.1. Task A Analysis

Following log transformation, a Shapiro-Wilk test affirmed data normality for task A data. In comparison to HOMER with a mean (M) of 3.56s and a standard deviation

(SD) of 5.10, my technique exhibited a faster attainment of the first optimal rotation (M = 2.12, SD = 1.65). Levene's test revealed a violation of the homogeneity of variances for both rotation time and rotation difference. As a result, I transitioned to the Welch ANOVA test to analyze the data, given its tolerance for unequal standard deviations. That ANOVA showed no significant efficiency difference between the two methods, $F(1, 14.19) = 0.20$, $p = .66$. However, when considering the rotation difference relative to the initial optimal target rotation, my technique(M = 9.80, SD = 11.80) demonstrated a lower value compared to HOMER (M = 21.81, SD = 6.58). The Welch ANOVA results highlighted a significant difference in accuracy, $F(1, 11.77) = 13.36$, $p = .0034$, $\omega^2 = 0.47$. Despite this, my technique required more time to reach the first optimal position, with a significant technique impact indicated by RM ANOVA, $F(1, 11) = 95.63$, $p < .0001$, $\omega^2 = 0.87$, albeit without a significant difference in accuracy, $F(1, 14.60) = 0.16$, $p = .69$.

5.8.2. Task B Analysis

For task B, and post-log transformation, a normal data distribution was confirmed via a Shapiro-Wilk test. Levene's test revealed a violation of variance homogeneity for rotation time, $F(1, 22) = 7.88$, $p = .010$, prompting a shift to Welch ANOVA, which detected no significant distinction between techniques (HOMER: M = 8.60, SD = 4.92; THFP: M = 6.29, SD = 3.51), $F(1, 14.91) = 1.20$, $p = .29$. However, a significant difference favoring my technique was observed for achieving the first optimal rotation, $F(1, 11.67) = 8.40$, $p = .014$, $\omega^2 = 0.36$. For position, variance homogeneity was not violated, and no significant difference between methods was found in either completion time nor accuracy to the first optimal position.

5.8.3. Task C Analysis

In Task C, the symmetric nature of the cube necessitated a modification in measuring rotation differences, which I adjusted to use the remainder of the angle modulo 90 degrees prior to log transformation. Shapiro-Wilk and Levene's tests indicated a normal distribution and homogeneity of variances, respectively. Welch ANOVA demonstrated no significant difference in rotation metrics between methods (HOMER: M =29.67, SD = 19.60; THFP: M = 24.84, SD = 25.18), $F(1, 16.71) = 2.00$ $p = .174$. Similar results were obtained for position metrics with RM ANOVA showing no significant differences in completion time (HOMER: M = 54.00, SD = 17.65; THFP: M =

53.58, SD = 12.09), $F(1, 11) = 0.04$ $p = .84$, and position differences (HOMER: M = 0.83, SD = 0.77; THFP: M = 0.61, SD = 0.35), $F(1, 11) = 0.65$ $p = .44$.

5.8.4. Task D Analysis

In Task D, normal data distribution was confirmed post-log transformation through a Shapiro-Wilk test. Despite a breach in the homogeneity of variances for the first optimal rotation difference, $F(1, 22) = 10.42$, $p = .0039$, both RM and Welch ANOVA indicated no significant difference between the techniques in competition time (HOMER: M =10.60, SD = 6.48; THFP: M = 14.67, SD = 7.78) and the first optimal rotation difference (HOMER: M =7.44, SD = 4.49; THFP: M = 12.05, SD = 13.62), respectively. However, a notable difference was observed in the accuracy of reaching the first best target position, $F(1, 11) = 31.14$, $p = .0002$, $\omega^2 = 0.70$, with the HOMER technique(M = 0.06, SD = 0.05) outperforming mine (M = 0.28, SD = 0.21).

5.8.5. Task E Analysis

In Task E, following a similar verification of normal data distribution, I found that THFP enabled participants to reach the first optimal rotation significantly quicker, $F(1, 11) = 8.10$, $p = .016$, $\omega^2 = .17$, (THFP: M = 21.42, SD = 15.58; HOMER: M = 43.70, SD = 30.39) although no significant difference was detected in accuracy for both rotation and position. The techniques also displayed no significant variation in the time taken to achieve the first optimal position (THFP: M =35.39, SD = 29.36; HOMER: M = 28.67, SD = 18.85) and the accuracy therein (THFP: M = 0.10, SD = 0.09; HOMER: M =0.09, SD = 0.11).

Task A	First Optimal Position	Mean	SD	First Optimal Rotation	Mean	SD
Completion Time	HOMER	1.34	0.63	HOMER	3.56	5.1
	THFP	7.13	4.37	THFP	2.12	1.65
Difference to Target	HOMER	0.37	0.11	HOMER	21.81	6.58
	THFP	0.49	0.35	THFP	9.8	11.8
Task B		Mean	SD		Mean	SD
Completion Time	HOMER	127	89.71	HOMER	110	89.39
	THFP	119.6	74.36	THFP	50.62	53.7
Difference to Target	HOMER	0.1	0.09	HOMER	7.94	14.8
	THFP	0.15	0.1	THFP	10.73	12.69
Task C		Mean	SD		Mean	SD
Completion Time	HOMER	54	17.65	HOMER	29.67	19.6
	THFP	53.58	12.09	THFP	24.84	25.18
Difference to Target	HOMER	0.83	0.77	HOMER	3.31	3.09
	THFP	0.61	0.35	THFP	3.22	2.9
Task D		Mean	SD		Mean	SD
Completion Time	HOMER	34.44	13.33	HOMER	10.6	6.48
	THFP	33.56	13.77	THFP	14.67	7.78
Difference to Target	HOMER	0.06	0.05	HOMER	7.44	4.49
	THFP	0.28	0.21	THFP	12.05	13.62
Task E		Mean	SD		Mean	SD
Completion Time	HOMER	28.67	18.85	HOMER	43.7	30.39
	THFP	35.39	29.36	THFP	21.42	15.58
Difference to Target	HOMER	0.09	0.11	HOMER	24.44	41.89
	THFP	0.1	0.09	THFP	18.4	36.94
Task F		Mean	SD		Mean	SD
Completion Time	HOMER	19.11	7.01	HOMER	21.56	12.47
	THFP	35.53	11.66	THFP	17.62	19.74
Difference to Target	HOMER	0.12	0.14	HOMER	12.01	17.08
	THFP	0.19	0.16	THFP	15.3	15.71

Figure 8. **Mean and SD of completion time and differences to target rotation and position with the HOMER interaction method vs. THFP interaction technique in the six tasks in the pilot study.**

Note: Significant differences are marked in red.

5.8.6. Task F Analysis

Task F analysis revealed a violation of the homogeneity of variances for rotation completion time, $F(1, 22) = 9.32$, $p = .0058$, yet both Welch and RM ANOVA demonstrated no significant difference between the interaction techniques in rotation metrics (Rotation completion time: HOMER: M =21.56, SD = 12.47; THFP: M = 17.62, SD = 19.74; Rotation difference: HOMER: M =12.01, SD = 17.08; THFP: M = 15.30, SD = 15.71). However, in position metrics, THFP was significantly slower in reaching the first optimal position, $F(1, 11) = 13.38$, $p = .0038$, $\omega^2 = .49$, (THFP: M = 35.53, SD = 11.66; HOMER: M =19.11, SD = 7.01) though accuracy remained comparable between the two techniques (THFP: M = 0.19, SD = 0.16; HOMER: M = 0.12, SD = 0.14).

5.8.7. Overall Analysis

The analysis of the data for Tasks A to F provides a comprehensive understanding of the comparative performance between THFP and the HOMER technique across various tasks (Figure 8). The key findings suggest that THFP generally facilitates quicker attainment of the first optimal rotation, albeit also exhibits some variance in the time required to achieve the first best position. This discrepancy may be attributed to the distinct interface dynamics or the learning curve associated with my technique, as discussed below.

5.9. SUS Analysis

Based on the SUS survey, the analyzed results indicate a positive reception of my technique's usability among participants. In the pilot study I did not assess the usability of HOMER. Out of the 12 participants, my technique achieved an average SUS score of 74.58 with a standard deviation of 9.28, which is notably above the benchmark of 68 [13] (Figure 9). This suggests that users generally found the technique to be user-friendly and intuitive.

Further individual investigation revealed that participants felt highly confident using THFP, with a mean score of 4.25. My technique's consistency and ease of use were also rated favorably, with mean scores of 4.08. Conversely, the perceived complexity and the need for technical support were areas identified for improvement, with mean scores of 2.25 and 2.33, respectively.

The results highlight THFP's strengths in terms of consistency, ease of use, and user confidence. However, it's crucial to take the standard deviation of 9.28 into account when looking at the usability score. This large variation indicates that while many users viewed the technique positively, some users possibly faced challenges or had concerns about the technique's usability, such as its complexity and the support provided to users. I also realized that a simple SUS survey does not provide insights into whether users experience less fatigue and physical demand when using THFP compared to HOMER. Therefore, I decided to perform further research to explore these issues.

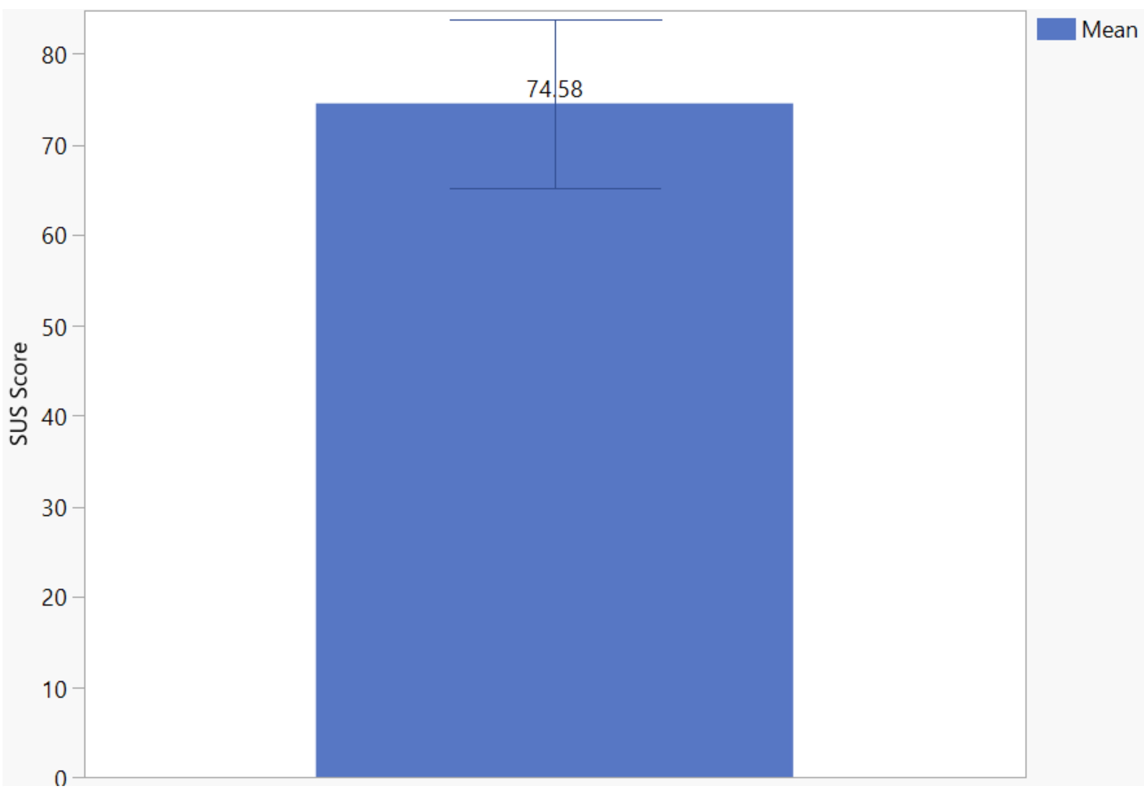

Figure 9. Mean of SUS test results of my technique for the pilot study.

5.10. Discussions and Limitations

Although, compared to HOMER, users took more time with my THFP technique to complete tasks, it demonstrated comparable performance in terms of accuracy for both positioning and rotation when participants had to decide on task completion. Notably, when I ignored errors caused by subsequent adjustments, THFP exhibited superior support for rotations, especially a faster attainment of the first optimal rotation during the execution of Task B and E, while still maintaining a comparable overall level of accuracy. Moreover, my technique showcased better precision during the rotation of distant objects in Task A. Also, participants reported a more relaxed experience as there was no need to constantly monitor their hands during interaction, given they are naturally aware of their hand positions through proprioception.

In summary, for complex scenarios requiring precise rotation, like tasks A and E, the pilot study version of my interaction technique likely surpasses the HOMER interaction method in terms of being the most efficient for achieving the initial optimal target rotation. Furthermore, the results suggest that my technique exhibits a high degree of usability. Consequently, my pilot user study findings partially validate my Research Questions (RQ) 1–4, even for participants who had limited familiarity with VR (9 out of 12). Still, the most significant issue identified was the pose error induced by subsequent adjustments after participants had already closely approached the target pose. Based on these outcomes, I identified some potentially confounding factors that affect my pilot study.

5.10.1. Hand Tracking Issues

During the pilot study, which utilized a first-generation Oculus Quest headset, my study identified critical tracking issues associated with the Oculus built-in vision system used for hand movement recognition. When fingers made contact with the palm surface, both displacement in the palm position and instances of frame drops were observed, indicating challenges related to occlusion. Furthermore, my intent to make the THFP technique ergonomic, by eliminating the need for users to stretch their arms, inadvertently moved user interactions into blind spots in tracking. Participants frequently positioned their left hand near their chest while interacting with their right hand, often moving it out of the headset camera's view and causing tracking loss. This highlights the pilot study technique's limitations in providing comprehensive tracking without blind spots across all positions. Additionally, the performance of my pilot version was compromised by limitations in the hand-tracking algorithms and the Heisenberg effect. Factors such as camera system resolution, inadequate lighting, occlusion, and reduced system efficiency (e.g., FPS drops) can negatively impact tracking accuracy. The Heisenberg effect, triggered by a sliding action over the palm with sufficient force [11], can further disturb tracking and result in unintended object movements. However, my overall technique design, which minimizes fingertip movement during object manipulation by allowing users to drag and stabilize a slide handle with their other hand, aims to mitigate these issues. I thus considered replacing the dragging and re-dragging actions with dragging and holding actions for the next iteration of my technique and to validate if this yielded an improvement.

5.10.2. Interaction Technique

In my pilot test, I observed that neither the knuckle-trackball nor the palm-handle achieved as precise object manipulation as I expected, with the primary challenge stemming from the fact that they were interacting with a virtual controller. Both the trackball and the handle have a certain thickness, and when the fingertip of the user passes through these virtual items as they are leaving the surface. Thus, even though it visually appears that the fingertip has left immediately, in reality, the fingertip travels briefly along the surface of these thick virtual items. During this short time, the object remains in a collision state, so if tracking experiences frame drops or jitter at this moment, it can result in the erroneous displacement of the object. In other words, although the object appears to have reached the designated position, the slight movement of the fingertip touching and immediately leaving the controller surface can still cause minor object movement. Improvements in this area await future innovations in the precision of hand-tracking technology. Meanwhile, reducing the thickness of the controller collider could serve as a mitigation measure. I also observed that users tend to move the palm-handle to its maximum speed in order to achieve rapid object movement. In subsequent improvements, I should thus consider how to better leverage this characteristic. For instance, the palm-handle could be used mainly to rapidly move an object to an approximate location, with further refinements relying on different interaction actions.

5.10.3. Manicure

A significant majority of participants in my experiment were female (9 out of 12). I received feedback from many participants regarding the difficulty in dragging the handle along the palm using the fingertip of the DH's index finger, as long fingernails posed an issue. This led to a gap between the fingertip and the palm, resulting in disengagement from the controller grasp. Initially, I aimed to achieve haptic feedback through the palm, yet I discovered that users with long or artificial nails did not enjoy this sliding and dragging method, as the nails scrape against the palm, which can cause discomfort. This observation prompted a deeper reflection on my part, emphasizing the need for a more inclusive approach during technique design. To enhance the technique's functionality, I thus considered shifting the contact point from the top of the fingertip to the dorsal side of the fingertip.

5.10.4. Undo Functionality

While the aforementioned approaches might reduce errors, I acknowledge that no system can perfectly evade all errors. Hence, I considered incorporating an undo function to assist users in swiftly reverting to a state prior to an error. I observed within my pilot study that users can efficiently attain the first optimal state concerning object rotation. Yet, rectifying errors post-occurrence necessitated a considerable amount of time. If I could diminish the time spent on error correction, that would improve my technique's efficiency. Nonetheless, the addition of an undo function would also increase the technique's complexity slightly, potentially even imposing additional cognitive load on the users. Therefore, further research is necessary to investigate this issue. In the main user study, I will incorporate an undo button into the technique to check whether this feature can mitigate this issue.

5.10.5. Constraint Integration (3D Spatial Awareness and Movement Constraints)

Participants also mentioned that sometimes they wished to move the object along a single axis at a time. However, the pilot version of my current technique did not support such constraints, only allowing users to freely move objects along the plane coordinate system of the non-dominant palm, which typically resulted in displacement along two axes at the same time. To move in a single direction, users would need to align the plane of their palm accurately with the axes of the world coordinate system, ensuring that the dragging direction either aligns with or is perpendicular to the palm orientation. Moreover, due to human anatomy, it is easier for my palms to face upwards, to the inside, or forward, while other poses can be uncomfortable.

Furthermore, since the hand's plane is two-dimensional while the object exists in a three-dimensional space, participants need to understand the mapping between the two-dimensional plane and the three-dimensional space while dragging the palm-handle. While some participants easily figured out how the hand orientation and drag direction work together, others found this very challenging and made many mistakes. I observed that participants unfamiliar with three-dimensional space often overlooked or became confused by the impact of the palm's three-dimensional orientation on the object's movement vector, which increased their cognitive load. Often, users wished for spatial movement to remain consistent, like for rotation, with constraints applied to movement along a single axis. In other words, users were not opposed to operations based on the palm plane but wished for restrictive operational forms similar to those with the finger-based interactions for rotations. Past research also indicates that enhancing degrees of freedom (DOF) separation can effectively improve operational precision [45]. Therefore, I speculated that adding a single-axis constraint for displacement to my technique and to distinguish between translation and rotation modes through a hand flip gesture.

5.10.6. Relative to User Distance

Auteri et al. [2] and Frees and Kessler [22], [23] suggested that traditional out-of-reach manipulation methods involve amplifying user movements to control distant objects, using a control/display (C/D) ratio based on the user's movement velocity. However, my experiments found that when using HOMER participants made significantly larger hand movements for more distant objects, amplifying inherent hand jitter and tracking errors. This phenomenon detrimentally impacts the efficiency and precision of control as it amplifies the inherent jitter of the hand and tracking. This observation may point to one of the factors why the HOMER method manifested superior performance when the target objects were in closer proximity to the user in my pilot study (as observed in tasks D, E, and F). Thus, the factor of object distance warrants further exploration to better understand its impact on user interaction and technique performance.To ensure that I am comparing HOMER and THFP fairly, I considered focusing only on objects that were initially more than 4 meters away from the user in my main user study.

5.10.7. Control Modes and Consistent Interaction

Concurrently, based on some of the user feedback, my pilot version exhibited some “consistency issues,” indicating user comprehension barriers. In the pilot version, to rotate objects, participants dragged a wheel or a trackball, but to move objects, they used a handle on the palm. This setup made the controls feel less coherent or unified.

Participants suggested that I could consider introducing mode switching; for instance, in movement mode, the three fingers could operate a specific single axis each, while the palm or knuckle position could be used for free movement. Similarly, in rotation mode, the three fingers could allow object rotation along a single axis, while the palm or knuckle position could be used for free rotation. These suggestions and hypotheses offer directions for further research. In chapter 5.12, I will provide a detailed description of the improvements that I made to my THFP technique.

5.11. Pilot Study Conclusion

From the preceding analysis, I can see the advantage of my technique in quickly reaching a position close to accuracy. However, due to the technique's operational complexity and lack of fault tolerance, users made many operational errors when making precise adjustments. Additionally, because I only implemented single-axis constraints for rotation and did not provide similar support for movement, it was evident that users encountered difficulties during precise displacement operations. This pilot test made me aware of the potential of my technique, but it also exposed some flaws. Based on user feedback from this test, I iterated on the THFP technique and conducted a main user study, as described in the following sections.

5.12. Revised Version of THFP for the Main User Study

The final version of my new THFP technique separates translation and rotation modes, while still allowing for swift transitions between them through a simple flip of the non-dominant hand. When the palm of the NDH faces the user, all controls on that hand surface are mapped to translation operation. Flipping the palm so the back of the NDH faces the user switches the mode of all operations to rotation.

In the THFP technique, while the 3D position of the NDH is continuously monitored, it is important to note that the movement of the NDH hand itself after contact does not influence the manipulation of the target object directly. Instead, the system focuses exclusively on the relative positioning of the DH's index fingertip to the NDH's surface. This interaction occurs through designated virtual controllers — 'handles' — on the NDH. When manipulating objects, the user's dominant hand interacts with these handles, which are mapped to translation and rotation control, based on which side of the NDH faces the user at the start of the interaction.

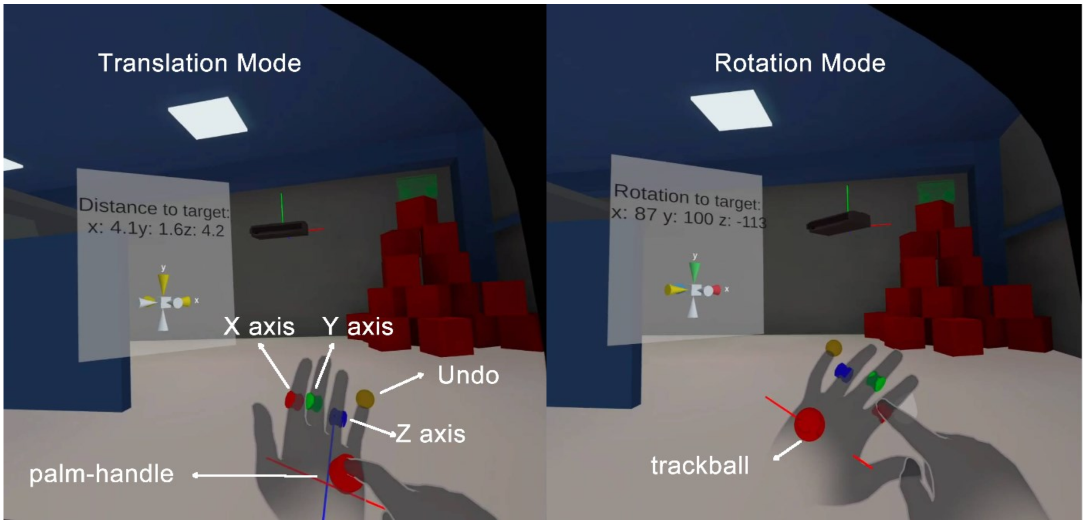

Figure 10. Illustration of the THFP technique used in the main user study.

In translation mode, there are primarily two options for interaction. The first is a palm-handle (virtual sliding handle) on the NDH's palm, which is manipulated with the fingertip of the index finger of the DH, enabling object movement along the plane of the current palm coordinate system. Its main purpose is to assist users in quickly moving objects to a proximate 3D target location. By changing the orientation of the palm plane, users can then freely move objects in 3D, even diagonally. For example, if the NDH's palm faces diagonally forward-left at a 45-degree angle, pushing forward on the palm-handle with the right index finger moves the object in a 45-degree left-upward direction. As mentioned above, translations of the NDH palm do not affect the interaction. The second option is designed for translating the object along the coordinate system axes and THFP offers three handles on the index, middle, and ring fingers of the non-dominant hand, corresponding to object movement along the x, y, and z axis respectively. These "wheel" handles, located at the fingers' intermediate joints, can be dragged forward or backward to adjust the object's position along a specific axis. The dragging distance proportionally controls the selected object's movement speed, with the maximum positive speed at the fingertip and the maximum negative speed at the proximal joint. If dragged beyond these limits, the wheel handle resets, and the object stops moving, requiring the user to reselect and drag the wheel handle for further movement. Since each of the three fingers constrains movement along a single axis, this method allows for precise translation adjustments.

For the rotation mode, which is mapped to the back of the NDH, my method offers a free rotation handle and three-wheel handles. For rotations, I incorporated a trackball on the back of the NDH for free rotation of distant objects, manipulated by sliding the dominant hand's index fingertip over it. This method is useful for quickly rotating the object to a rough orientation rather than precise adjustments. The three wheel handles on the backs of the fingers afford precise rotation around the x, y, and z axis, with rotation speed and direction controlled similarly to those in translation mode. This is loosely inspired by Zhao et al.'s research [82], which advocated the mouse wheel for user-friendly and precise 3D object rotation.

To make it easier to recover from manipulation errors, I added a yellow undo button at the tip of the little “pinky” finger of the NDH, which can be activated by touching it with the index finger of the dominant hand to swiftly revert the object to any of its previous states. Such functionality enhances the precision with which users can manipulate an object (Figure 10).

Additionally, my UI employs visual cues for axis identification: the colour of each axis displayed at the object matches the colour of the wheel handle associated with that axis. This feature allows users to easily identify how to correct any differences between the object's current position and the target at a distance, as well as to recognize which rotation axis might be needed to get the object into the correct target orientation.

Furthermore, my THFP technique incorporates a floating user interface (UI) anchored to the top of the user's DH, providing real-time feedback on the object translation or rotation, as appropriate for the active control mode. It displays any discrepancies in distance or rotational angles along the x, y, and z axis between the object and its intended target position.

As my THFP technique allows the DH's fingers to touch the palm and fingers of the NDH, it incorporates a tactile element. Mine et al. noted the challenge of precise virtual object manipulation due to the absence of physical work surfaces and haptic feedback in VEs [46]. Typically, users manipulate virtual objects with extended arms without support in most current VR systems. For precise manipulation in the real world, it is common to rest the forearm, elbow, or arm's palm on a support surface to stabilize hand motion and reduce mental and physical strain. My approach leverages this principle, allowing users to stabilize their fingers on the other hand, eliminating the need for visual tracking of the hand's position to understand its spatial location. My method not only enables precise control of the translation and rotation of remote objects in VR but does so by relying only on bare-handed interactions. Overall, my innovative method for manipulating remote objects in VR also tackles the precision challenges highlighted by Bowman et al [12] and provides the flexibility to move and rotate objects both freely and precisely.

Chapter 6.

Main User Study

6.1. Research Questions

Based on the issues identified from the pilot study, I modified the THFP technique. To evaluate the redesign, I performed another comparison of my method with HOMER in terms of time and accuracy for completing experimental tasks, while also investigating several new research questions:

RQ1: How does my Two-Hand Fingertip-Palm (THFP) technique perform in terms of time efficiency relative to the HOMER technique for the manipulation of distal objects?

RQ2: Does THFP method perform similarly to or better than the HOMER technique in terms of accuracy when moving a distant object?

RQ3: Does my THFP technique perform as well or better than the HOMER technique in terms of precision when rotating a distant object?

RQ4: Does my new THFP method have a high degree of usability?

RQ5: Does my new THFP technique invoke higher mental and physical demand than the HOMER technique?

These research questions are aimed at exploring the effectiveness, usability, and demand of my newly iterated THFP technique compared to the established HOMER technique. By addressing these questions in my main study, I aim to deepen my understanding of the THFP technique's capabilities and identify areas for further improvement.

6.2. Participants

In this study, distinct from my previous pilot, I recruited an entirely new and diverse group of twelve volunteers from the local university community. This group consisted of three males and nine females. The age distribution among the participants was as follows: three individuals were aged between 18 to 24 and nine were aged between 25 to 34. Their engagement with video games varied, with four playing daily, one a few times a week, four approximately once a week, and three less than once a month. Additionally, I assessed their familiarity with 3D software and VR: six participants had never used 3D software, whereas two engaged with it daily. Regarding VR experience, eight participants had no prior experience, while two had used VR a few times a week. All participants were right-handed. Each participant was briefed about the study's objectives and procedures and provided informed consent before the experiment commenced.

6.3. Apparatus

Similar to the pilot study, my THFP interaction technique was implemented within a Unity application. However, this time I utilized a more modern headset, specifically the Oculus Quest Pro. The Unity application leverages the Oculus Integration SDK for hand-tracking, which operates through the headset's built-in cameras. The whole application enables participants to complete tasks, records task completion time wirelessly in a database, and logs all object movement and rotation data locally at a rate of 10 times per second. For distant 3D object selection, I make use of the SDK's built-in curved raycast feature. Confirmation of selection is achieved by the participants pinching the index finger and thumb together. For distant 3D object manipulation, I compared HOMER and my newly developed THFP interaction technique.

6.4. Procedure and Experimental Design

Mirroring the approach of the pilot study, my within-subjects study explored the efficiency and accuracy of two interaction methods—HOMER and my THFP technique—across six diverse tasks (A-F) involving manipulating objects of varying sizes and shapes in a VE. These tasks were designed to test the interaction techniques' capability to manage complex movements and rotations across different axes and distances. Unlike the pilot test, I replaced the box in task C with a speaker to make it easier to identify the correct orientation. The tasks included (Figure 11):

- Moving a toy car to a diagonally rotated parking spot (Task A)
- Hanging a large tilted painting on a slanted wall (Task B)
- Placing a speaker atop a stack of boxes at a specific angle (Task C)
- Repositioning and rotating a chair from one side of a table to the other (Task D)
- Hanging a chandelier from a table to a ceiling spot (Task E)
- Moving and rotating a kettle from a fireplace to a table (Task F)

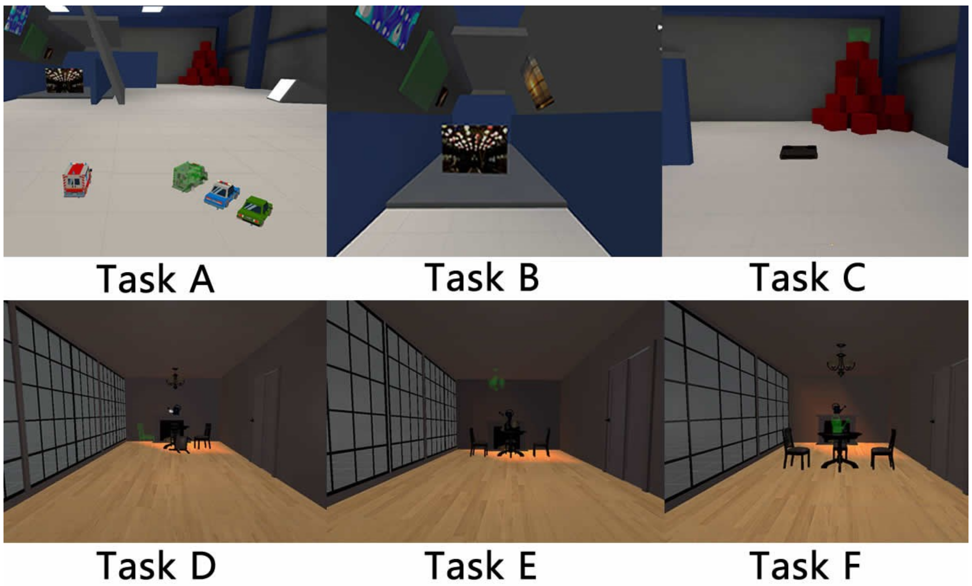

Figure 11. Screenshots of the six tasks used in my main user study.

I adjusted the objects' initial distances to range from 4 to 9 meters, with target distances spanning 4 to 16 meters, encompassing lateral, vertical, and diagonal movements, as well as rotations around different axes. The dependent variables were task completion time, positional variance ("position difference"), and rotational variance ("rotation difference") between the manipulated and target objects. These tasks collectively aimed to evaluate the effectiveness and versatility of the systems in handling various object manipulation challenges in VE.

After giving consent and completing a pre-assessment questionnaire, participants underwent a tutorial and two free training phases on each technique. They then performed the six tasks, presented in random order, using one interaction method, followed by the same tasks using the other method, in a counterbalanced manner. The system logged all position and rotation changes, and participants indicated task completion by pressing a red button. The researcher noted participants' inquiries, remarks, or confusion during the tasks.

Upon completion, participants, in addition to filling out the System Usability Scale (SUS), completed NASA Task Load Index (NASA-TLX) surveys [44] to evaluate both techniques, addressing research questions RQ4 and RQ5, respectively. Finally, I conducted a semi-structured interview to gather insights into participants' experiences. This comprehensive approach provided a more in-depth understanding of both the usability and the cognitive and physical demands of the methods.

6.5. Data Analysis

I used JMP 15 to conduct quantitative analysis of the collected data. Shapiro-Wilk tests indicated a non-normal distribution for task completion time, position difference, and rotation difference across various conditions ($p < .0001$).

During the user study, numerous comments regarding my technique were recorded. Many participants expressed that they experienced a heightened sense of control, leading to an inclination to spend additional time on the technique to attain precise results. Conversely, for the HOMER technique some participants found it challenging to ascertain proximity to the target goal due to the object being distant (and thus visual differences being small), leading to increasing pose errors after having already reached an optimal pose. I also noted a variance of strategies, with participants alternating between initiating translation or rotation in different orders.

6.6. Results

6.6.1. Task Completion Time Analysis

A Shapiro-Wilk test on task completion time identified that the data was distributed normally after log transform. Levene's test confirmed the homogeneity of variances for interaction techniques, $F(1, 142) = 3.45$, $p = .07$, and task types, $F(5, 138) = 0.90$, $p = .48$. RM ANOVA revealed significant differences between interaction techniques, $F(1, 11) = 25.97$, $p = .0003$, $\omega^2 = 0.66$, and task types, $F(5, 55) = 3.18$, p =

.0136, $\omega^2 = 0.15$. Overall, my technique exhibited a longer mean completion time (M = 121.51, SD = 38.12) compared to the HOMER technique (M = 83.92, SD = 44.92). While the interaction between task type and technique was not significantly different, overall my technique was approximately 44.79% slower than the HOMER technique (Figure 12).

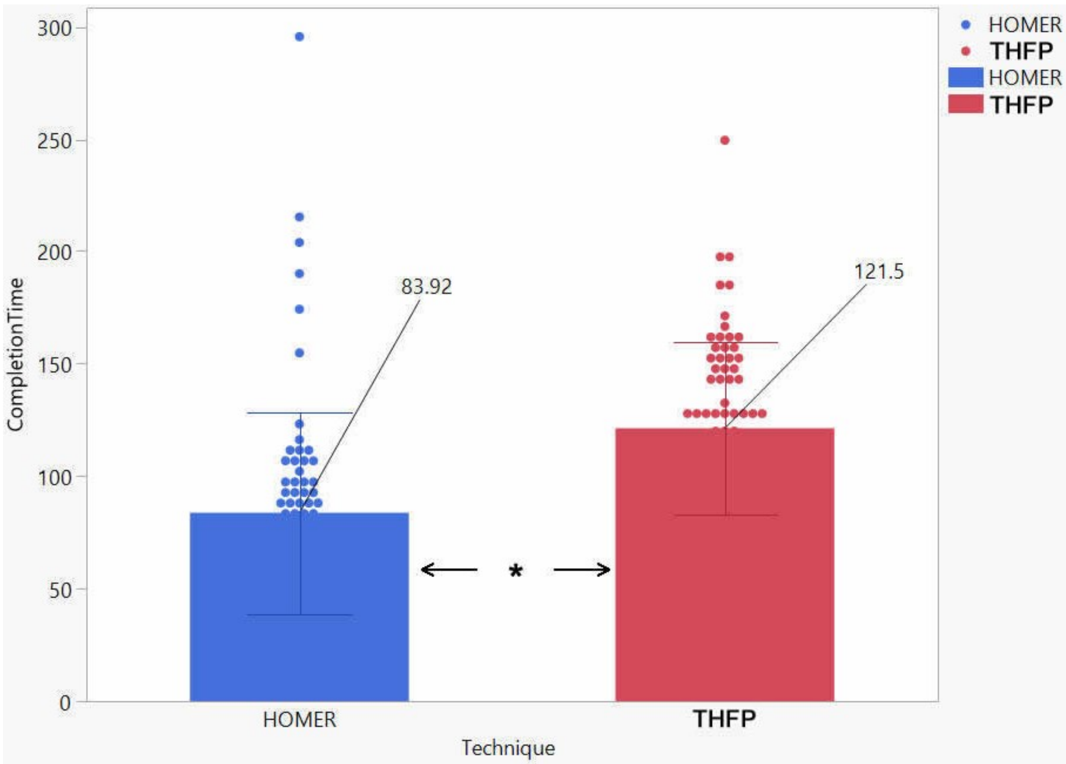

Figure 12. A scatter plot of task completion time and average task completion time with HOMER vs. THFP technique.

6.6.2. Position Difference Analysis

Following log transformation, a Shapiro-Wilk test affirmed data normality for position differences. Levene's test revealed a violation of the homogeneity of variances, $F(1, 142) = 19.21$, $p < .0001$. As a result, I transitioned to a Welch ANOVA to analyze the data, given its tolerance for unequal standard deviations. The Welch ANOVA identified that the THFP technique has significantly higher accuracy for position than the HOMER technique, $F(1, 108.1) = 33.37$, $p < .0001$, $\omega^2 = 0.23$. THFP exhibited a significantly lower mean position difference (M = 0.04, SD = 0.03) in comparison to the HOMER technique (M = 0.17, SD = 0.33) (Figure 13). Using the THFP technique, there was an improvement of approximately 76.47% in distance accuracy compared to HOMER.

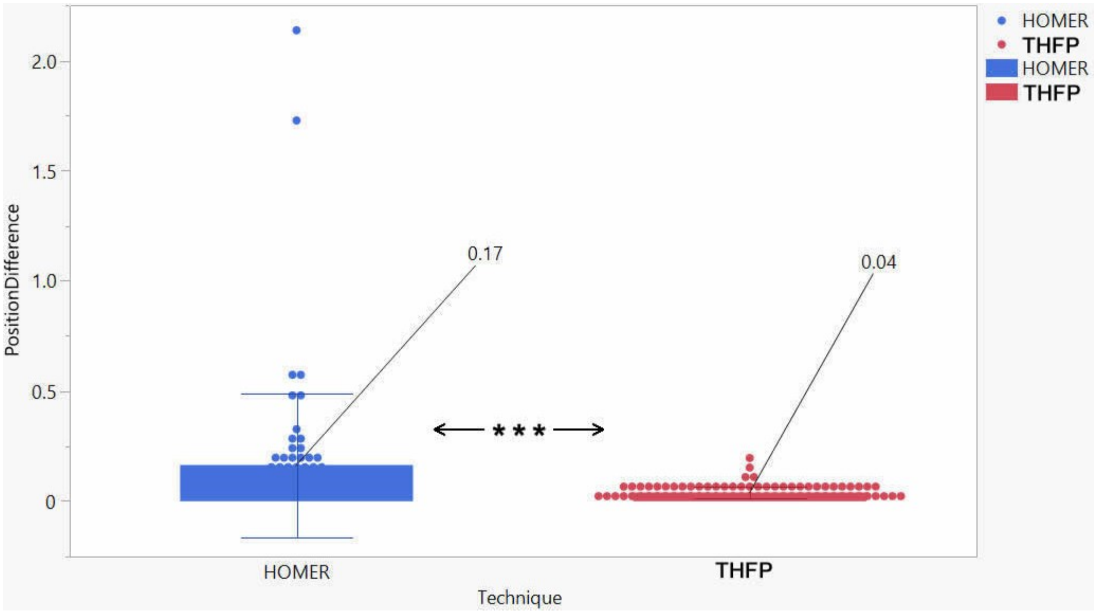

Figure 13. Average position difference with HOMER vs. my THFP technique.

6.6.3. Rotation Difference Analysis

A similar trend was observed for rotation differences, with THFP showing a lower mean (M = 1.77, SD = 1.12) relative to the HOMER technique (M = 11.53, SD = 14.36) (Figure 14). After log transformation, the data was normal but Levene's test revealed a violation of variance homogeneity, $F(1, 142) = 4.62$, $p = .03$, prompting a shift to Welch ANOVA, which detected a significant difference between techniques, $F(1, 133.06) = 137.34$, $p < .0001$, $\omega^2 = 0.50$. The result indicated my technique improved rotation accuracy by approximately 84.65% compared to the HOMER technique.

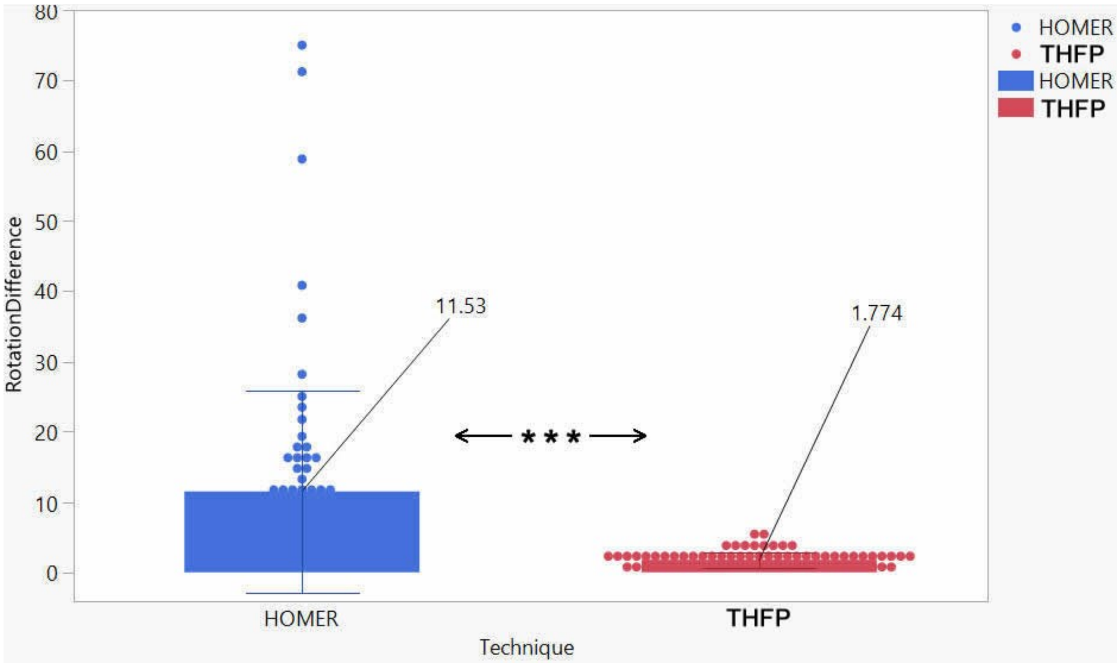

Figure 14. Average rotation difference with HOMER vs. my THFP technique.

6.6.4. Task A Analysis

Following log transformation, a Shapiro-Wilk test affirmed data normality and Levene's tests indicated homogeneity of variances for task A. My technique achieved a higher accuracy in terms of position and rotation differences (Position: M = 0.03, SD = 0.01 ; Rotation: M = 1.52, SD = 0.96, respectively) compared to the HOMER technique (Position: M = 0.07, SD = 0.06 ; Rotation: M = 10.28, SD = 5.06, respectively). The RM ANOVA identified a significant difference between the two techniques, in position, $F(1, 11) = 20.74$, $p = .0008$, $\omega^2 = 0.60$ and in rotation, $F(1, 11) = 47.54$, $p < .0001$, $\omega^2 = 0.78$. My technique required more time to complete task A (HOMER: M = 88.64, SD = 50.69; THFP: M = 100.37, SD = 42.96), albeit without a significant difference in completion time, $F(1, 11) = 0.66$, $p = .43$ (Figure 19).

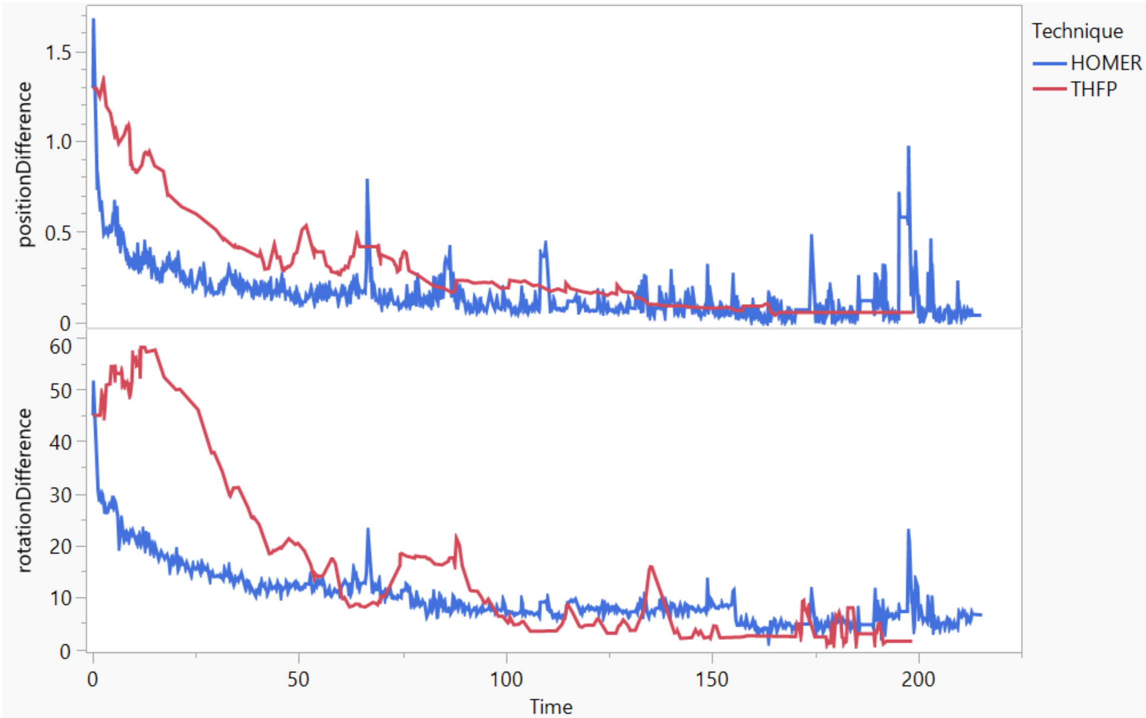

Figure 15. Mean (Position difference, rotation difference over twelve participants with two techniques in Task A) vs. Time.

Given that I recorded the position and rotation differences of objects relative to their target positions at 0.1-second intervals, I visualized the mean performance of all participants in Task A across all time points in a linear fashion (Figure 15). I discovered that, despite users spending approximately the same amount of time using both the HOMER technique and my technique, over time the position and rotation differences tended to converge to a lower value with my technique. Roughly from the midpoint of the average task duration onwards, the average position and rotation differences attained with HOMER were generally higher compared with my technique. This means that, as tasks neared completion, the accuracy of my technique consistently maintained a relatively better level compared to HOMER.

6.6.5. Task B Analysis

Post-log transformation, a normal data distribution was confirmed via a Shapiro-Wilk test for the data for task B. Variance homogeneity was also validated via a Levene's test. My technique required more time (M = 116.28, SD = 32.69) to complete the task than HOMER technique (M = 84.02 , SD = 45.82), with a significant impact of interaction technique indicated by RM ANOVA, $F(1, 11) = 5.27$, $p = .04$, $\omega^2 = 0.25$. However, a significant difference favoring my technique was observed for accuracy in both translation (HOMER: M = 0.13, SD = 0.14 ; THFP: M = 0.04, SD = 0.02), $F(1, 11) = 7.08$, $p = .022$, $\omega^2 = 0.32$ and rotation (HOMER: M = 3.97, SD = 1.86; THFP: M = 1.49, SD = 0.67), $F(1, 11) = 33.64$, $p < .0001$, $\omega^2 = 0.72$ (Figure 19).

I also created a similar visualization as for Task A, and identified a similar pattern, specifically in the time period approaching the completion of the task, where the curves representing the rotation and displacement differences appeared more jagged for the HOMER technique compared to my technique. This suggests that users were repeatedly making minor adjustments to their position and rotation but were unable to achieve a relatively precise location and orientation (Figure 16).

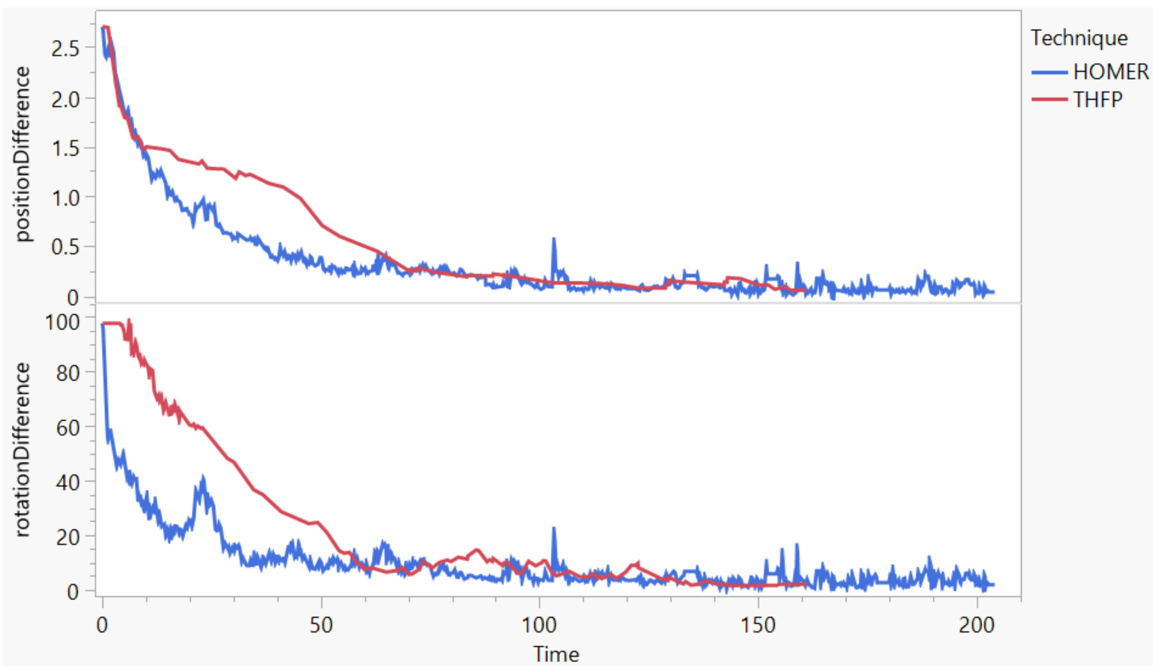

Figure 16. Mean (Position difference, rotation difference over twelve participants with two techniques in Task B) vs. Time.

6.6.6. Task C Analysis

In Task C, following a similar verification of normal data distribution and variance homogeneity, RM ANOVA reflected that my technique (M = 137.97, SD = 33.92) was also slower than the HOMER technique (M = 98.95, SD = 34.34) with significant variation, $F(1, 11) = 7.43$, $p = .02$, $\omega^2 = 0.33$, but maintained a significantly higher accuracy in translation (HOMER: M = 0.48, SD = 0.70; THFP: M = 0.04, SD = 0.01), $F(1, 11) = 13.81$, $p = .0034$, $\omega^2 = 0.50$, and rotation (HOMER: M = 11.56, SD = 10.16; THFP: M = 2.15, SD = 1.65), $F(1, 11) = 16.18$, $p = .0020$, $\omega^2 = 0.54$ (Figure 19).

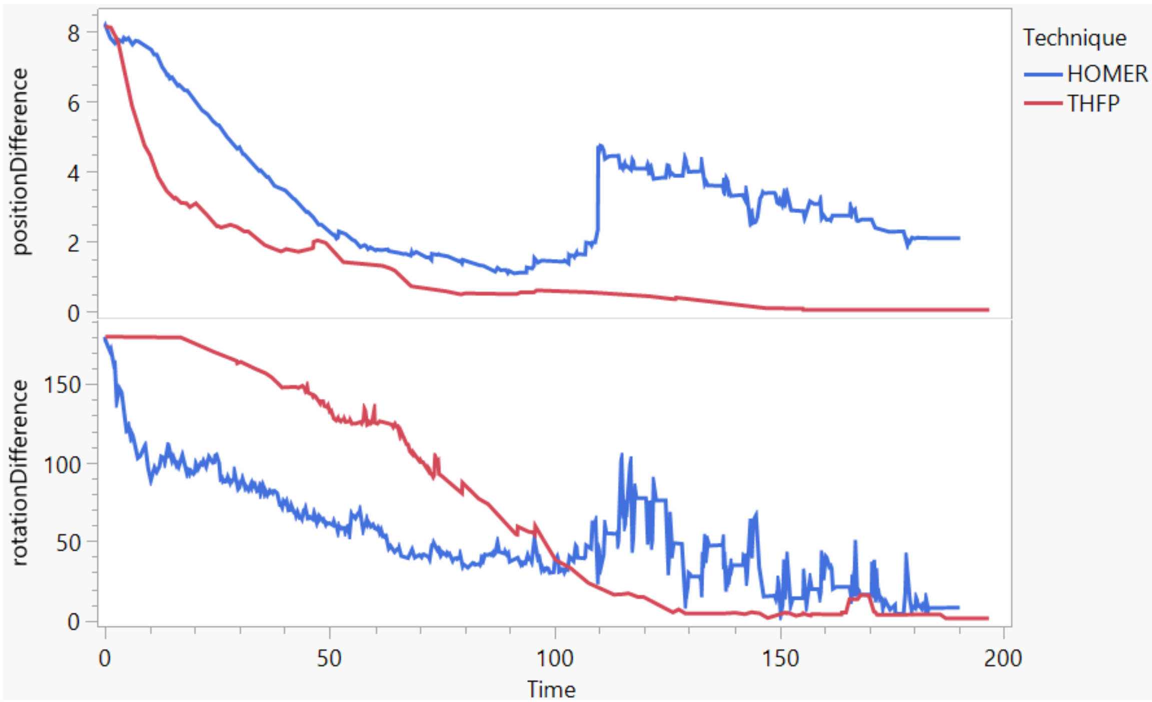

Figure 17. **Mean (Position difference, rotation difference over twelve participants with two techniques in Task C) vs. Time.**

Using again a visualization of performance over time, I confirmed the same pattern as in tasks A and B. Yet, for task C the discrepancy between my technique and HOMER is much more noticeable in the terms of position differences. The significant improvement in position accuracy for my technique is evident by the larger disparity in the two curves (Figure 17). Additionally, as indicated by the red line for rotations, my technique demonstrates greater stability, whereas the blue curve for HOMER exhibits continuous fluctuations. This suggests that participants attempted to adjust rotation to a precise state, but these attempts were not consistently successful. Ultimately, in terms of precision, the HOMER technique falls short of my technique's performance.

6.6.7. Task D Analysis

In Task D, the Shapiro-Wilk test verified normal distribution for post-log transformation data, confirming its suitability for further statistical analysis. However, I observed a significant departure from homogeneity of variance in position difference, $F(1, 22) = 5.84$, $p = .024$, in contrast to the satisfactory variance homogeneity in time to completion and rotation difference. RM ANOVA revealed that my technique had significant slower completion times (M = 135.69, SD = 45.30) compared to the HOMER technique (M = 80.38 , SD = 69.40), consistent with a pattern seen in previous tasks, $F(1, 11) = 11$, $p = .0069$, $\omega^2 = 0.43$. Despite the slower times, my technique (M = 1.97, SD = 1.06) exhibited a significant improvement in rotation accuracy over the HOMER technique (M = 8.78, SD = 6.91) according to RM ANOVA results, $F(1, 11) = 32.95$, $p < .0001$, $\omega^2 = 0.71$. Although Welch ANOVA revealed no statistically significant difference in position difference, my system still achieved lower average position differences (M = 0.05, SD = 0.03) than the HOMER technique (M = 0.14, SD = 0.16), albeit insignificantly so (Figure 19).

6.6.8. Task E Analysis

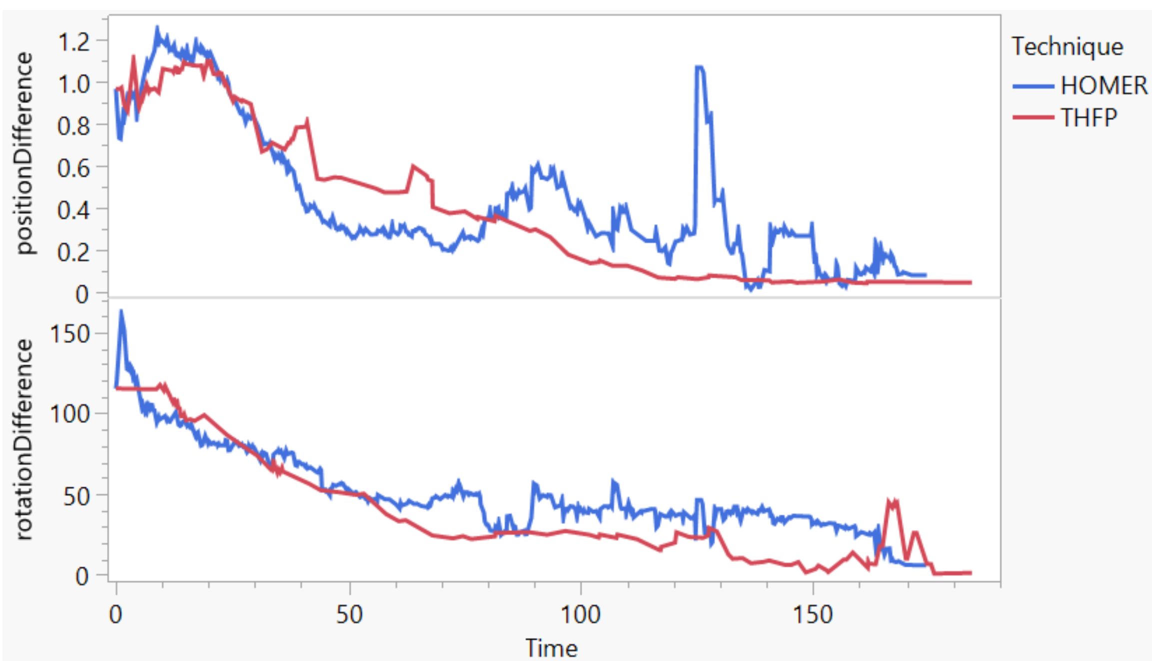

Figure 18. Mean (Position difference, rotation difference over twelve participants with two techniques in Task E) vs. Time.

In Task E, after data transformation, Shapiro-Wilk and Levene's tests confirmed a normal distribution and variance homogeneity, respectively. RM ANOVA aligned with previous results, illustrating that my technique (M = 131.78, SD = 22.26) required significantly more time compared to HOMER (M = 83.10, SD = 33.05), $F(1, 11) = 22.95$, p = .0021, $\omega^2 = 0.53$. Simultaneously, my method (Position: M = 0.04, SD = 0.02; Rotation: M = 1.84, SD = 1.10) showed higher accuracy in both position (with no significant variance) and rotation (with a significant difference, $F(1, 11) = 8.25$, $p = .015$, $\omega^2 = 0.36$) compared to the HOMER technique (Position: M = 0.10, SD = 0.13; Rotation: M = 23.14, SD = 25.75), (Figure 19). This result again indicates a consistent pattern where my technique, despite slower completion times, achieves greater precision in positioning and rotation.

6.6.9. Task F Analysis

In Task F, Shapiro-Wilk confirmed a normal distribution after log transformation of the data. However, significant deviations from variance homogeneity were observed in completion time ($F(1, 22) = 4.64$, $p = .0425$) and position difference ($F(1, 22) = 7.31$, $p = .0129$). The Welch Test indicated that my technique (M = 107, SD = 32.12) required a longer time to complete the task than HOMER technique (M = 68.42, SD = 26.81), with a significant impact, $F(1, 19.84) = 10.85$, $p = .0037$, $\omega^2 = 0.31$. Despite this, the position difference of my technique (M = 0.04; SD = 0.02) was lower than that of HOMER (M = 0.08; SD = 0.06), though the impact was not significant. As there was no violation of variance homogeneity in rotation difference, RM ANOVA demonstrated that my technique (M = 1.67; SD = 1.16) achieved significantly higher accuracy than HOMER technique (M = 11.47; SD = 16.37) in rotation, $F(1, 11) = 18.60$, $p = .0012$, $\omega^2 = 0.58$ (Figure 19).

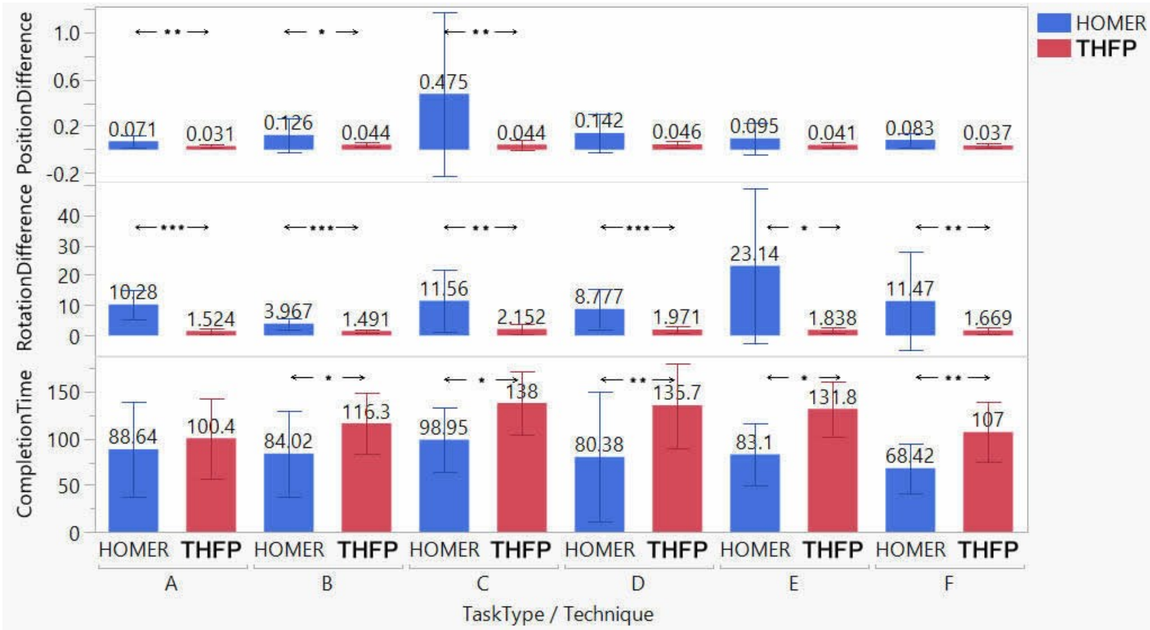

Figure 19. **Mean (PositionDifference) & Mean (RotationDifference) & Mean (CompletionTime) across all Tasks & Techniques.**

6.6.10. Overall Task Performance Analysis

The analysis from Task A to F offers a detailed comparison of the performance between my technique and HOMER technique across a range of tasks. The primary outcomes indicate that, although THFP typically leads to a slower achievement of task completion, it markedly enhances accuracy in both position and rotation. While the improvement in position accuracy for certain tasks for my technique is not always significant (Task D, E, F), the increase in rotational accuracy consistently remains substantial. In my visual analysis, which averaged all data in each task and plotted this average for both techniques, I discovered that, compared to the HOMER technique, my technique frequently achieves superior accuracy during the final stages of fine-tuning the target position and rotation. Even though my technique might experience occasional outliers, such as substantial movements due to overly rapid operations, my methods's easy-to-use handles, the undo feature, and the UI with numerical feedback, all help users quickly correct deviations (Figure 18). This assistance significantly enhances the accuracy upon task completion, surpassing the performance of the HOMER technique. I will address the reasons behind this in further detail in the subsequent discussion section.

6.6.11. SUS Analysis

Based on the SUS survey, my analysis showed a favorable reception towards my THFP technique's usability among participants. Shapiro-Wilk confirmed normal distribution, whereas Levene's test revealed a significant deviation from variance homogeneity, $F(1, 22) = 13.50$, $p = .0013$. Among the 12 participants, THFP achieved an average SUS score of 80.83 with a standard deviation of 4.81, significantly surpassing the well-acknowledged usability benchmark of 68 [13]. In comparison to the HOMER technique, Welch ANOVA highlighted that my THFP technique received significantly higher SUS ratings than the HOMER technique ($M = 44.38$, $SD = 14.23$), with $F(1, 13.479) = 70.73$, $p < .0001$, $\omega^2 = 0.15$, indicating users generally found my technique to be more user-friendly and intuitive (Figure 20).

Further investigation revealed that participants reported high confidence in the THFP technique, with a mean score of 4.33, significantly outperforming HOMER's mean of 2.92. My THFP technique also scored higher in terms of consistency and ease of use, with mean scores of 4.58, compared to HOMER's 2.25. However, the perceived complexity and the initial learning required before effectively using THFP were noted as areas for improvement, both receiving a mean score of 2. While these results highlight my technique's strengths in user confidence, consistency, and usability, a mean score of 1.92 for cumbersomeness points to a variation in user experiences. Despite overall positive feedback, some users encountered challenges, highlighting the need for addressing usability concerns and potentially simplifying the user experience in future iterations.

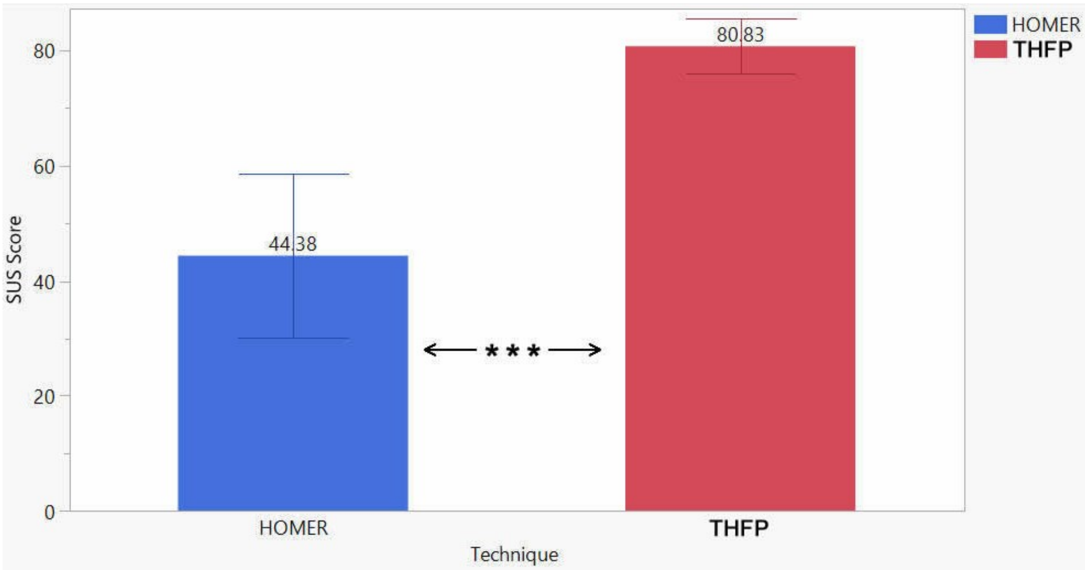

Figure 20. Mean(SUS Score) vs. Technique.

6.6.12. NASA-TLX Analysis

In the comparative analysis of NASA-TLX results between HOMER and THFP, my technique demonstrates a clear advantage. The RM ANOVA results show a significantly lower overall workload score for THFP (18.75) compared to HOMER (28.67), with a substantial statistical effect ($F(1, 11) = 55.73$, $p < .0001$, $\omega^2 = 0.81$). These findings indicate that THFP significantly reduces perceived workload across all categories, underlining its efficiency and user-friendly design. Supporting these quantitative results from the NASA-TLX, participant's verbal feedback further emphasizes the practical benefits of THFP. Users described THFP as "relaxing" and praised its precise control features, with comments highlighting the ease of "dragging objects with their fingers" which felt more "precise and relaxing." In contrast, feedback on HOMER highlighted its physical demands, with users reporting increased arm fatigue and likening the experience to a workout due to its manual intensity. This qualitative feedback aligns seamlessly with the statistical data, reinforcing the superior user experience offered by THFP.

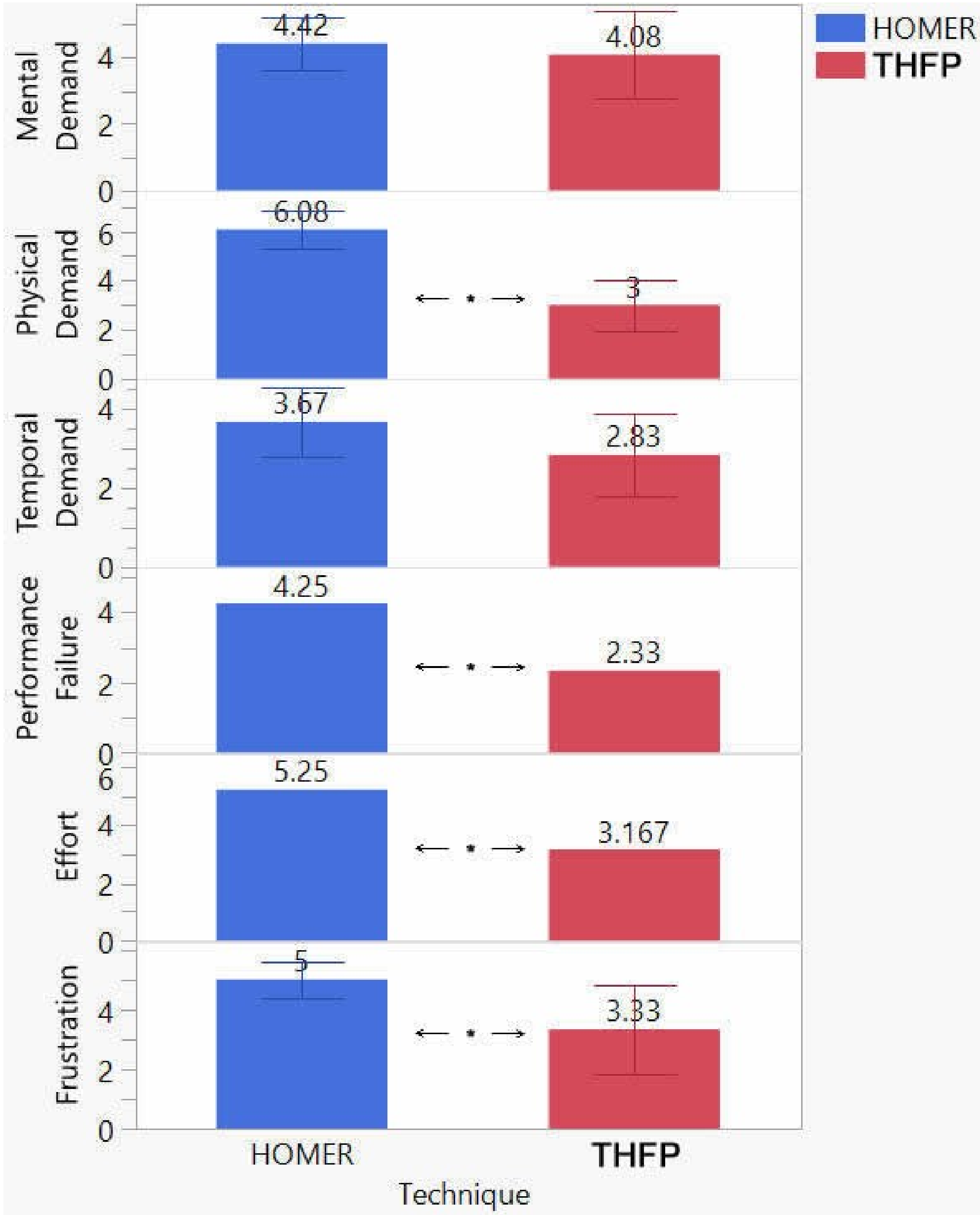

Figure 21. **Mean(NASA-TLX) vs. Technique.**

A Wilcoxon Signed-Rank Test demonstrated that my technique significantly improved task completion success, with a lower average performance score of 2.33, indicating fewer failures, compared to HOMER's average of 4.25. This was substantiated to be statistically different by a z-value of -2.86 and a p-value of less than .05. Additionally, my technique required a significantly lower physical demand from users, evidenced by an average score of 3, as opposed to HOMER's 6.08, and this difference was significantly different, with a z-value of -4.11 and a p-value of less than .05. Regarding user effort and frustration, my technique showed significantly lower scores

(Effort: M = 3.17; Frustration: M = 3.33) than HOMER (Effort: M = 5.25; Frustration: M = 5), with z-values of -3.28 and -2.82, respectively, both with p-values of less than .05 (Figure 21). These statistical analyses indicate that my technique not only enhances task completion rate and user satisfaction but does so with less physical demand and frustration on the user, distinguishing it as a more effective and user-centric approach.

6.6.13. Equal-time Comparison

To analyze the performance of all participants over time, I recorded the position and rotation differences of objects relative to their target positions and rotations at 0.1-second intervals. In Figure 22 I show the mean position and rotation differences per user at every 0.1-second interval. The extended duration of the blue curves (HOMER technique) compared to the red curves (THFP technique) does not imply a longer overall completion time for HOMER. Instead, it reflects the fact that a few individual users dedicated more time to the HOMER technique, as also visible by the outliers in Figure 12.

Yet, despite some users allocating more time to HOMER, they did not achieve higher accuracy than with THFP. In cases where users spent similar amounts of time with both techniques, my approach consistently yielded superior positional accuracy over time. Moreover, both position and rotation differences exhibit a similar pattern approaching task completion, with the blue curves for HOMER appearing more jagged compared to the smoother red curves of THFP. This suggests that users made more frequent minor adjustments to position and rotation with HOMER but still struggled to achieve a stable and accurate location and orientation.

The visualization also reveals that THFP's advantage in rotation difference is not as consistent as its advantage in position difference. Reflecting many users' strategy of prioritizing position adjustments over fine-tuning rotation. THFP begins to show an advantage in terms of orientation (lower mean deviation) only approximately 80 seconds into the task. Once detailed adjustments begin, THFP converges to significantly more accurate results. Although THFP experiences some fluctuations due to rapid operations or erroneous manipulations, these fluctuations are generally less pronounced than HOMER's in both magnitude and duration. This indirectly demonstrates the challenges users face in achieving accurate rotational control with HOMER, even with repeated adjustments.

The increased fluctuation of the data in the right half of the graph is due to fewer data points being available at longer time intervals (and thus less averaging across data), causing more spikes to appear after individual data trace endpoints (Figure 22). Thus, the spike in position discrepancy at ~200 seconds is attributable to a large manipulation error and subsequent corrections by a user, rather than a systematic issue with my technique.

As a result, while HOMER may have a faster overall completion time, it exhibits considerable variation and plateaus at lower accuracy levels for both position and rotation. Conversely, THFP achieves consistently better overall pose accuracy. Furthermore, THFP has the potential for further improvements with advancements in finger tracking and other optimizations, whereas HOMER appears to be limited by a human's ability to accurately control their hand pose. My approaches' incorporation of user-friendly controls, an undo feature, and a UI with numerical feedback enables users to swiftly detect and correct deviations, as illustrated by the sharp triangular shapes in the graph. These features significantly enhance accuracy upon task completion, surpassing the performance of the HOMER technique.

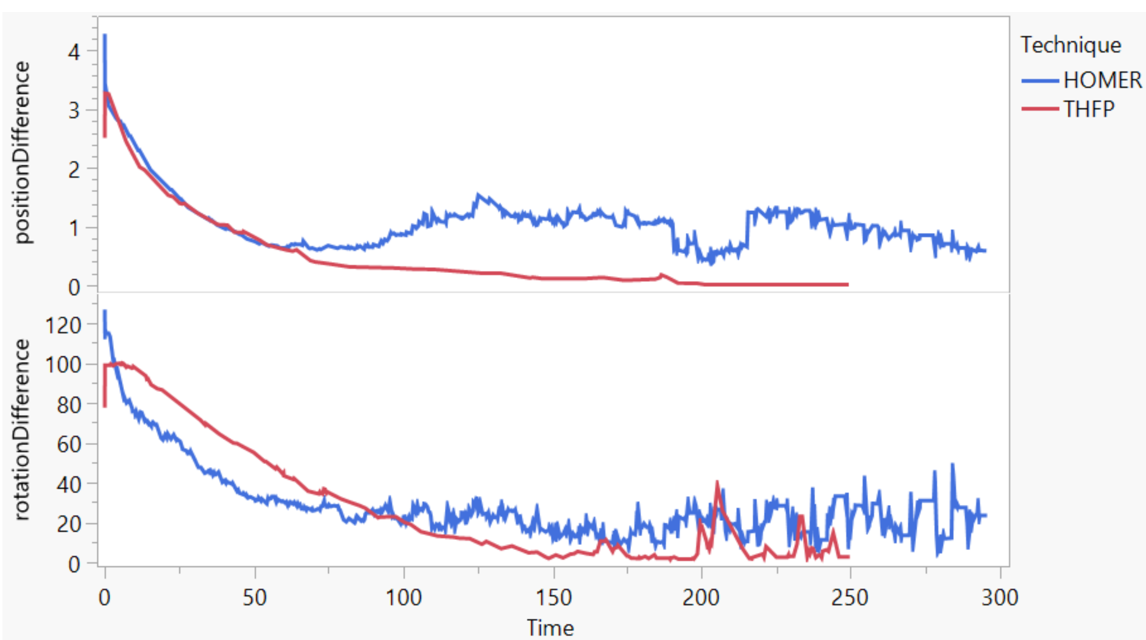

Figure 22. Mean (positionDifference & rotationDifference) vs. Time.

6.6.14. Individual Differences Comparison

To analyze whether individual differences among users impact their effectiveness with the HOMER and THFP techniques, I investigated users' performance metrics recorded during experiments in relation to their demographic data, visualizing this in charts to identify potential trends (Figure 23, 24, 25). For the three sets of engagement level data (with 3D software, video games, or VR familiarity), it appears that familiarity with immersive, interactive 3D environments—whether through 3D software, video games, or VR—correlates with an individual's ability to effectively utilize both HOMER and THFP techniques.

Across all three domains, I saw a clear pattern: individuals who frequently use 3D software, play video games, or engage with VR tend to achieve higher accuracy (lower position and rotation differences) when using both HOMER and THFP techniques. For example, individuals who engage with games or 3D software daily tend to complete tasks faster and with fewer fluctuations than those who engage less than once a month. Regarding VR familiarity, individuals who use VR a few times a week perform better than those who engage less frequently, suggesting that skills acquired in these areas are transferable and beneficial across different interactive 3D platforms. Overall, this is not surprising, as familiarity with 3D systems is likely to make other 3D manipulation tasks easier, too.

Additionally, the HOMER technique exhibits a clear improvement in time for frequent 3D/VR users, leading to quickly falling deviations over time. This indicates that HOMER's performance may benefit from existing practice and familiarity. In contrast, the THFP technique displays less variance in both rotation and position differences, exhibiting less variation across different levels of user engagement. This consistency suggests that THFP may be more user-friendly or less dependent on the frequency of 3D interactions, making it a potentially better option for novice or casual users.

For both techniques, there is a general trend of improvement over time in all user groups, highlighting a learning effect. However, the rate of improvement and its variation differ, with frequent users showing steadier progress. This analysis underscores the importance of considering user backgrounds with 3D systems when evaluating the effectiveness and user-friendliness of different VR interaction techniques.

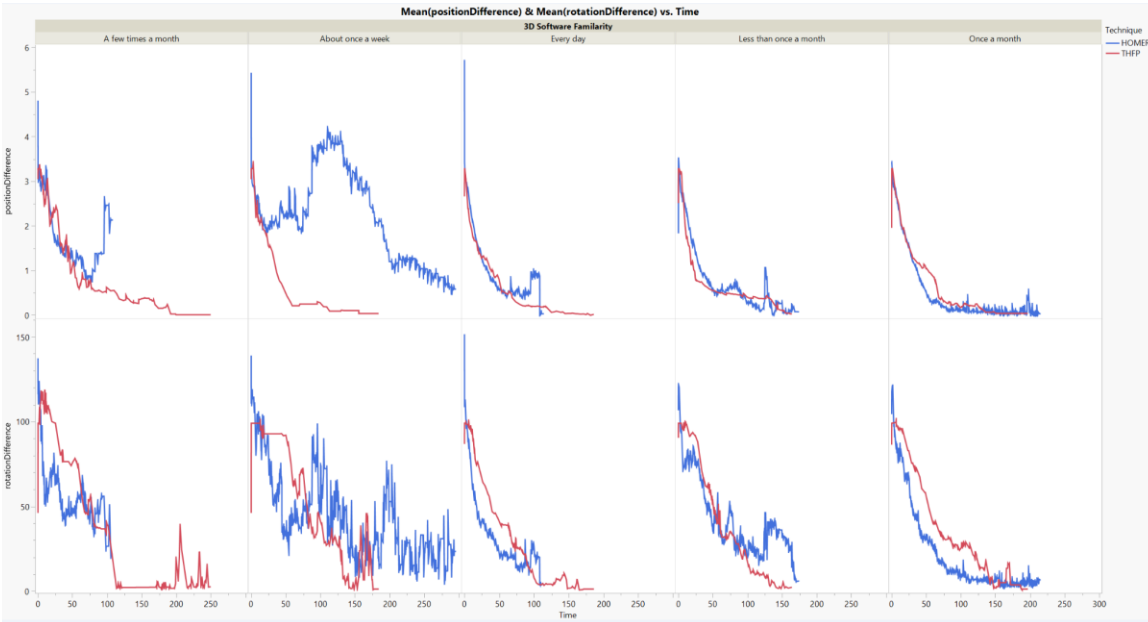

Figure 23. Mean (positionDifference & rotationDifference) vs. Time & 3D Software Engagement Levels.

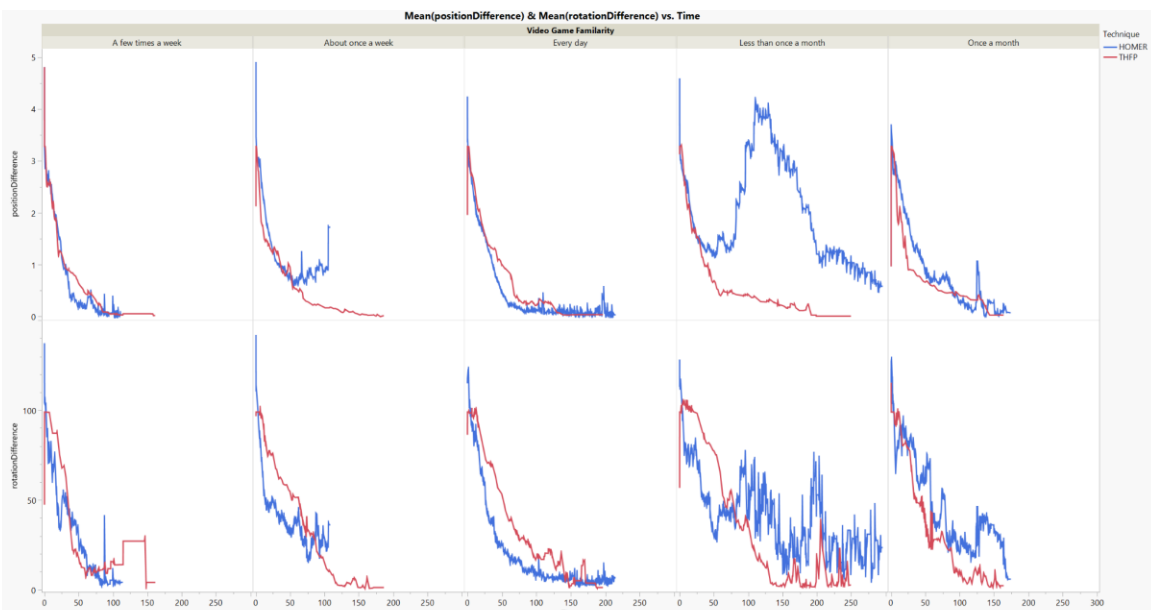

Figure 24. Mean (positionDifference & rotationDifference) vs. Time & Video Game Engagement Levels.

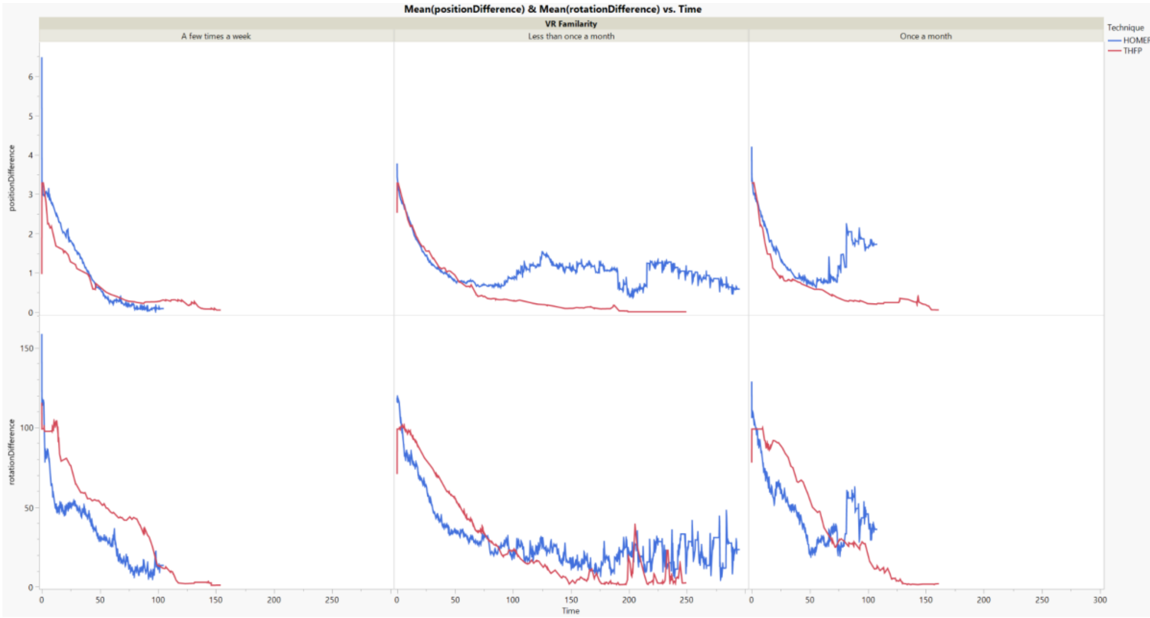

Figure 25. Mean (positionDifference & rotationDifference) vs. Time & VR Engagement Levels.

Rather than solely focusing on levels of familiarity with 3D environments, I also created a new chart that segments the performance of the HOMER and THFP techniques by gender (Figure 26). Based on the results, both genders show a general decline in position and rotation difference over time with both techniques. This suggests that regardless of gender, users improve their proficiency with the techniques as they continue to use them. The performance trends for HOMER and THFP, relative to the mean position and rotation differences over time, are consistent with previous findings.

However, I also observe that, in general, the male group chose to terminate tasks earlier with HOMER, while the female group used both techniques for longer. In terms of translation, it seems that males performed better with HOMER, while females did better with THFP. For rotation, both males and females performed better with THFP, but females exhibited more variation with HOMER. Despite the fact that male participants' data trails off earlier than the female participants' data, we cannot conclude that males are better than females in using these two techniques, as both user groups achieve broadly comparable final outcomes. Further, the gender ratio was highly unbalanced (3 females for each male) in my study, and the results are thus unlikely to be significant.

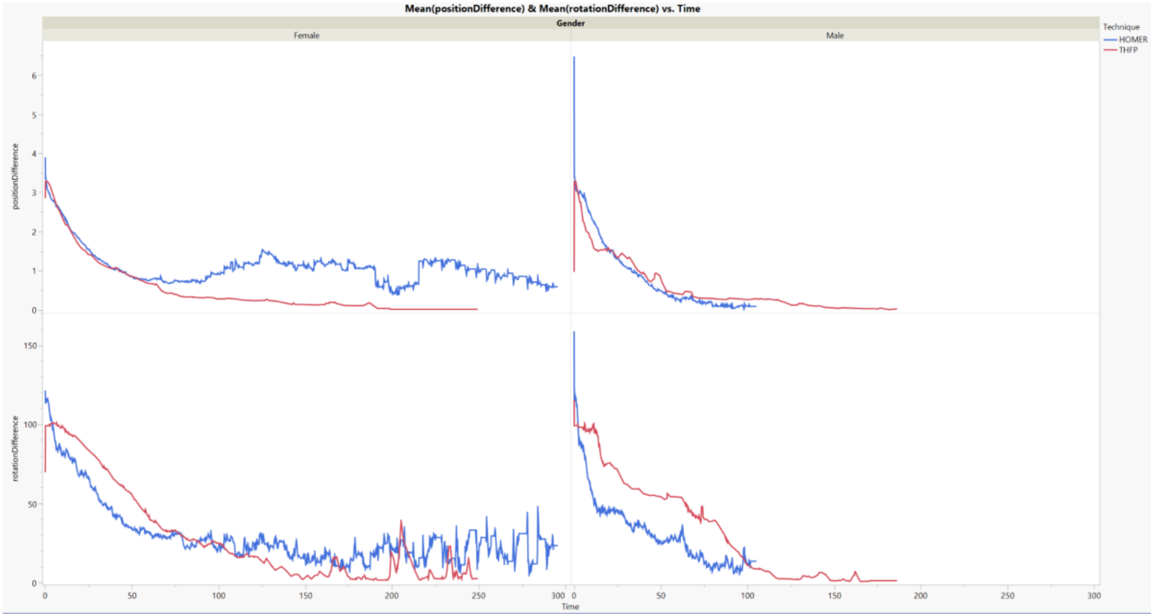

Figure 26. Mean (positionDifference & rotationDifference) vs. Time & Gender.

This limitation underscores the need for cautious interpretation, as the available data only offers a snapshot for a small and gender-unbalanced participant group. Future research could thus compare the performance of an equal number of males and females with similar backgrounds. Additionally, verifying the impact of all aspects of the demographic on the performance trends of the techniques, including familiarity with 3D environments and gender, would require a substantially larger number of subjects.

6.7. Discussion and Limitations

In summary, my bare-hand THFP technique significantly outperformed HOMER in accuracy for manipulating the position orientation of remote objects. Even if some tasks did not show a significant improvement in movement accuracy, the average errors were still smaller than with HOMER. Despite my THFP technique requiring more time to complete tasks, user feedback indicated that the provision of single-axis movement and rotation restrictions afforded more precise control. Consequently, users were willing to spend additional time making fine adjustments, which also contributed to longer times. Furthermore, analysis of individual tasks revealed that THFP tended towards more stable results towards the end of tasks, whereas the HOMER technique still suffered from small-scale object fluctuations, likely due to minor hand tremors or tracking issues. Consequently, my user study findings partially validate my RQs 2–4, even for participants with limited VR familiarity (8 out of 12). The analysis of the NASA-TLX results also showed that, compared to the HOMER Technique, my technique demanded lower physical and mental effort, thereby corroborating my RQ5. Further, and unlike HOMER, which requires significant physical movement, THFP facilitates precise VR interaction even within confined spaces [74] due to its minimal need for arm movement.

Although my THFP technique requires more time to complete the tasks compared to the HOMER method, when participants had the option to decide when the task was complete, it demonstrated comparable performance in accuracy control for both positioning and rotation. Through the pilot study presented above, I identified that while matching the orientation approximately can be fairly fast, it is necessary to also provide direct visual feedback on the object's angles to potentially achieve higher precision, especially if the object is far away. However, due to the limitations of HOMER in terms of hand tremors and/or tracking, participants were not able to match the results achievable with my THFP technique. Additionally, participants reported a more relaxed experience with THFP, as the need for constant hand pose monitoring was eliminated, due to natural proprioception of the finger positions.

My new technique could also assist users with gaze-based object selection technologies [49]. For example, by incorporating eye-tracking similar to the Apple Vision Pro [77], users would not need to focus their operational attention on their hands but could instead gaze at the object they wish to manipulate for precise remote operations.

Simultaneously, given the potential for efficiency and accuracy improvements, my main user study also identified some potentially confounding factors and improvement ideas.

6.7.1. Hand Tracking Issues

My pilot study revealed persistent challenges in hand tracking, particularly when fingers contacted the palm surface, causing occlusion and displacement issues. I aimed to reduce occlusions by promoting single-finger interactions with the opposing palm and switching to the Oculus SDK hand tracking version 2.1, which improved hand tracking performance. The Quest Pro also demonstrated superior FPS performance compared to the Quest, highlighting the importance of hardware improvements in enhancing hand tracking reliability. Despite these significant advancements [52], I could not completely eliminate tracking issues, as users occasionally relaxed their hands, fully opening their palms, which, along with quick hand movements, led to frame drops.

Similar to the issue I observed in the pilot test, I also noticed that participants would often place their non-dominant hand (NDH) close to their chest while operating with their dominant hand (DH) on the NDH palm, as this is a comfortable and ergonomic position. However, with such hand positioning, the hands can move out of view of the headset cameras, resulting in loss of tracking. This blind-spot issue underscores the need for further advancements in hand-tracking systems.

Additionally, to minimize the impact of ambient light on tracking, I conducted the new user tests in a brightly lit environment. However, I recognize the limitation that users might not always have access to brightly lit environments when using VR. Furthermore, to mitigate the influence of the Heisenberg effect, my technique allows users to drag and hold a slide handle, minimizing fingertip movement during object manipulation. Users can stabilize their finger on their other hand to further reduce unintended movements. Despite these measures, I could not completely eliminate instances where some users applied excessive force when gripping the handle with their index finger, leading to NDH movement that is too rapid for the hand tracking system and/or causes image blur, resulting in finger tracking issues, which in turn then can result in unexpected object movements.

Towards more reliable and accurate hand-based interactions in virtual reality applications, these observations underscore the necessity for ongoing advances in hand tracking precision and addressing issues related to occlusion, blind-spots, and environmental factors.

6.7.2. Control Modes and Consistent Interaction

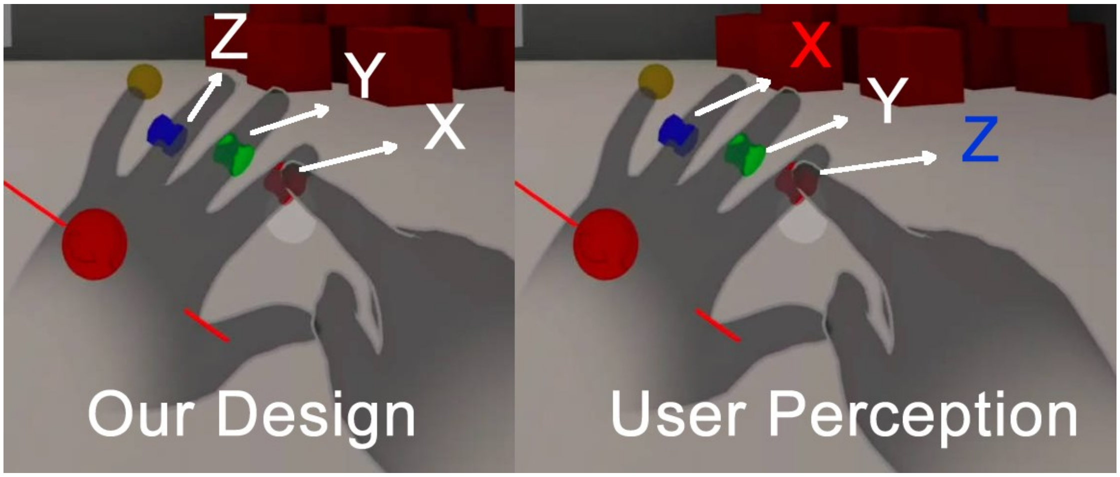

Figure 27. My THFP Technique Design vs. Design Favoured by a Third of Participants.

Building on insights from the pilot study, I enhanced the THFP interaction technique by incorporating additional axial constraints for the translation function, along with the capability to toggle between translation and rotation modes. In the main study, the THFP technique employed a distinct hand flip gesture for mode switching, an approach that users found exceptionally intuitive. This innovation facilitated a seamless transition between modes, enabling users to operate without any confusion or overlap in interaction options.

Despite the initial positive feedback on the intuitiveness of my mode-switching approach, my study revealed unexpected challenges. I aimed to map specific axes to colors and consistently assign them to different fingers, regardless of hand orientation. Yet, approximately one-third of my participants struggled to internalize this concept, relying heavily on proprioception for determining finger positions due to their focus on the remote object. When users flipped their palm, altering the hand orientation, they mistakenly believed that the control assigned to the three fingers should swap, i.e., the ring finger should govern the X-axis based on the fingers' left-to-right arrangement (Figure 27). This confusion resulted in incorrect object rotations, necessitating frequent use of the Undo function and reliance on additional UI feedback for corrections. Future research could explore aligning my design even more closely with user intuition to improve both understanding and efficiency. Additionally, it is essential to consider how potential user variables such as gender and previous VR experience may influence proprioception in VR settings.

6.7.3. Constraint Integration (3D Spatial Awareness and Movement Constraints)

In my pilot study, I realized that separating degrees of freedom (DOF) for both translation and rotation can improve operational precision. Consequently, in the main study, I modified my THFP technique by implementing a mode switch that allows object movement to be constrained along the x, y, and z axes using three fingers. However, for enhanced precision, the control through the wheel-handle must be accurate, implying that displacement speed cannot be too rapid, as this could lead to reduced efficiency. To counteract this, the THFP technique facilitates rapid manipulation through the palm-handle for quick, approximate positioning, while the wheel-handles are designated for fine control. Despite these provisions, I observed that many users relied solely on the wheel-handles for single-axis movement, which led to longer completion times. This underscores the necessity of balancing between speed and precision in the design of interaction techniques for virtual environments.

In addition, test results and user feedback also revealed confusion regarding the palm-handle's function. Although designed to allow users to control object movement along the current plane of the NDH's palm, facilitating free movement along all three axes simultaneously, participants needed to understand the mapping between the two-dimensional plane (dragging palm-handle direction) and the three-dimensional space (hand pose orientation). Some participants easily grasped this interaction, while others found it challenging and made frequent errors. Those unfamiliar with three-dimensional space often misunderstood the impact of the palm's orientation on the object's movement vector, indicating increased cognitive load.

Some users suggested simplifying the palm-handle to a trigger mechanism, moving the object in the direction the palm's fingertips are pointing. However, human anatomy limitations make certain hand poses more comfortable than others, making it difficult to move in the opposite direction of the palm by touch alone. Future research could consider constraining the palm-handle to drag only in the direction of the fingertips or the opposite direction, simplifying its use by reducing it to two possible directions: along or against the palm's orientation.

6.7.4. Interaction Technique

In the THFP technique, participants interact with virtual objects by manipulating a widget along their fingers, causing movements in distal objects along specified axes. To optimize the physical space utilized and enhance precision, I employed a rate control strategy similar to that used in joystick operations. This allows for finer control mappings on the fingers and coarser manipulations via the palm or back of the hand. Initially, I set the manipulation speed to a maximum of 1 cm per second to balance control precision and responsiveness. Despite this, participants often maximized the force exerted on the handle, leading to frequent overshooting errors. This observation likely reflects the speed-accuracy tradeoff [21], suggesting that higher interaction speeds, even at a fairly controlled rate, can easily compromise precision, highlighting the complex dynamics in human perception and motor control in VR environments.

The palm-handle was designed for quick repositioning of objects, but challenges arose when the direction of movement was ambiguous, leading to user errors and hesitation in subsequent attempts. Hurried attempts to rectify the movement can exacerbate such errors, contributing to increased task durations. These findings suggest a classic example of the speed-accuracy tradeoff in task performance [65]. Although reducing the maximum speed or adjusting the control-display ratio could mitigate this issue, it may decrease manipulation efficiency. Future work could incorporate the Adaptive Gain concept from Liu et al.'s work [42], applying a translation gain greater than 1 when the object is far from the target and a gain less than 1 when the object is close, enhancing both efficiency and precision.

Consistent with findings from the pilot test, I found that both the trackball and palm-handle lacked the precision needed for accurate object manipulation, attributed to the inherent challenges of interfacing with a virtual controller. The requisite thickness for these controls, intended to offset imprecise hand tracking, inadvertently leads to unintentional pose changes when the user's fingertip loses contact with the control surface. Progress in this domain is heavily dependent on future enhancements in the accuracy of hand-tracking technology. The use of physical sensors, such as haptic gloves, presents a promising alternate avenue [18]. Such gloves, capable of detecting touch initiation and termination, could potentially overcome the limitations associated with not sensing the physical presence of a virtual collider's thickness.

My analysis also revealed challenges with positional movements during Task F, as participants found it difficult to effectively move the palm-handle towards the wrist. This issue arises because the palm-handle moves along a flat plane, while human palms are not entirely flat. In future work, I plan to re-evaluate the shape of this interaction surface and redesign the palm-handle feature to match each individual's palm surface.

6.7.5. On The Applicability of Fitts' Law

In my study, and in line with (most of) the 3D manipulation literature, I elected not to investigate performance through Fitts' law, as it is currently not clear if Fitts' law is directly applicable to the tasks that I investigate. Fitts' law, traditionally used to model performance in 1D or 2D pointing tasks, may not sufficiently capture the complexities inherent in 3D interactions in VR HMDs, especially if objects are beyond arms' reach and there are potential depth perception issues due to the way stereo displays work. Prior work has considered the applicability of Fitts' law to 3D manipulations of objects within arm's reach [73], but did not propose a methodology to investigate 3D manipulation of objects at greater distances through Fitts' law. Further, there is evidence that Fitts' law does not describe human performance for 3D pointing at objects beyond arms' reach in current stereo display systems adequately [8], likely due to the presence of a vergence-accommodation conflict.

Moreover, my experimental design intentionally omitted variable-sized targets, which are essential for calculating the index of difficulty in Fitts' law experiments. Given the nature of the VR tasks designed for this study, including such targets could also have detracted from my primary research focus. As discussed in Batmaz et al.'s study [8], integrating traditional metrics like variable-sized targets might have skewed the objectives of the research, thereby emphasizing conventional assessments over the exploration of novel VR-specific interactions. On the other hand, in the future I could consider allowing users to manipulate objects of a variety of different sizes and shapes. This approach would then strengthen the generality of the results, ensuring that my findings are robust across different contexts and not limited to a specific set of conditions or object types. This would enhance the applicability of my techniques and findings to a broader range of real-world applications.

In conclusion, the decision to exclude Fitts' law from my study is based on a critical examination of its applicability in 3D VR contexts. This approach not only aligns with the current practice in the field of 3D user interface research but also points towards potential future research on developing and validating new models and methodologies that more accurately assess human performance in VR, thus enriching our broader understanding of human-technology interactions.

6.7.6. Manicure

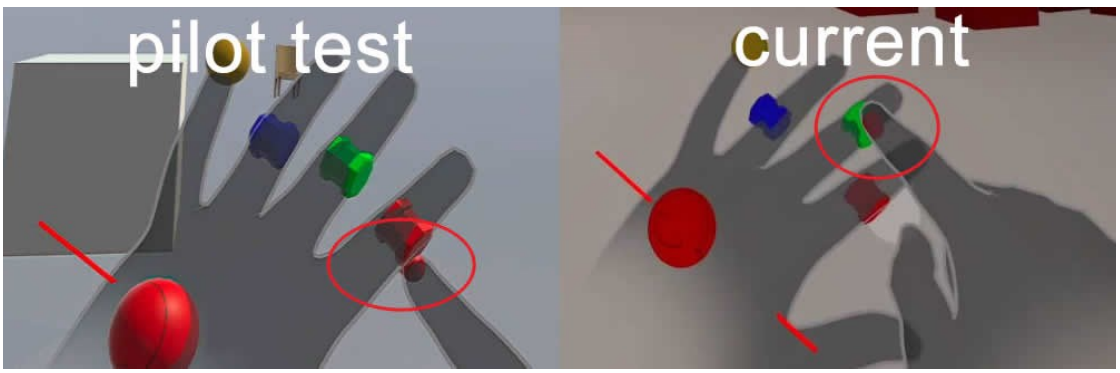

Figure 28. Contact point of fingertip in pilot test vs. current.

This observation in my pilot test prompted a deeper reflection on the THFP technique, emphasizing the need for a more inclusive approach for its interaction design. In my main user study, a significant majority of the participants in my experiment was female (9 out of 12). Therefore, in the refined version evaluated in the main user study, I continued testing the results of shifting the contact point to the dorsal side of the fingertip and found that this modification enhanced the technique’s usability (Figure 28).

6.7.7. Undo Functionality

While the aforementioned modifications, such as constraining axes and altering fingertip contact positions, typically reduce errors, I believe that recognition-based systems are currently unlikely to always work perfectly. Hence, based on feedback from pilot tests, I integrated an undo function to assist users in quickly reverting to a state before an error occurred (Figure 10). Within my study, I observed that users used the undo function effectively to correct excessive displacements or rotations caused by tracking issues or system sensitivity. While the average deviation plot showed that my technique still experienced significant deviations as the task neared completion, where users wished to make fine adjustments, the curve quickly returned to the starting point, forming sharp, triangular peaks rather than a continuous wave like pattern as with HOMER (Figure 18, Figure 22). According to my system logs, this was precisely the time when users employed the undo button, indicating that users could utilize this additional feature to prevent errors and improve accuracy.

Additionally, while the undo feature is a valuable addition to my THFP technique, it is important to note that only half of the participants utilized this function during the experiments. This suggests that the undo feature, although beneficial, did not play a critical role in the overall results. Furthermore, even if the undo feature were to be incorporated into the HOMER technique, it is unlikely to significantly impact the outcome of my comparative study, as half of the participants did not rely on it. Therefore, the inclusion of the undo feature in THFP does not diminish the validity of my findings or the superiority of my approach compared to HOMER. In future research, I plan to further investigate the impact of the undo function on technique accuracy.

6.7.8. Visual Discrimination for Distant Object Poses

Despite observations from a pilot test that the distance of objects impacts the accuracy of the HOMER technique—with greater distances leading to lower accuracy—this was not the primary focus of my research. My thesis concentrated on comparing the manipulation of distant objects using HOMER and THFP for the same tasks with objects at the same distances. Nevertheless, my study outcomes identified that when using HOMER, participants typically continued to make larger corrections for more distant objects, which might be due to HOMER amplifying inherent hand jitter and tracking errors. This adversely affects the efficiency and precision of control, providing a likely reason why HOMER exhibits more fluctuations in positional and rotational errors over time compared to my THFP technique. Although HOMER did not show significant differences in displacement accuracy compared to THFP in tasks D, E, and F, there was a notable difference in orientation. This indicates that for distant objects, even minor movements might cause additional unwanted rotational changes. While HOMER allows for simultaneous control of movement and rotation, this capability can also render the interaction technique more unstable. In contrast, THFP, by allowing only movement or rotation at any given time, sacrifices the potential efficiency of simultaneous approximate manipulation to gain a significant increase in accuracy.

My study also demonstrates that the superiority of my THFP technique over HOMER is not just due to the inclusion of the position/angles panel. User feedback revealed that even with a similar panel added to HOMER, users would still face significant challenges in achieving the same level of accuracy due to the inherent limitations of HOMER, where subtle hand movements can cause substantial deviations in object placement, making it time-consuming and difficult to attain accuracy, especially for distant objects. The key issue lies in the challenges users face when attempting to accurately place the object at the intended position, rather than their inability to perceive the correct position.

Previous work has frequently utilized UI widgets to enhance operational precision, yet the effectiveness of these widgets often hinges on their integration with supplementary mechanisms [3], [38], [50], [72], [81]. For instance, Lee et al.'s UMSR technique necessitates the use of additional hand-held menus to facilitate translation, rotation, and scaling operations with clear separation of degrees of freedom. Similarly, their BMSR approach also requires the creation of a scaled object replica and a bounding box around the object replica intended for manipulation, enabling users to interact with faces, edges, or vertices of the object [38]. These examples illustrate the need for combining UI widgets with other tools or systems to achieve desired levels of accuracy and functionality in user interactions.

Thus, while comparing HOMER with an added feedback panel or HOMER with NFSRV to my THFP technique would be an intriguing direction for future research, I believe that the presence of UI feedback alone does not directly determine the superiority of my approach. Instead, it is the combination of THFP's inherent properties, such as its ability to mitigate the impact of subtle involuntary hand movements on object placement, and the enhanced feedback provided by my UI, which complements these inherent advantages by offering users valuable visual references to quickly identify and more accurately correct positional and rotational deviations along each axis, which enables users to achieve higher levels of accuracy in object manipulation tasks at a distance.

6.7.9. 6Visual Environment

The screenshots of the scene used in my work appears darker in this document than the actual VR experience (Figure 11). In my studies, no users mentioned that the lighting in the scene might have affected their ability to see distant objects or the axes clearly. Post-test interviews revealed that many users did not rely on these axes; instead, they simply confirmed the axes of rotation and movement through direct interaction, i.e., they simply tried all axes until they found the correct one. Regarding the scene rendering, we opted for flat shading to avoid potential issues with rendering performance, which could impact interaction performance. However, with advancements in VR headset rendering performance, future experiments should consider more realistic scene settings. We also plan to evaluate whether our technique can maintain its consistent advantages in more complex and crowded environments.

6.7.10. Immersion in VR

One purpose of VR is to immerse users into an environment. In that context, it is crucial to consider how new object control methods either support or complicate such immersion. In my technique, the interaction between hands is intuitive because proprioception allows users to easily perceive the positions of their own hands and fingers. Operating objects through bare-hand interactions not only maintains the embodiment illusion but also enables users to focus entirely on the objects they are manipulating, potentially strengthening the place illusion [35], [67].

However, my technology introduces non-diegetic UI elements and control handles on the fingers, which do not involve direct physical contact with objects to receive haptic feedback from the environment [64]. Such a user interface can diminish the plausibility illusion [26]. Therefore, future work needs to consider how to integrate my UI more seamlessly into the actual environment to make its presence more logical and enhance the immersive experience. Additionally, VR immersion extends beyond simulating real spaces; it can foster belief in new superhuman abilities, such as remote control over objects [62]. During my experiments, users' misperceptions about controlling the x, y, z positions of their fingers paradoxically demonstrated their new-formed cognitive awareness of the capacity to remotely manipulate objects through THFP. This phenomenon, where users instinctively mapped spatial positions from left to right to define the x, y, z axes without considering a mapping to specific fingers, suggests that they subconsciously regard their fingers as part of their ability to remotely control objects. This intriguing finding warrants further exploration, also since it may indicate that the background of a user might also impact proprioceptive differences in VR.

Moreover, while direct hand-to-object interaction remains the most straightforward approach, engaging with distant objects invariably also introduces the need for tools, either moving oneself towards the object or bringing the object closer, such as through scaled replicas, widgets, or the HOMER technique. Many such methods introduce additional challenges related to human spatial orientation, extrapersonal space, and space constancy in VR, potentially causing spatial disorientation as discussed in works by Ruddle et al. [63] and Loomis & Knapp [43]. My THFP approach seeks a balance where users can accurately move and rotate objects beyond their immediate reach without relying on extra spatial movement. Future work should thus consider using the level of immersion as a criterion to quantitatively assess how new object manipulation techniques impact the VR experience, which is crucial for the practical application of our THFP technique within the VR field.

6.7.11. Limited sample size and Participant types

One significant limitation of my study is the sample size and composition. The experiment was conducted with a small cohort of 12 participants, predominantly female and right-handed. This sample size, while seemingly modest, aligns with similar studies in the field of 3D user interface research, e.g., [57], [70]. I acknowledge that the sample composition may not fully represent the broader population, potentially introducing biases and limiting the generalizability of my findings. It is thus important to interpret my results with caution due to these potential biases.

This also highlights the importance of including more diverse and larger samples in future studies to enhance the robustness and applicability of the findings. Additionally, while my experiment did not include trial repetitions, it would be interesting to add repetitions to a future study to investigate how participants adapt to my technique over time with repeated trials, which could provide deeper insights into the long-term efficacy and user acceptance of the approach.

Moreover, I also considered the inclusivity of the technique for left-handed users. Adapting the technique is straightforward, requiring merely a switch in the assignments between dominant and non-dominant hand, thus enhancing accessibility and ensuring a more representative experience across diverse populations.

Chapter 7.

Conclusion

In conclusion, this research presented a novel two-handed interaction technique for manipulating distant 3D objects in VEs. By extending the interactive zone to the entire NDH, users can intuitively and precisely control an object's position and rotation using their DH's index finger on the NDH's fingers and palm. A user study comparing my new THFP method with the HOMER technique revealed significant improvements in accuracy (at least 75% better), particularly for complex tasks such as hanging a painting on a slanted wall. Although the new technique was slower than HOMER (about 47% slower), it demonstrated less variation in completion time and still demonstrated substantially higher accuracy at the average HOMER end time.

Moreover, my new method received higher ratings on the System Usability Scale, indicating a more user-centric and intuitive design. The NASA Task Load Index further supports the effectiveness of the technique, showing reduced cognitive, physical, and temporal demands on users, as well as improved task completion rates and user satisfaction. The results of my user study also provide insights into options for error prevention, constraint integration, and the facilitation of consistent UI widgets.

My findings contribute to the advancement of intuitive and effective interaction techniques in VR/AR, expanding the possibilities for bare-handed interaction in these environments. As VR/AR technologies continue to evolve, the integration of this method into current systems can enable more accurate 3D object manipulation, ultimately enriching the user experience and increasing the potential for adoption in various application areas, such as computer-aided design and architecture. Consequently, my research lays a solid foundation for future developments in the field of controller-free VR systems, paving the way for more natural and immersive interactions in virtual environments.

References

[1] F. Argelaguet and C. Andujar, “A survey of 3D object selection techniques for virtual environments,” *Comput. Graph.*, vol. 37, no. 3, pp. 121–136, 2013.

[2] C. Auteri, M. Guerra, and S. Frees, “Increasing Precision for Extended Reach 3D Manipulation,” *Int. J. Virtual Real.*, vol. 12, no. 1, pp. 66–73, Jan. 2013, doi: 10.20870/IJVR.2013.12.1.2859.

[3] S. V. Babu, C. University, and W.-A. Hsieh, “The Benefits of Near-Field Manipulation and Viewing to Distant Object Manipulation in VR,” 2024.

[4] F. Bach, H. Çakmak, and H. Maaß, “Vision-based Hand Representation and Intuitive Virtual Object Manipulation in Mixed Reality,” *Biomed. Eng. Biomed. Tech.*, vol. 57, no. SI-1-Track-B, p. 000010151520124086, Aug. 2012, doi: 10.1515/bmt-2012-4086.

[5] F. Bacim, R. Kopper, and D. A. Bowman, “Design and evaluation of 3D selection techniques based on progressive refinement,” *Int. J. Hum.-Comput. Stud.*, vol. 71, no. 7–8, pp. 785–802, Jul. 2013, doi: 10.1016/j.ijhcs.2013.03.003.

[6] E. Balaa, M. Raynal, Y. B. Issa, and E. Dubois, “Classification of Interaction Techniques in the 3D Virtual Environment on Mobile Devices,” in *Virtual, Augmented and Mixed Reality. Designing and Developing Virtual and Augmented Environments*, vol. 8525, R. Shumaker and S. Lackey, Eds., in Lecture Notes in Computer Science, vol. 8525. , Cham: Springer International Publishing, 2014, pp. 3–13. doi: 10.1007/978-3-319-07458-0_1.

[7] S. Bateman, R. L. Mandryk, C. Gutwin, and R. Xiao, “Analysis and comparison of target assistance techniques for relative ray-cast pointing,” *Int. J. Hum.-Comput. Stud.*, vol. 71, no. 5, pp. 511–532, May 2013, doi: 10.1016/j.ijhcs.2012.12.006.

[8] A. U. Batmaz, R. Turkmen, M. Sarac, M. D. Barrera Machuca, and W. Stuerzlinger, “Re-investigating the Effect of the Vergence-Accommodation Conflict on 3D Pointing,” in *29th ACM Symposium on Virtual Reality Software and Technology*, Christchurch New Zealand: ACM, Oct. 2023, pp. 1–10. doi: 10.1145/3611659.3615686.

[9] A. Bellarbi *et al.*, “A 3D interaction technique for selection and manipulation distant objects in augmented reality,” in *2017 5th International Conference on Electrical Engineering - Boumerdes (ICEE-B)*, Boumerdes: IEEE, Oct. 2017, pp. 1–5. doi: 10.1109/ICEE-B.2017.8192012.

[10] F. Bettio, A. Giachetti, E. Gobbetti, F. Marton, and G. Pintore, “A practical vision based approach to unencumbered direct spatial manipulation in virtual worlds,” 2007.

[11] D. A. Bowman, C. A. Wingrave, J. M. Campbell, V. Q. Ly, and C. J. Rhoton, "Novel Uses of Pinch Gloves™ for Virtual Environment Interaction Techniques," *Virtual Real.*, vol. 6, no. 3, pp. 122–129, Oct. 2002, doi: 10.1007/s100550200013.

[12] 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 - SI3D '97*, Providence, Rhode Island, United States: ACM Press, 1997, p. 35-ff. doi: 10.1145/253284.253301.

[13] J. Brooke, "SUS - A quick and dirty usability scale," 1996.

[14] G. Buckingham, "Hand Tracking for Immersive Virtual Reality: Opportunities and Challenges," *Front. Virtual Real.*, vol. 2, p. 728461, Oct. 2021, doi: 10.3389/frvir.2021.728461.

[15] W. Büschel, A. Mitschick, T. Meyer, and R. Dachselt, "Investigating Smartphone-based Pan and Zoom in 3D Data Spaces in Augmented Reality," in *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*, Taipei Taiwan: ACM, Oct. 2019, pp. 1–13. doi: 10.1145/3338286.3340113.

[16] F. M. Caputo, M. Emporio, and A. Giachetti, "The Smart Pin: An effective tool for object manipulation in immersive virtual reality environments," *Comput. Graph.*, vol. 74, pp. 225–233, Aug. 2018, doi: 10.1016/j.cag.2018.05.019.

[17] I. Cho and Z. Wartell, "Evaluation of a bimanual simultaneous 7DOF interaction technique in virtual environments," in *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, Arles, France: IEEE, Mar. 2015, pp. 133–136. doi: 10.1109/3DUI.2015.7131738.

[18] C. M. J. Cox, B. Hicks, J. Gopsill, and C. Snider, "FROM HAPTIC INTERACTION TO DESIGN INSIGHT: AN EMPIRICAL COMPARISON OF COMMERCIAL HAND-TRACKING TECHNOLOGY," *Proc. Des. Soc.*, vol. 3, pp. 1965–1974, Jul. 2023, doi: 10.1017/pds.2023.197.

[19] B. R. De Araújo, G. Casiez, J. A. Jorge, and M. Hachet, "Mockup Builder: 3D modeling on and above the surface," *Comput. Graph.*, vol. 37, no. 3, pp. 165–178, May 2013, doi: 10.1016/j.cag.2012.12.005.

[20] A. O. S. Feiner, "The flexible pointer: An interaction technique for selection in augmented and virtual reality," in *Proc. UIST*, 2003, pp. 81–82.

[21] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement.," *J. Exp. Psychol.*, vol. 47, no. 6, pp. 381–391, 1954, doi: 10.1037/h0055392.

[22] S. Frees, G. D. Kessler, and E. Kay, “PRISM interaction for enhancing control in immersive virtual environments,” *ACM Trans. Comput.-Hum. Interact. TOCHI*, vol. 14, no. 1, pp. 2-es, 2007.

[23] S. Frees and G. D. Kessler, “Precise and rapid interaction through scaled manipulation in immersive virtual environments,” in *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, IEEE, 2005, pp. 99–106.

[24] P. C. Gloumeau, W. Stuerzlinger, and J. Han, “PinNPivot: Object Manipulation Using Pins in Immersive Virtual Environments,” *IEEE Trans. Vis. Comput. Graph.*, vol. 27, no. 4, pp. 2488–2494, Apr. 2021, doi: 10.1109/TVCG.2020.2987834.

[25] E. S. Goh, M. S. Sunar, and A. W. Ismail, “3D Object Manipulation Techniques in Handheld Mobile Augmented Reality Interface: A Review,” *IEEE Access*, vol. 7, pp. 40581–40601, 2019, doi: 10.1109/ACCESS.2019.2906394.

[26] M. Gottsacker, N. Norouzi, K. Kim, G. Bruder, and G. Welch, “Diegetic Representations for Seamless Cross-Reality Interruptions,” in *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Bari, Italy: IEEE, Oct. 2021, pp. 310–319. doi: 10.1109/ISMAR52148.2021.00047.

[27] A. Hameed, A. Perkis, and S. Möller, “Evaluating hand-tracking interaction for performing motor-tasks in vr learning environments,” in *2021 13th international conference on quality of multimedia experience (QoMEX)*, IEEE, 2021, pp. 219–224.

[28] H. M. Hammouri, R. T. Sabo, R. Alsaadawi, and K. A. Kheirallah, “Handling Skewed Data: A Comparison of Two Popular Methods,” *Appl. Sci.*, vol. 10, no. 18, p. 6247, Sep. 2020, doi: 10.3390/app10186247.

[29] C. Hand, “A Survey of 3D Interaction Techniques,” *Comput. Graph. Forum*, vol. 16, no. 5, pp. 269–281, Dec. 1997, doi: 10.1111/1467-8659.00194.

[30] D. Hayatpur, S. Heo, H. Xia, W. Stuerzlinger, and D. Wigdor, “Plane, Ray, and Point: Enabling Precise Spatial Manipulations with Shape Constraints,” in *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, New Orleans LA USA: ACM, Oct. 2019, pp. 1185–1195. doi: 10.1145/3332165.3347916.

[31] O. Hilliges, S. Izadi, A. D. Wilson, S. Hodges, A. Garcia-Mendoza, and A. Butz, “Interactions in the air: adding further depth to interactive tabletops,” in *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, Victoria BC Canada: ACM, Oct. 2009, pp. 139–148. doi: 10.1145/1622176.1622203.

[32] J. Jankowski and M. Hachet, “A Survey of Interaction Techniques for Interactive 3D Environments,” 2013.

[33] M. Kari and C. Holz, “HandyCast: Phone-based Bimanual Input for Virtual Reality in Mobile and Space-Constrained Settings via Pose-and-Touch Transfer,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg Germany: ACM, Apr. 2023, pp. 1–15. doi: 10.1145/3544548.3580677.

[34] 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*, vol. 8, no. 3, p. 60, Sep. 2021, doi: 10.3390/informatics8030060.

[35] K. Kilteni, R. Groten, and M. Slater, “The Sense of Embodiment in Virtual Reality,” *Presence Teleoperators Virtual Environ.*, vol. 21, no. 4, pp. 373–387, Nov. 2012, doi: 10.1162/PRES_a_00124.

[36] S. Kim, “Evaluating an In-Hand Ball-Shaped Controller for Object Manipulation in Virtual Reality,” 2024.

[37] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev, *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017.

[38] C.-Y. Lee, W.-A. Hsieh, D. Brickler, S. V. Babu, and J.-H. Chuang, “Design and Empirical Evaluation of a Novel Near-field Interaction Metaphor on Distant Object Manipulation in VR,” in *Symposium on Spatial User Interaction*, Virtual Event USA: ACM, Nov. 2021, pp. 1–11. doi: 10.1145/3485279.3485296.

[39] S. Lee, J. Seo, G. J. Kim, and C.-M. Park, “Evaluation of pointing techniques for ray casting selection in virtual environments,” in *Third International Conference on Virtual Reality and Its Application in Industry*, SPIE, Apr. 2003, pp. 38–44. doi: 10.1117/12.497665.

[40] H. N. Liang, C. Williams, M. Semegen, W. Stuerzlinger, and P. Irani, “An investigation of suitable interactions for 3D manipulation of distant objects through a mobile device,” *Int. J. Innov. Comput. Inf. Control*, vol. 9, no. 12, pp. 4737–4752, 2013.

[41] J. Liang and M. Green, “JDCAD: A highly interactive 3D modeling system,” *Comput. Graph.*, vol. 18, no. 4, pp. 499–506, Jul. 1994, doi: 10.1016/0097-8493(94)90062-0.

[42] X. Liu, L. Wang, S. Luan, X. Shi, and X. Liu, “Distant Object Manipulation with Adaptive Gains in Virtual Reality,” in *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Singapore, Singapore: IEEE, Oct. 2022, pp. 739–747. doi: 10.1109/ISMAR55827.2022.00092.

[43] J. M. Loomis and J. M. Knapp, “Virtual and Adaptive Environments: Applications, Implications, and Human Performance Issues”.

[44] H. Lovasz-Bukvova, M. Hölzl, G. Kormann-Hainzl, T. Moser, T. Zigart, and S. Schlund, "Usability and Task Load of Applications in Augmented and Virtual Reality: How Applicable are the Technologies in Corporate Settings?," in *Systems, Software and Services Process Improvement: 28th European Conference, EuroSPI 2021, Krems, Austria, September 1–3, 2021, Proceedings 28*, Springer, 2021, pp. 708–718.

[45] D. Mendes, F. M. Caputo, A. Giachetti, A. Ferreira, and J. Jorge, "A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments," in *Computer graphics forum*, Wiley Online Library, 2019, pp. 21–45.

[46] M. R. Mine, F. P. Brooks Jr, and C. H. Sequin, "Moving objects in space: exploiting proprioception in virtual-environment interaction," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, 1997, pp. 19–26.

[47] M. R. Mine, "Working in a Virtual World: Interaction Techniques Used in the Chapel Hill Immersive Modeling Program," 1996.

[48] A. Mossel, B. Venditti, and H. Kaufmann, "3DTouch and HOMER-S: intuitive manipulation techniques for one-handed handheld augmented reality," in *Proceedings of the virtual reality international conference: laval virtual*, 2013, pp. 1–10.

[49] A. K. Mutasim, A. U. Batmaz, and W. Stuerzlinger, "Pinch, Click, or Dwell: Comparing Different Selection Techniques for Eye-Gaze-Based Pointing in Virtual Reality," in *ACM Symposium on Eye Tracking Research and Applications*, Virtual Event Germany: ACM, May 2021, pp. 1–7. doi: 10.1145/3448018.3457998.

[50] T.-T. H. Nguyen, T. Duval, and C. Pontonnier, "A New Direct Manipulation Technique for Immersive 3D Virtual Environments," *ICAT-EGVE 2014 - Int. Conf. Artif. Real. Telexistence Eurographics Symp. Virtual Environ.*, p. 8 pages, 2014, doi: 10.2312/VE.20141367.

[51] Oculus Developer Portal, "Distance Hand Grab Interactions," 2023. [Online]. Available: https://developer.oculus.com/documentation/unity/unity-isdk-distance-hand-grab-interaction/

[52] Oculus Developer Portal, "Dive into the Latest Hand Tracking Improvements with v2.1 Update.," Jan. 2023. [Online]. Available: https://developer.oculus.com/blog/hand-tracking-improvements-v2-1

[53] N. Ouramdane, S. Otmane, F. Davesne, and M. Mallem, "FOLLOW-ME: a new 3D interaction technique based on virtual guides and granularity of interaction," in *Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications*, Hong Kong China: ACM, Jun. 2006, pp. 137–144. doi: 10.1145/1128923.1128945.

[54] R. Pausch, T. Burnette, D. Brockway, and M. E. Weiblen, “Navigation and locomotion in virtual worlds via flight into hand-held miniatures,” in *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques - SIGGRAPH ’95*, Not Known: ACM Press, 1995, pp. 399–400. doi: 10.1145/218380.218495.

[55] S. Pei, A. Chen, J. Lee, and Y. Zhang, “Hand Interfaces: Using Hands to Imitate Objects in AR/VR for Expressive Interactions,” in *CHI Conference on Human Factors in Computing Systems*, New Orleans LA USA: ACM, Apr. 2022, pp. 1–16. doi: 10.1145/3491102.3501898.

[56] S. A. Penumudi, V. A. Kuppam, J. H. Kim, and J. Hwang, “The effects of target location on musculoskeletal load, task performance, and subjective discomfort during virtual reality interactions,” *Appl. Ergon.*, vol. 84, p. 103010, 2020.

[57] J. S. Pierce and R. Pausch, “Comparing Voodoo Dolls and HOMER: Exploring the Importance of Feedback in Virtual Environments,” 2002.

[58] K. Pietroszek, “Virtual Hand Metaphor in Virtual Reality,” in *Encyclopedia of Computer Graphics and Games*, N. Lee, Ed., Cham: Springer International Publishing, 2018, pp. 1–3. doi: 10.1007/978-3-319-08234-9_178-1.

[59] K. Pietroszek, L. Tahai, J. R. Wallace, and E. Lank, “Watchcasting: Freehand 3D interaction with off-the-shelf smartwatch,” in *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, Los Angeles, CA, USA: IEEE, 2017, pp. 172–175. doi: 10.1109/3DUI.2017.7893335.

[60] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa, “The go-go interaction technique: non-linear mapping for direct manipulation in VR,” in *Proceedings of the 9th annual ACM symposium on User interface software and technology - UIST ’96*, Seattle, Washington, United States: ACM Press, 1996, pp. 79–80. doi: 10.1145/237091.237102.

[61] I. Poupyrev and T. Ichikawa, “Manipulating objects in virtual worlds: Categorization and empirical evaluation of interaction techniques,” *J. Vis. Lang. Comput.*, vol. 10, no. 1, pp. 19–35, 1999.

[62] R. S. Rosenberg, S. L. Baughman, and J. N. Bailenson, “Virtual Superheroes: Using Superpowers in Virtual Reality to Encourage Prosocial Behavior,” *PLoS ONE*, vol. 8, no. 1, p. e55003, Jan. 2013, doi: 10.1371/journal.pone.0055003.

[63] R. A. Ruddle, E. Volkova, and H. H. Bülthoff, “Learning to walk in virtual reality,” *ACM Trans. Appl. Percept.*, vol. 10, no. 2, pp. 1–17, May 2013, doi: 10.1145/2465780.2465785.

[64] P. Salomoni, C. Prandi, M. Roccetti, L. Casanova, L. Marchetti, and G. Marfia, “Diegetic user interfaces for virtual environments with HMDs: a user experience study with oculus rift,” *J. Multimodal User Interfaces*, vol. 11, no. 2, pp. 173–184, Jun. 2017, doi: 10.1007/s12193-016-0236-5.

[65] R. A. Schmidt, H. Zelaznik, B. Hawkins, J. S. Frank, and J. T. Quinn, “Motor-Output Variability: A Theory for the Accuracy of Rapid Motor Acts”.

[66] C. Shaw and M. Green, “Two-handed polygonal surface design,” in *Proceedings of the 7th annual ACM symposium on User interface software and technology - UIST ’94*, Marina del Rey, California, United States: ACM Press, 1994, pp. 205–212. doi: 10.1145/192426.197517.

[67] M. Slater, “Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 364, no. 1535, pp. 3549–3557, Dec. 2009, doi: 10.1098/rstb.2009.0138.

[68] P. Song, W. B. Goh, W. Hutama, C.-W. Fu, and X. Liu, “A handle bar metaphor for virtual object manipulation with mid-air interaction,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Austin Texas USA: ACM, May 2012, pp. 1297–1306. doi: 10.1145/2207676.2208585.

[69] R. Stoakley, M. J. Conway, and R. Pausch, “Virtual reality on a WIM: interactive worlds in miniature,” in *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI ’95*, Denver, Colorado, United States: ACM Press, 1995, pp. 265–272. doi: 10.1145/223904.223938.

[70] J. Sun and W. Stuerzlinger, “Extended Sliding in Virtual Reality,” in *25th ACM Symposium on Virtual Reality Software and Technology*, Parramatta NSW Australia: ACM, Nov. 2019, pp. 1–5. doi: 10.1145/3359996.3364251.

[71] B. G. Tabachnick, L. S. Fidell, and J. B. Ullman, *Using multivariate statistics*, Seventh edition. NY, NY: Pearson, 2019.

[72] Thi Thuong Huyen Nguyen and T. Duval, “Poster: 3-Point++: A new technique for 3D manipulation of virtual objects,” in *2013 IEEE Symposium on 3D User Interfaces (3DUI)*, Orlando, FL: IEEE, Mar. 2013, pp. 165–166. doi: 10.1109/3DUI.2013.6550230.

[73] E. Triantafyllidis, W. Hu, C. McGreavy, and Z. Li, “Metrics for 3D Object Pointing and Manipulation in Virtual Reality: The Introduction and Validation of a Novel Approach in Measuring Human Performance,” *IEEE Robot. Autom. Mag.*, vol. 29, no. 1, pp. 76–91, Mar. 2022, doi: 10.1109/MRA.2021.3090070.

[74] W.-J. Tseng, S. Huron, E. Lecolinet, and J. Gugenheimer, “FingerMapper: Mapping Finger Motions onto Virtual Arms to Enable Safe Virtual Reality Interaction in Confined Spaces,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Apr. 2023, pp. 1–14. doi: 10.1145/3544548.3580736.

[75] J.-N. Voigt-Antons, T. Kojić, D. Ali, and S. Möller, “Influence of Hand Tracking as a way of Interaction in Virtual Reality on User Experience.” arXiv, Apr. 27, 2020. Accessed: Mar. 16, 2024. [Online]. Available: http://arxiv.org/abs/2004.12642

[76] J. Wagner, W. Stuerzlinger, and L. Nedel, “Comparing and Combining Virtual Hand and Virtual Ray Pointer Interactions for Data Manipulation in Immersive Analytics,” *IEEE Trans. Vis. Comput. Graph.*, vol. 27, no. 5, pp. 2513–2523, May 2021, doi: 10.1109/TVCG.2021.3067759.

[77] E. Waisberg *et al.*, “Apple Vision Pro and why extended reality will revolutionize the future of medicine,” *Ir. J. Med. Sci. 1971-*, vol. 193, no. 1, pp. 531–532, 2024.

[78] C. A. Wingrave, Y. Haciahmetoglu, and D. A. Bowman, “Overcoming World in Miniature Limitations by a Scaled and Scrolling WIM,” 2006.

[79] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins, “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: ACM, May 2011, pp. 143–146. doi: 10.1145/1978942.1978963.

[80] P. Yao, T. Zhu, and M. Zyda, “Adjustable Pointer in Virtual Reality for Ergonomic Interaction,” in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, Atlanta, GA, USA: IEEE, Mar. 2020, pp. 828–829. doi: 10.1109/VRW50115.2020.00265.

[81] X. Zhang, W. He, M. Billinghurst, L. Yang, S. Feng, and D. Liu, “Design and Evaluation of Bare-Hand Interaction for Precise Manipulation of Distant Objects in AR,” *Int. J. Human–Computer Interact.*, pp. 1–15, Dec. 2022, doi: 10.1080/10447318.2022.2158527.

[82] Y. J. Zhao, D. Shuralyov, and W. Stuerzlinger, “Comparison of multiple 3D rotation methods,” in *2011 IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems Proceedings*, Ottawa, ON, Canada: IEEE, Sep. 2011, pp. 1–5. doi: 10.1109/VECIMS.2011.6053855.