AMP-IT and WISDOM: Improving 3D Manipulation for High-Precision Tasks in Virtual Reality

Francielly Rodrigues [1,2] * Alexander Giovannelli [2] † Leonardo Pavanatto [2] ‡ Haichao Miao [3] §
Jauvane C. de Oliveira [1] ¶ Doug A. Bowman [2] ‖

[1] National Laboratory of Scientific Computing, Brazil
[2] Center for Human-Computer Interaction, Virginia Tech, USA
[3] Lawrence Livermore National Laboratory, USA

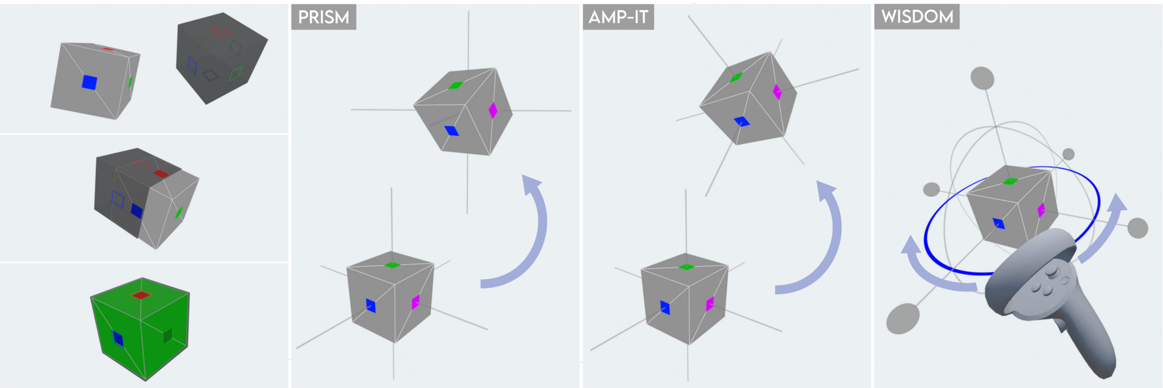


Figure 1: Left: Visualization of the Docking Task for assessing precise object manipulation in VR. Right: PRISM, AMP-IT, and WISDOM manipulation techniques.

ABSTRACT

Precise 3D manipulation in virtual reality (VR) is essential for effectively aligning virtual objects. However, state-of-the-art VR manipulation techniques have limitations when high levels of precision are required, including the unnaturalness caused by scaled rotations and the increase in time due to degree-of-freedom (DoF) separation in complex tasks. We designed two novel techniques to address these issues: AMP-IT, which offers direct manipulation with an adaptive scaled mapping for implicit DoF separation, and WISDOM, which offers a combination of Simple Virtual Hand and scaled indirect manipulation with explicit DoF separation. We compared these two techniques against baseline and state-of-the-art manipulation techniques in a controlled experiment. Results indicate that WISDOM and AMP-IT have significant advantages over best-practice techniques regarding task performance, usability, and user preference.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—Human computer interaction (HCI)—Interaction techniques

1 INTRODUCTION

Manipulation in virtual environments usually involves the application of spatial transformations to selected objects [3], including changes in position, orientation, and scale [5]. Mid-air interaction is a common manipulation method in virtual environments, offering an intuitive alternative to traditional 2D input methods like mouse and keyboard. However, unlike in real-world scenarios, physical properties such as the objects' weight and inertia are not naturally available in the virtual world [9]. In addition, such systems suffer from hand instability, imprecise tracking systems and difficulties in mapping the user's movements to virtual objects [18, 19]. These challenges are particularly evident in tasks that require high levels of control and accuracy, known as precise manipulation tasks.

A common use case for such systems is when comparing an ideal digital model against an imperfect physical object created based on that specification. In industrial applications such as additive manufacturing (AM), this alignment is essential for the defect inspection of printed components. Even a minor deviation from the measured geometry can indicate a potential defect in the virtually printed part. Achieving precise 3D alignment in order to do such inspections presents a demanding task with strict error tolerances.

In this context, although several studies have explored precise manipulation in virtual reality (VR), the existing techniques still present limitations, particularly when such exceptionally high precision is required. The challenge lies in achieving natural translational and rotational movements, minimizing errors, and maintaining efficiency in terms of time and workload. We intend to fill this gap by designing techniques that allow the user to naturally achieve a high level of precision without compromising time and the amount of effort required to learn and apply the technique.

This work presents an evaluation of two novel manipulation techniques through two separate experiments. The first experiment targeted standard precision tasks, while the second focused on extremely high precision tasks. Initially, our goal was to design a single novel technique, AMP-IT, which is why our initial experiment exclusively contrasts this technique with existing literature. However, when dealing with precise manipulation tasks, we recognized the

---

*e-mail: francielly@vt.edu, fmunique@posgrad.lncc.br
†e-mail: agiovannelli@vt.edu
‡e-mail: lpavanat@vt.edu
§e-mail: miao1@llnl.gov
¶e-mail: jauvane@acm.org
‖e-mail: dbowman@vt.edu

need to enhance one of the existing techniques from the literature to maintain a fair basis for comparison. This resulted in a second novel technique, called WISDOM. The outcomes of our investigation showcase AMP-IT's efficacy in standard precision tasks, and highlight the unexpectedly strong performance of WISDOM when compared to the others.

2 RELATED WORK

Numerous techniques have been developed for manipulating objects in 3D virtual environments. The strategies developed, however, vary greatly depending on the particularities of the problem being tackled, and there is not a consolidated technique that works well for most cases. In this context, some of the most relevant approaches include: breaking the task into several sub-tasks [1,4], zooming or scaling the environment (World in Miniature metaphor) [21,24,27], using control points or pins attached to the object's vertices [10,14], exploring Gesture-to-Force Mappings for Remote Manipulation [31], implementing auxiliary objects such as shape constraints [12], exploring bi-manual gestures, anchors or virtual handles [8,10,17, 26], enabling explicit DoF separation through collaboration [25] or through widgets [19,20], and scaling the movement by adjusting the control-display ratio [2,7,9,10].

For scenarios that do not require high precision, the impact of hand instability and the inaccuracies introduced by the input device or method may be less significant in terms of task completion. In this case, isomorphic techniques such as the Simple Virtual Hand (SVH) [5] metaphor may be sufficient. SVH is a direct manipulation technique that enables exact mapping, or 1:1 C/D (Control to Display) Ratio, of the user's physical hand movement to the movement of a virtual hand within the environment. The C/D Ratio quantifies the relationship between the physical movement of an input device and the resulting displacement of a virtual object. With a 1:1 C/D ratio, we have a direct correspondence between the input device and the virtual environment coordinate system [5,16].

Although very intuitive, as the level of precision required increases, the SVH technique suffers from its extreme sensitivity to hand instability and restricted reach limited by arm size [5]. Several studies proposed changes on the transfer function to mitigate those issues, either by magnifying the movement to achieve long-distance reach and controlled manipulation [28,31], or by scaling down the motion to increase accuracy and control [9,15].

The PRISM technique, developed by Frees et al. [9], uses a scaled mapping that dynamically adjusts the C/D ratio based on the user's hand speed to provide a higher level of control when moving slowly, and direct, unrestricted manipulation when moving quickly. Although the results are promising, the main limitation of this technique is that, because they use global axes to calculate the scaling factors for translation, diagonal movements are much more scaled than movements along a principal axis. Additionally, scaled rotation was reported as unnatural by the users.

The aforementioned techniques allow for direct object manipulation in six degrees-of-freedom (6DoF) simultaneously. Alternatively, one can limit the simultaneous degrees of freedom. Mendes et al. [19] proposed to employ virtual widgets attached to objects. Users interact by grabbing a sphere linked to an axis, enabling single DoF translation or rotation. Compared to PRISM and SVH, their results show that DoF separation benefits precision in spatial manipulations, but considerably increases time for complex tasks.

The same authors later developed MAiOR [20], a bi-manual technique that allows for six, three, and one DoF separation with custom axis locking using the non-dominant hand. Although the new technique did not compromise completion time, the traditional widgets had the best performance overall. Single DoF manipulations were mostly not used, and rotation remained challenging to understand.

We aim to improve the existing work in several ways. First, we explore the benefits of an implicit DoF separation, capturing user intent intuitively without the need for explicit axis selection as seen in traditional widgets. Our goal is to design a natural technique that remains time-efficient and does not impose extra effort on the user. Second, we propose changing the frame of reference from the global space to the local space of the controlled object. By doing so, we expect the manipulation to align with the user's expectations, particularly when the object is rotated. Third, our aim is to develop a technique that effectively explores the benefits of scaled manipulation, as proposed by PRISM, but enhanced with implicit DoF separation for both rotation and translation, enabling a more natural and controlled motion perception. Finally, we strive for a technique that excels in high-precision scenarios without compromising the motion when high precision is not a requirement, offering a balanced manipulation experience in various scenarios.

3 TECHNIQUE DESIGN

We designed two novel techniques, AMP-IT and WISDOM. Initially, we intended to design one direct manipulation technique to compare with existing state-of-the-art techniques, including SVH, PRISM, and mid-air indirect manipulation through widgets [19]. The SVH and PRISM techniques were implemented as benchmarks for comparing with our methods. SVH was implemented based on [5], and PRISM followed the description from [9]. However, an initial assessment revealed the unsuitability of the original widgets approach for our task (see 3.2 for further information). Since we were still interested in comparing the trade-offs between direct and indirect manipulation in our scenario, we substantially improved the technique, effectively creating another novel approach.

3.1 AMP-IT Technique

Based on the insights from section 2, we designed AMP-IT: Adaptive Mapping for Precise InTeraction. As AMP-IT was specifically designed as an improvement to PRISM, both techniques share significant similarities. Apart from the distinctions outlined below, the implementation can be inferred as derived from PRISM.

Similar to PRISM, AMP-IT assumes that the user's hand speed reflects their desired level of precision. In this context, a higher hand speed indicates a less precise intention, while a slower hand speed suggests a desire for controlled motion. By continuously monitoring the hand speed, such techniques dynamically adjust the C/D ratio to match the user's intent. This filters out hand instability while allowing direct manipulation when needed. The user's hand speed is computed per frame by measuring the controller's displacement within a predefined time interval $T = 500ms$ (see [9]). Thresholds $MinS$ (Minumum Speed) and SC (Scaling Constant) are used to determine the appropriate response for different hand speeds, and a scaling factor K is calculated based on these thresholds, controlling the movement behavior.

$$K = 1/CD = \begin{cases} 1 & \text{for } S_{hand} \geq \text{SC} \\ f(S_{hand}) & \text{for MinS} < S_{hand} < \text{SC} \\ 0 & \text{for } S_{hand} \leq \text{MinS} \end{cases} \tag{1}$$

$$D_{object} = K \cdot D_{hand} \tag{2}$$

From Equations 1 and 2, when the hand speed S_{hand} is below $MinS$, the object remains still ($K = 0$). Between $MinS$ and SC, the movement is scaled by a function of the hand speed, ($K = f(S_{hand})$). When the hand speed reaches or exceeds SC, the object's movement aligns with the hand's motion ($K = 1$). In Equation 2, D_{object} is the distance the controlled object will move, and D_{hand} is the hand displacement since the last frame. Through experimentation, threshold values for AMP-IT were determined to be $MinS$ = 0.02 m/s and SC = 0.9 m/s. Similarly, for PRISM, the most suitable values were found to be $MinS$ = 0.05 m/s and SC = 0.45 m/s (see [9]).

Different functions can be used to scale down the movement when the hand speed is between $MinS$ and SC. PRISM uses a linear

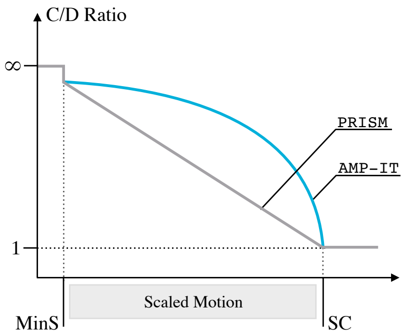


Figure 2: Comparison of the mapping between C/D Ratio and hand speed for PRISM and AMP-IT

function $f(S_{hand}) = S_{hand}/SC$, but this approach has limitations. If a significant movement occurs along one axis with minimal movement on others, an optimal $MinS$ value would effectively filter out these secondary motions. However, when aiming for high precision, the intentional execution of very small movements is common. In such cases, even the hand speed on the desired axis may fall below the $MinS$ threshold. To accommodate such small movements, an extremely small value for $MinS$ would be required, but setting $MinS$ too low would result in minimal or no filtering whatsoever.

To address this issue, AMP-IT implements an exponential function, $f(S_{hand}) = (0.02 \times 4^{S_{hand}})/SC$. When users aim for precision, their hand speeds along unwanted axes tend to be similar to the hand speed on the intended axis, all close to the $MinS$ threshold. Initially, we lower the $MinS$ value to avoid filtering out these movements entirely. Subsequently, by incorporating an exponential function, we achieve a smoother decrease in the C/D ratio for hand speed values near $MinS$, effectively minimizing undesired motions (see Fig. 2).

Another key difference between AMP-IT and PRISM lies in their approach to translation. PRISM operates on each world axis independently, breaking down the controller displacement into x, y, and z components within the global coordinate system to calculate hand speed, yielding three scaling factors: K_x, K_y, and K_z. AMP-IT also operates on each axis individually, but using local rather than global axes. The decision to use the object's local space for translation calculations in AMP-IT comes from observing user interactions. We observed that users naturally moved perpendicular to object faces when aligning rotated objects, following its local axis instead of global coordinates. If we were to use the global axes for translation calculations, the resulting scaled movement would still be diagonal, not aligned with the user's intention. Consequently, it would be challenging to effectively eliminate undesired motion. AMP-IT thus provides a more accurate and intuitive translation experience, allowing precise manipulation along the intended axes. This approach also addresses the $MinS$ constant issue, reducing diagonal movements and facilitating the exclusion of axes below the threshold.

AMP-IT also enhances rotation compared to PRISM, making it more intuitive and precise. While PRISM represents the rotation using a global quaternion and calculates a single angular hand speed and scaling factor K_r, AMP-IT takes a different approach. Similar to translation, we aim to convert rotations from global to local space to apply them separately to each rotational axis. Our approach involves initially converting the global rotation, represented by the quaternion, into a global angle-axis representation. The axis is subsequently transformed into the local space of the object. This transformation is straightforward, and the resulting local angle-axis representation can be then converted back into a new quaternion that captures the original rotation, but is specific to the object's local space. We can then decompose the quaternion into Pitch, Yaw, and Roll, the x, y, and z rotational axes. That way, we avoid the use of Euler angles at the conversion and are able to prevent Gimbal lock [13].

For our implementation of AMP-IT, the $MinS$ and SC values obtained through experimentation for rotation were 5 and 50 degrees per second, respectively.

3.2 WISDOM Technique

We initially reproduced the traditional widgets approach based on the state-of-the-art implementation presented by [19]. The only difference was in the design of rotation handles, where we chose a widely adopted approach seen in 3D modeling and game engines: the use of circular gizmos (i.e., Fig. 1). Manipulation using this approach was very straightforward. Users could manipulate only one degree of freedom (DoF) at a time. For translation, they grabbed one of the spheres connected to the x, y, or z axis to move the object solely along that axis. Similarly, for rotation, users grasped the gizmo positioned perpendicular to the desired axis and performed circular motions either clockwise or counterclockwise. However, we found that the original widget technique was notably less effective than other methods, especially for tasks requiring extremely precise manipulation, because users were always restricted to manipulating a single degree of freedom at a time, even when they did not have precision in mind. This resulted in an increased workload, longer task completion times, and higher levels of fatigue.

Thus, we decided to create a hybrid approach that enabled users to explicitly switch between SVH for coarse manipulation and widgets for precise fine-tuning of the object's position/orientation. We also found that isolating the degrees of freedom alone was insufficient to address the issue of hand instability, especially for rotation, which was particularly sensitive to small, unintentional hand movements. The lack of a precise motion mode made it nearly impossible to fine-tune the object's pose once it was in close proximity to the target. Thus, we applied the same scaled mapping concepts used in AMP-IT and PRISM to each of the translation and rotation axes separately, resulting in the development of a second novel technique called WISDOM: Widget-based Indirect Scaled mapping with Direct Object Manipulation. In this case, since we were dealing with a single DoF at a time and we did not need to filter or minimize unwanted movements on other axes, we opted to retain the linear function and thresholds employed by PRISM.

Unlike other direct manipulation techniques discussed earlier, the widgets technique does not require wrist movements for rotation. All physical movements using widgets are translational, thereby overcoming possible limitations related to the wrist and forearm range of motion [22]. Additionally, we did not encounter the same challenges in rotation as faced in the AMP-IT technique, as hand speed was calculated in a similar manner to translation, considering hand displacement over a given time interval.

4 EXPERIMENTAL APPROACH

Two experiments were conducted to compare techniques for manipulating a 3D object in VR with varying precision requirements. A docking task was implemented to evaluate performance and usability, considering the trade-offs between 1:1 and N:1 scaled mapping as well as direct and indirect manipulation.

4.1 Environment

Participants were standing and were allowed to walk within a limited indoor area (approximately $1m^2$). They were immersed in a virtual room with a white background and a floor area that served both as a reference point for the beginning of each task and as a delimiter of the area in which the participant could move. Participants were also able to see a representation of the handheld controller. The environment was implemented using the Unity game engine[1] version 2021.3.10f1 and the XR Interaction Toolkit[2] version 2.3.

---

[1] https://unity.com

[2] https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@2.3

4.2 Apparatus

In both experiments, participants were equipped with a Head-Worn Display (HWD), the Meta Quest 2. This HWD features an LCD Screen with a resolution of 1832 x 1920 per eye and a refresh rate of 90Hz [3]. Participants used a handheld Meta Quest 2 controller in their dominant hand to interact. All interactions required users to hold the trigger button on the controller while manipulating the object; releasing the trigger caused the virtual object to be released. Only the widget interactions required the user to touch the widgets with the controller; all direct manipulations occurred on trigger press whether the controller was touching the object or not.

4.3 Task

Participants performed docking tasks that involved manipulating a 3D cube with sides measuring 10cm. The objective of each task was to position the cube completely inside a designated target. Successful completion of the task required both proper positioning and the correct orientation of the cube within the target. Participants were encouraged to physically move around both the cube and the target object to obtain a better visual understanding.

The target object was set to be slightly transparent, allowing users to maintain visibility of the moving object even when it was completely inside the target. The moving cube was rendered with an opaque gray material and a white wireframe. With this configuration, users were also able to use the different shades of gray to facilitate the perception of the environment's depth, as an indication of whether a specific part of the moving object was inside or outside the target. To ensure distinct identification of each face, different colors were assigned such that both objects had corresponding faces sharing the same color. The moving object had solid squares filled with the assigned color on each face, while the target object had slightly larger squares with only the colored borders (see Fig. 1).

The positions of both objects were set at the beginning of each trial so that the moving cube was oriented to face the participant, with no initial rotation, its x and z positions set to 0, and its y position adjusted to match the participant's eye level. The target object was positioned according to a predefined set of poses (position and orientation) per trial. These poses were randomly generated once and remained consistent across all participants, ensuring that they had a clear visual perspective of all faces of the target object during the task. The distance from the user to the target was fixed at 50cm along the z axis, and was constrained to not surpass 50cm for both the x and y axes.

The target was uniformly scaled so that the task difficulty was determined by the target size, which defined the acceptable error margin. Once participants were able to fit the moving object completely inside the target matching the colors on each corresponding face, the moving object would turn green, as illustrated in Fig. 1. This was a visual indication that the object was in the right position. Once the participant released the controller trigger button and the object remained green, the task was considered a success.

4.4 Procedure

Both experiments received approval, as required, from the local Institutional Review Board. Participants went through three phases. In the *pre-study phase*, participants provided verbal consent and completed a background questionnaire, which gathered information on their VR experience, age, gender, area of specialization, and dominant hand. Next, a short presentation was given, explaining the study's objective, the task at hand, and the manipulation techniques in detail. Participants were then introduced to the HWD and provided instructions on adjusting it to fit. They were also guided on using the controller to manipulate the object during the trials.

For the *study phase*, participants were instructed to put on the headset while standing at the center of the delimited area, facing forward. Both objects were then positioned based on the participant's height (see section 4.3). Participants were asked to walk around observing the objects from various angles in order to understand the importance of maneuvering around the objects for the task and to establish a sense of safety in the virtual environment.

Before each condition, participants engaged in a training session to become acquainted with the technique. The participant was instructed by the investigator to freely move the object, followed by executing specific translation and rotation transformations with respect to each axis. The investigator observed the participant's movements to assess their comprehension of the technique's mapping. The participant was considered proficient once they were able to successfully separate degrees of freedom and control their hand speed to precisely move the object. This was further confirmed by the second part of the training, where the participant was asked to place the object inside the target, completing the task.

Following the training session, participants completed the trials for the respective condition. They were informed about the number of trials and the time limit for each trial and were instructed to finish each trial as quickly as possible. Once participants finished all trials for a specific condition, they were instructed to remove the HWD. Subsequently, they were provided with a custom usability questionnaire based on the System Usability Scale (SUS). This questionnaire aimed at providing insights into subjective experiences, such as how easy the techniques were to learn and to use, and included 10 questions with response options ranging from ”Strongly disagree” (1) to ”Strongly agree” (5). Furthermore, participants were asked to complete the unweighted NASA-TLX survey [11] to measure workload, as detailed in Section 5.

In the *post-study phase*, participants completed a survey to rank the manipulation techniques based on the following criteria: 1) personal preference, 2) efficiency, 3) control, 4) naturalness of the technique during rotation, and 5) naturalness of the technique during translation. Following the survey, a brief recorded interview was conducted to gain insight into the reasoning behind their rankings. All equipment was cleaned after each session as a safety measure against COVID-19.

5 EXPERIMENT 1: STANDARD PRECISION TASKS

Previous studies have demonstrated that techniques that offer direct manipulation with exact mapping, such as SVH, are commonly perceived as more natural and efficient for coarse manipulation tasks [6]. However, when the task requires a certain level of precision, scaled manipulation tends to outperform SVH, as the latter is usually no longer sufficient, mainly due to hand instability. Since we aimed to develop an adaptive technique that achieved the right balance between accuracy and naturalness, our first experiment addressed the following research question:

RQ1. How does scaled mapping influence user performance and perceived usability in tasks where high precision is not a requirement?

We were interested in whether AMP-IT could deliver an efficient and similarly natural 6DoF experience as SVH, considering that SVH is expected to be preferred and perceived as more natural, while AMP-IT may offer superior efficiency and control. We had the following hypotheses for RQ1:

H1.1 The AMP-IT technique will not negatively impact the user's performance compared to SVH for tasks that do not require high precision.

H1.2 The perceived usability and workload will not be significantly worse for AMP-IT compared to SVH.

---

[3] https://www.meta.com/quest/products/quest-2/tech-specs

5.1 Experiment design

This experiment employed a within-subjects experimental design to investigate the effects of two independent variables: *technique* (SVH and AMP-IT) and *difficulty* (Very Easy, Easy, and Medium). The dependent variables included both objective and subjective measures. The objective measures were automatically recorded by the system. These measures included the task completion time, the number of completed trials, and the number of clutches (releasing/regrapsing the object during manipulation). The subjective measures were the participant's ratings from the usability questionnaire and the subscales from the NASA-TLX questionnaire, as well as a post-study ranking conditions questionnaire.

For task completion time, we chose not to exclude failed trials, as they could provide insight into the scenario where it might take longer to complete tasks than the allotted time using a specific technique. If we were to consider only successful trials, our analysis could become biased due to potential "lucky trials."

There were nine trials for each of the two techniques, three for each level of difficulty: Very Easy, where the target size was 15% larger than the moving object; Easy (target size 10% larger); and Medium (target size 5% larger). Participants were allotted a time limit of 90 seconds to complete each trial. The selection of target sizes and time limit for the trials were based on pilot runs conducted prior to the study. The order of the techniques was alternated between participants, ensuring that each technique was experienced first an equal number of times. Additionally, the order of the difficulty levels was counterbalanced using a Balanced Latin Square design [29]. A study session lasted between 30 and 60 minutes.

5.2 Participants

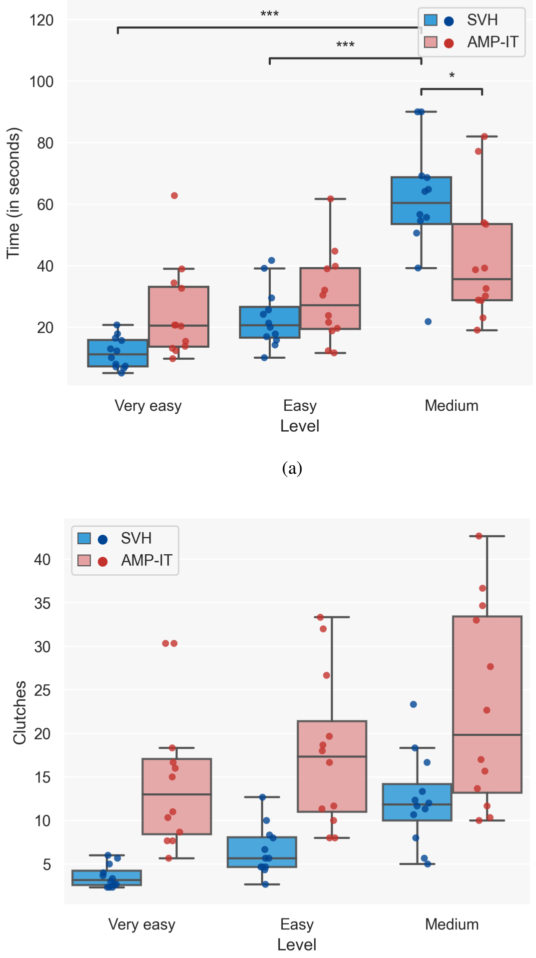
Twelve participants were recruited from academic courses and received volunteer credit for participation. Their ages ranged from 18 to 35 years ($\mu = 22.3, \sigma = 4.56$); nine participants were male and three were female. Two participants reported they had never used VR before, four had used VR 1-3 times, three had used VR 5-10 times, and three had used VR more than ten times. Ten participants were right-hand dominant and two were left-hand dominant.

5.3 Results

After confirming the normality of the data, we conducted two-way ANOVAs to investigate the effects of the independent variables on task completion time and number of clutches. Post-hoc analyses used the Tukey Honestly Significant Difference (HSD) test, allowing pairwise comparisons while controlling for the Type I error rate. All analyses were performed at a 95% confidence level.

The analysis of task completion time revealed main effects for both *difficulty* (F(2, 70) = 31.779, $p < 0.001$) and the interaction between *difficulty* and *technique* (F(2, 70) = 7.184, $p = 0.002$). The average task completion time for the Medium difficulty ($\mu = 51.34, \sigma = 21.49$) was significantly higher compared to Very easy ($\mu = 18.13, \sigma = 12.95$) and Easy ($\mu = 26.33, \sigma = 12.66$) difficulties, both with $p < 0.001$. For the interaction effect, AMP-IT was significantly faster on average ($\mu = 42.23, \sigma = 20.44$) compared to SVH ($\mu = 60.45, \sigma = 19.15$) for tasks with a Medium difficulty level ($p = 0.045$). Additionally, SVH showed significant differences on average task completion times between difficulty levels (both $p = 0.001$), with the Very easy ($\mu = 11.69, \sigma = 5.09$) and Easy ($\mu = 23.03, \sigma = 9.69$) difficulties being significantly faster than the Medium difficulty (see Fig. 3a).

For clutches, there was a significant main effect of *difficulty* (F(2, 70) = 8.39, $p < 0.001$). The number of clutches was significantly higher for Medium difficulty ($\mu = 17.67, \sigma = 10.3$) compared to both Very Easy ($\mu = 9.19, \sigma = 8.16$) with $p = 0.001$, and Easy ($\mu = 12.16, \sigma = 8.64$) with $p = 0.02$. There was a significant main effect of *technique* (F(1, 71) = 41.641, $p < 0.001$), demonstrating that AMP-IT had significantly more clutches ($\mu = 18.54, \sigma = 9.99$) than SVH ($\mu = 7.48, \sigma = 5.02$) (see Fig. 3b).



(a)

(b)

Figure 3: Performance measures in Experiment 1. Significantly different pairs are marked with * when $p \leq 0.05$, ** when $p \leq 0.01$ and *** when $p \leq 0.001$. (a) Task completion time. (b) Number of clutches.

A non-parametric Mann-Whitney U Test was conducted on task completion rate, which showed a significant difference ($U = 29.5, p = 0.008$) between techniques. AMP-IT had a higher task completion rate ($\mu = 8.75, \sigma = 0.62$) than SVH ($\mu = 7.75, \sigma = 1.05$). We note that this average difference is specific to the Medium difficulty level, as all trials were completed at the other levels. Additionally, out of the 12 participants, ten successfully completed all trials within the allotted time using AMP-IT, compared to only three for SVH.

To assess perceived usability and workload, we conducted Mann-Whitney U Tests for each question on the usability questionnaire and on the NASA-TLX subscales. Analysis of the questionnaire revealed a significant difference between the techniques on the *Effectiveness* score ($p = 0.023$). AMP-IT was perceived as more effective ($\tilde{x} = 5$, IQR $= 0.25$) than SVH ($\tilde{x} = 4$, IQR $= 2$). For the NASA-TLX, we found a significant difference in *Performance* ($p = 0.049$), with AMP-IT having better perceived performance ($\tilde{x} = 1.5$, IQR $= 1.25$) than SVH ($\tilde{x} = 2.5$, IQR $= 2$). Although no other statistical significance was reached, AMP-IT scored higher in terms of control, confidence, effectiveness, and fun, while also being perceived as less awkward and complex. SVH received higher scores for ease of use, expected behavior of translation, and ease of learning. The post-study ranking questionnaire revealed a higher preference for AMP-IT (58.3%) compared to SVH. AMP-IT was also considered more efficient (66.6%) and controllable (66.6%), while SVH was ranked higher regarding "Natural for translation" and "Natural for rotation" with 66.6% and 75%, respectively.

5.4 Discussion

We hypothesized that user performance would be similar between SVH and AMP-IT in tasks that do not require high precision (*H1.1*). Analyzing task completion time, the number of clutches, and the number of completed trials, we found partial support for *H1.1*.

Although we can observe that SVH was faster than AMP-IT for Very Easy and Easy tasks, this difference was not statistically significant. However, as the precision level slightly increased, AMP-IT showed significantly better performance (see Fig. 3a). This suggests that the implemented scaled mapping strikes a good balance between precision and performance, enabling users to complete very easy tasks in reasonable time while being significantly more efficient in precise tasks. Additionally, the significance observed between difficulties for SVH indicates that, as expected, increasing the task difficulty leads to longer completion times with SVH. Although the same trend was observed for AMP-IT, no significance was found to confirm that, suggesting that the effect of difficulty on the completion time may vary depending on the technique used.

The number of clutches proved to be significantly higher for AMP-IT compared to SVH, likely due to the fact that the N:1 mapping causes the movement of AMP-IT to be totally dependent on the hand speed. SVH's sensitivity to hand movement allows users to simply grab and hold the object when placing it, as repeatedly releasing and grabbing it does not make a substantial difference. In contrast, AMP-IT benefits from incremental positioning, enabling users to make adjustments during each clutch for precise alignment.

These findings further demonstrate the suitability of AMP-IT for tasks with standard precision requirements. Even when a slight increase in precision is needed, AMP-IT significantly outperforms SVH, enabling users to successfully complete all tasks.

Hypothesis *H1.2* aimed to address concerns related to usability and workload. We expected both techniques to be perceived similarly in terms of subjective measures. Our results, based on the NASA-TLX survey, usability questionnaire, and ranking, partially support *H1.2*, with a tendency for AMP-IT to outperform SVH. Surprisingly, the significance observed in performance ratings for the NASA-TLX survey indicated that users perceived themselves as more successful when using AMP-IT. This could be attributed to the higher number of completed trials achieved with this technique.

As anticipated, SVH received higher scores for the naturalness of translation and rotation. However, participants noted that AMP-IT was more efficient and offered better object control, even during simpler tasks. Overall, 7 out of 12 participants preferred AMP-IT. This aligns with the perception of AMP-IT being more effective than SVH across all trials and difficulties, demonstrating the suitability of scaled mapping for tasks with varying precision. Finally, the absence of a difference between the techniques in terms of expected rotation behavior indicates the positive reception of the novel rotation implementation in AMP-IT.

6 EXPERIMENT 2: HIGH PRECISION TASKS

After assessing the suitability of AMP-IT for typical manipulation tasks, the second experiment focused on very high precision tasks. The goal was to compare the state-of-the-art approach against our two techniques, as well as compare the two techniques against each other. We had two research questions, as follows:

RQ2. How effective is the use of local coordinates and decomposed rotation, as in AMP-IT, compared to using world space coordinates and a single component for rotation, as in PRISM?

We wanted to investigate how AMP-IT's proposed improvements (section 3) influence user performance, perceived usability, and workload when compared to PRISM. With that in mind, we had the following hypotheses for RQ2:

H2.1 AMP-IT will outperform PRISM in tasks that demand a high level of precision. Furthermore, users will prefer it, find it more efficient, and have better control of the object.

H2.2 Scaling the movement in the local space of the object will result in a more natural translation experience in AMP-IT compared to PRISM.

H2.3 Decomposing the rotation into Pitch, Yaw, and Roll components within the local space of the object will result in a more natural rotation experience in AMP-IT compared to PRISM.

H2.4 AMP-IT will have superior perceived usability and lower workload compared to PRISM.

RQ3. How effective is the use of direct manipulation with implicit degree of freedom separation, as in AMP-IT, compared to indirect manipulation with explicit DoF separation combined with SVH, as in WISDOM?

Both techniques offer direct scaled manipulation and DoF separation, but in different ways. Based on prior research on the benefits of direct and natural 3D manipulation [5, 18], we anticipated that AMP-IT would be superior. We had the following hypotheses for RQ3:

H3.1 AMP-IT will outperform WISDOM in tasks that demand a high level of precision. Furthermore, users will prefer it, find it more efficient, and have better control of the object.

H3.2 The direct manipulation provided by AMP-IT will result in more natural translation and rotation experiences compared to the indirect manipulation of WISDOM.

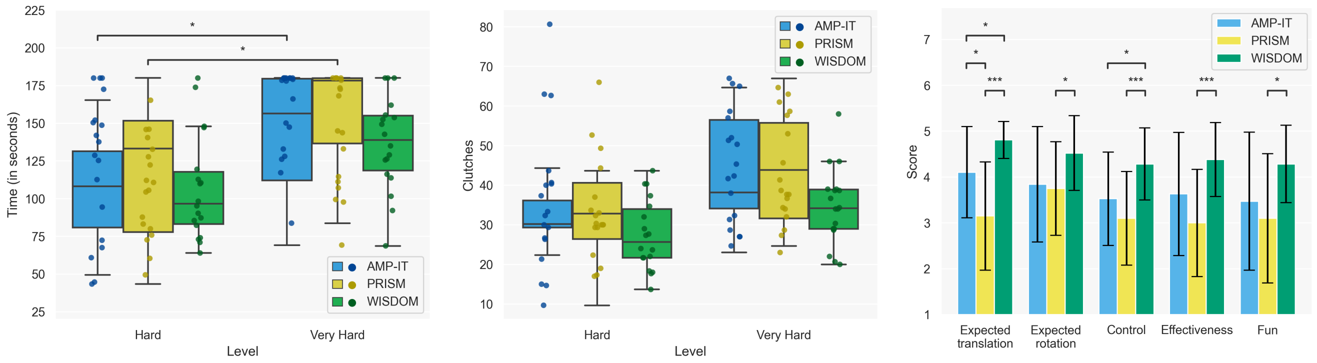
H3.3 AMP-IT will have superior perceived usability and lower workload when compared to WISDOM.

6.1 Experiment design

This experiment employed a within-subjects experimental design to investigate the effects of two independent variables: *technique* (AMP-IT, PRISM, and WISDOM) and *difficulty* (Hard and Very hard). The dependent measures remained the same as in experiment 1. A total of six trials were conducted with each of the three techniques (three for each difficulty level). The difficulty levels were classified as "hard", with a target size that was 1% larger than the moving object, and "very hard", with a target size 0.5% larger. Participants were allotted a time limit of 180 seconds to complete each trial. The selection of target sizes and time limits for the trials were based on pilot runs conducted prior to the study. The order of the difficulty levels was alternated between participants, and the order of the techniques was counterbalanced using a Balanced Latin Square design. The study session lasted about 60-90 minutes.

6.2 Participants

We recruited 18 participants from diverse academic courses, who received volunteer credit for their participation. The ages of the participants ranged from 19 to 33 years ($\mu = 22.63, \sigma = 3.49$); 13 participants were male and 5 were female. Four participants reported they had never used VR before, eight had used VR 1-3 times, four had used VR 5-10 times, and three had used VR more than ten times. Only one participant was left-hand dominant.



(a) (b) (c)

Figure 4: Performance and usability in Experiment 2. Significantly different pairs are marked with * when $p \leq 0.05$, ** when $p \leq 0.01$ and *** when $p \leq 0.001$. (a) Task completion time. (b) Number of clutches. (c) Usability questionnaire scores.

6.3 Results

Two-way ANOVA tests were utilized to examine the main effects of the independent variables on task completion time and the number of clutches per task. Post-hoc pairwise comparisons were performed using the Tukey HSD test.

The analyses revealed that the task completion time differs significantly between the two *difficulties* (F(1, 107) = 25.68, $p < 0.001$) with the Very Hard difficulty ($\mu = 146.78, \sigma = 33.14$) taking longer on average than the Hard one ($\mu = 111.85, \sigma = 38.73$). Both PRISM ($p = 0.034$) and AMP-IT ($p = 0.021$) had significance between difficulty levels (Fig. 4a). No significant main effect was observed for technique, nor the interaction between them.

For number of clutches, we found a main effect for *difficulty* (F(1, 107) = 11.28, $p = 0.001$) and *technique* (F(2, 106) = 4.932, $p = 0.009$). The number of clutches for the Very Hard difficulty ($\mu = 40.82, \sigma = 13.12$) was significantly higher than the Hard difficulty ($\mu = 32.28, \sigma = 14.04$). When comparing techniques, WISDOM ($\mu = 31.03, \sigma = 9.84$) had a significantly lower number of clutches compared to PRISM ($\mu = 40.34, \sigma = 16.76$), with $p < 0.01$, and showed a trend of having fewer clutches than AMP-IT ($\mu = 38.28, \sigma = 13.72$) as well, with $p = 0.056$ (Fig. 4b).

To investigate WISDOM hybrid approach, we analyzed how the movement was distributed between the two manipulation modes. On average, participants covered 93.63% of the total translational distance and 91.12% of the total rotational distance using the SVH mode. However, the SVH mode accounted for only 1.96% of the total manipulation time. The majority of the time was dedicated to using widgets (57.62% for translation and 40.42% for rotation).

In terms of the task completion rate, a Kruskal-Wallis non-parametric test demonstrated a significant effect of *technique*. A Post-hoc Dunn test revealed WISDOM ($\mu = 4.28, \sigma = 1.78$) was only trending to be significantly better than PRISM ($\mu = 2.83, \sigma = 1.68$), with $p = 0.052$. No significance was found between AMP-IT and PRISM; however, we noted that ten participants completed more trials with AMP-IT compared to PRISM, while only three completed more trials with PRISM than AMP-IT. To investigate this further, we conducted an additional analysis taking into account the order in which participants experienced the techniques. Using the Mann-Whitney U Test, we found that participants completed significantly more tasks with AMP-IT when it was experienced after PRISM (U = 69.5, $p = 0.01$). Additionally, we found that 70% of the rotational transformations performed with AMP-IT were using 1DOF.

The Kruskal-Wallis test, followed by post-hoc analysis with the Dunn test, were conducted on each question of the usability questionnaire and subscale of the NASA-TLX. For the *expected behavior of translation*, WISDOM ($\tilde{x} = 5$, IQR $= 0$) showed a significant advantage over both AMP-IT ($\tilde{x} = 4$, IQR $= 1$), with $p = 0.034$, and PRISM ($\tilde{x} = 3$, IQR $= 2$), with $p < 0.001$. Also, AMP-IT was rated significantly higher than PRISM ($p = 0.034$). For the *expected behavior of rotation*, WISDOM ($\tilde{x} = 5$, IQR $= 1$) was rated significantly better than PRISM ($\tilde{x} = 4$, IQR $= 1$), $p = 0.015$. For the *control* question, WISDOM ($\tilde{x} = 4$, IQR $= 1$, $\mu = 4.29$, $\sigma = 0.78$) was significantly better than both PRISM ($\tilde{x} = 3$, IQR $= 2$), $p < 0.001$, and AMP-IT ($\tilde{x} = 4$, IQR $= 1$, $\mu = 3.53$, $\sigma = 1.02$), with $p = 0.043$. For this last case, we included the mean and standard deviation as well, since the median and IQR did not capture the observed significance. WISDOM ($\tilde{x} = 5$, IQR $= 1$) also exhibited a significant advantage over PRISM ($\tilde{x} = 3$, IQR $= 2$) regarding *effectiveness* ($p < 0.001$). Finally, WISDOM ($\tilde{x} = 4$, IQR $= 1$) was also considered more *fun* than PRISM ($\tilde{x} = 3$, IQR $= 2$), ($p = 0.02$).

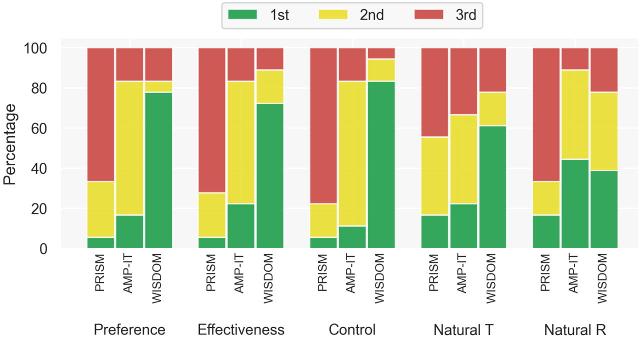


Figure 5: Reported post-study rankings for each of the five questions

No significance was found for any of the NASA-TLX subscales. Although there was a trend for a significant effect on *temporal demand* ($p = 0.66$), a closer look through the Dunn test revealed no significantly different pairs.

We conducted descriptive statistical analysis on the post-study ranking conditions questionnaire, as shown in Fig. 5. Each technique was evaluated based on five aspects: *Preference*, *Efficiency*, *Control*, *Naturalness for translation* (Natural T), and *Naturalness for rotation* (Natural R). WISDOM appeared most often in the first position for overall preference (77.78%), efficiency (72.22%), control (83.33%) and natural for translation (61.11%). AMP-IT was ranked as the most natural for rotation, with 44.44% of the votes. Also, AMP-IT emerged as the second-place performer for most of the rankings. PRISM was most often ranked last for all questions.

6.4 Discussion

In the second experiment, *H2.1* and *H3.1* stated that AMP-IT would demonstrate better performance than both PRISM and WISDOM in terms of task completion time, number of clutches, and task completion rate. However, the results did not support our hypotheses.

Although not statistically significant, our analysis of task completion time revealed a trend suggesting potential superiority of the WISDOM technique compared to the other techniques. This trend can be attributed to the nature of manipulation tasks, which typically involve two distinct phases: an initial phase characterized by fast but imprecise motion towards the target, followed by a final phase of slower and more precise movement to acquire the target [30]. By combining indirect and direct manipulation, WISDOM leverages the benefits of SVH during the initial phase, allowing for faster coarse manipulations through direct and unconstrained movement.

Indeed, our investigation on the use of each mode in WISDOM confirmed that participants covered a significant portion of both translational and rotational distances very quickly using SVH. Subsequently, they dedicated a considerable amount of time to performing precise manipulation through widgets. The lack of statistical significance in the overall time difference suggests the possibility of participants requiring more time during the second phase, particularly when interacting with widgets to achieve precision. This observation aligns with earlier findings which indicate an increase in time during widget usage [19].

The influence of the technique on the number of clutches can also be attributed to the integration of SVH with widgets in WISDOM. As seen in Experiment 1, AMP-IT required a significantly higher number of clutches compared to SVH. By reducing this number in the initial phase of the task, WISDOM may reduce substantially the clutches required for task completion.

In terms of perceived usability, WISDOM also outperformed the other techniques, with significant differences in several categories. Those scores can be explained by two factors: the level of success that the participants had with the technique, which was slightly greater with WISDOM (though not statistically significant), and the fact that it was easy to know what to expect when grabbing a widget. With WISDOM's restricted movement to one DoF, participants found it easy to correct any unintended transforms by simply making the opposite movement with their hand. This was not the case with the other techniques, as making a reverse movement might not bring the object back to its original position due to natural differences in the movement itself and in hand speed.

Comparing the two direct manipulation techniques, a further look into the task completion rates suggests that the improvements in AMP-IT were advantageous, as participants that experienced PRISM before AMP-IT were able to leverage their experience with PRISM to effectively utilize the enhancements offered by AMP-IT. Participant 5 mentioned *"This [AMP-IT] is much more accurate [than PRISM]. It's really precise how it applies my movement ... it's really awesome"*. Furthermore, users actively explored 1DOF rotations in AMP-IT, which demonstrates the advantages of independent manipulation along each axis, supporting hypothesis *H2.3*. When considering subjective measures, AMP-IT received significantly higher scores than PRISM, specifically in terms of the expected behavior of translation. This supports our hypothesis *H2.2* that using the local coordinates of the object improved the responsiveness of translational movement. Also, although no significance was found, AMP-IT showed a trend of outperforming PRISM in all aspects, indicating a positive trend towards hypothesis *H2.4*.

Regarding the post-study ranking conditions questionnaire, WISDOM generally performed better in almost every aspect, except for the perceived naturalness of rotation, in which AMP-IT showed superiority. This aligns with previous research that found that users had difficulty understanding rotation with PRISM and widgets [9,19,20], which supports *H2.3* and partially *H3.2*, indicating that the separation of the rotation in its three axes and the direct manipulation in the local space of the object offered by AMP-IT does show improvements over the other two methods.

Overall, the demonstrated superiority of WISDOM strongly indicates the effectiveness of a hybrid approach over "one-size-fits-all" approaches for highly precise manipulation tasks. Moreover, although maintaining a similar completion rate to the implicit approach, the explicit separation of degrees of freedom coupled with the visual feedback in WISDOM enhanced subjective perceived performance and usability. Significant advantages are achieved by providing optimal object manipulation tailored for each of the two phases of the task instead of relying on the balance offered by scaled mapping alone. However, given the considerable time dedicated to widget interaction exclusively, the indirect manipulation approach may not be the most suitable for the second phase. In such instances, it would be prudent to explore the potential benefits of incorporating scaled mapping through the design of an AMP-IT technique that explicitly incorporates an SVH mode.

7 LIMITATIONS AND FUTURE WORK

Based on our hypothesis that AMP-IT would outperform WISDOM, we did not compare WISDOM against SVH. However, the promising performance of WISDOM for high-precision tasks suggests the need for evaluating the technique in standard-precision scenarios as well. Participants highlighted the lack of visual feedback as one of the biggest limitations of AMP-IT compared to WISDOM. Future research could investigate the impact of visual aids on performance. Also, while AMP-IT effectively filtered unwanted movements on the axes, it still allowed unintentional translations during rotation and vice versa. Future work could focus on methods to identify the intended transform before applying scaling to individual axes. Additional investigation is also necessary to evaluate the performance of the novel techniques when aligning complex 3D meshes. This would provide valuable insights into the effectiveness and usability of the techniques in handling more intricate and detailed objects, enhancing their applicability in various domains and tasks. Last, for all techniques, users suffered with release precision. A further step towards that direction would be to apply automatic release corrections, as suggested by Osawa et al. [23].

8 CONCLUSION

Motivated by the need for extremely precise six DoF manipulation in some XR applications, we introduced two novel techniques aimed at addressing issues with state-of-the-art 3D manipulation techniques.

Interestingly, our findings revealed that although AMP-IT may exhibit slightly worse performance than SVH for very easy tasks, it offers a compelling trade-off between performance and usability. AMP-IT demonstrated superior usability and a workload similar to SVH, making it a suitable choice for a wide range of manipulation tasks, including coarse manipulation tasks.

Furthermore, our second study revealed an unexpected result regarding the effectiveness of WISDOM, which employs indirect manipulation with a switchable direct manipulation mode, compared to both AMP-IT and PRISM. Surprisingly, WISDOM showed a trend to perform better than the other techniques in tasks that required a high level of precision. The ability to employ direct manipulation and subsequently fine-tune the position with precise one DoF control proved to be promising. Also, AMP-IT demonstrated its superiority as the most natural technique for rotation, surpassing both PRISM and WISDOM. These findings highlight the potential of both AMP-IT and WISDOM for precise manipulation in VR, and suggest opportunities for future enhancements and analysis.

ACKNOWLEDGMENTS

This study received partial funding from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, and was also supported by the US DOE LLNL-LDRD 23-SI-003, performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 (LLNL-CONF-850723).

REFERENCES

[1] F. Bacim, R. Kopper, and D. A. Bowman. Design and evaluation of 3d selection techniques based on progressive refinement. *International Journal of Human-Computer Studies*, 71(7-8):785–802, 2013.

[2] R. Blanch, Y. Guiard, and M. Beaudouin-Lafon. Semantic pointing: improving target acquisition with control-display ratio adaptation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 519–526, 2004.

[3] D. A. Bowman and L. F. Hodges. Formalizing the design, evaluation, and application of interaction techniques for immersive virtual environments. *Journal of Visual Languages & Computing*, 10(1):37–53, 1999.

[4] D. A. Bowman, R. Kopper, and F. Bacim. Effortless 3d selection through progressive refinement., 2019.

[5] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison Wesley Longman Publishing Co., Inc., USA, 2004.

[6] D. A. Bowman, R. P. McMahan, and E. D. Ragan. Questioning naturalism in 3d user interfaces. *Commun. ACM*, 55(9):78–88, sep 2012. doi: 10.1145/2330667.2330687

[7] L. Dominjon, A. Lécuyer, J.-M. Burkhardt, P. Richard, and S. Richir. Influence of control/display ratio on the perception of mass of manipulated objects in virtual environments. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pp. 19–25. IEEE, 2005.

[8] J. Feng, I. Cho, and Z. Wartell. Comparison of device-based, one and two-handed 7dof manipulation techniques. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*, SUI '15, p. 2–9. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2788940.2788942

[9] S. Frees, G. D. Kessler, and E. Kay. Prism interaction for enhancing control in immersive virtual environments. *ACM Trans. Comput.-Hum. Interact.*, 14(1):2–es, may 2007. doi: 10.1145/1229855.1229857

[10] P. C. Gloumeau, W. Stuerzlinger, and J. Han. Pinnpivot: Object manipulation using pins in immersive virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 27(4):2488–2494, 2021. doi: 10.1109/TVCG.2020.2987834

[11] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988.

[12] 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*, UIST '19, p. 1185–1195. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3332165.3347916

[13] D. Hoag. Apollo guidance and navigation: Considerations of apollo imu gimbal lock. *Canbridge: MIT Instrumentation Laboratory*, pp. 1–64, 1963.

[14] P. Klacansky, H. Miao, A. Gyulassy, A. Townsend, K. Champley, J. Tringe, V. Pascucci, and P.-T. Bremer. Virtual inspection of additively manufactured parts. In *2022 IEEE 15th Pacific Visualization Symposium (PacificVis)*, pp. 81–90. IEEE, 2022.

[15] W. A. König, J. Gerken, S. Dierdorf, and H. Reiterer. Adaptive pointing–design and evaluation of a precision enhancing technique for absolute pointing devices. In *Human-Computer Interaction–INTERACT 2009: 12th IFIP TC 13 International Conference, Uppsala, Sweden, August 24-28, 2009, Proceedings, Part I 12*, pp. 658–671. Springer, 2009.

[16] 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.

[17] D. P. Mapes and J. M. Moshell. A two-handed interface for object manipulation in virtual environments. *Presence: Teleoper. Virtual Environ.*, 4(4):403–416, jan 1995. doi: 10.1162/pres.1995.4.4.403

[18] 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*, vol. 38, pp. 21–45. Wiley Online Library, 2019.

[19] D. Mendes, F. Relvas, A. Ferreira, and J. Jorge. The benefits of dof separation in mid-air 3d object manipulation. In *Proceedings of the 22nd ACM conference on virtual reality software and technology*, pp. 261–268, 2016.

[20] D. Mendes, M. Sousa, R. Lorena, A. Ferreira, and J. Jorge. Using custom transformation axes for mid-air manipulation of 3d virtual objects. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pp. 1–8, 2017.

[21] M. R. Mine, F. P. Brooks, and C. H. Sequin. *Moving Objects in Space: Exploiting Proprioception in Virtual-Environment Interaction*, p. 19–26. ACM Press/Addison-Wesley Publishing Co., USA, 1997.

[22] B. F. Morrey. *The elbow and its disorders*. Elsevier Health Sciences, 2009.

[23] N. Osawa. Automatic adjustments for efficient and precise positioning and release of virtual objects. In *Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications*, pp. 121–128, 2006.

[24] J. S. Pierce, B. C. Stearns, and R. Pausch. Voodoo dolls: seamless interaction at multiple scales in virtual environments. In *Proceedings of the 1999 symposium on Interactive 3D graphics*, pp. 141–145, 1999.

[25] M. S. Pinho, D. A. Bowman, and C. M. Freitas. Cooperative object manipulation in immersive virtual environments: Framework and techniques. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '02, p. 171–178. Association for Computing Machinery, New York, NY, USA, 2002. doi: 10.1145/585740.585769

[26] 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*, CHI '12, p. 1297–1306. Association for Computing Machinery, New York, NY, USA, 2012. doi: 10.1145/2207676.2208585

[27] 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*, pp. 265–272, 1995.

[28] C. Wilkes and D. A. Bowman. Advantages of velocity-based scaling for distant 3d manipulation. In *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology*, VRST '08, p. 23–29. Association for Computing Machinery, New York, NY, USA, 2008. doi: 10.1145/1450579.1450585

[29] E. J. Williams. Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry*, 2(2):149–168, 1949.

[30] R. S. Woodworth. Accuracy of voluntary movement. *The Psychological Review: Monograph Supplements*, 3(3):i, 1899.

[31] R. Yu and D. A. Bowman. Force push: Exploring expressive gesture-to-force mappings for remote object manipulation in virtual reality. *Frontiers in ICT*, 5:25, 2018.