Move'n'Hold Pro: Consistent Spatial Interaction Techniques for Object Manipulation with Handheld and Head-mounted Displays in Extended Reality

Vera M. Memmesheimer
v.memmesheimer@rptu.de
Human Computer Interaction Lab,
RPTU Kaiserslautern-Landau
Kaiserslautern, Germany

Kai J. Klingshirn
klingshi@rptu.de
Human Computer Interaction Lab,
RPTU Kaiserslautern-Landau
Kaiserslautern, Germany

Cindy Herold
herold@rptu.de
Human Computer Interaction Lab,
RPTU Kaiserslautern-Landau
Kaiserslautern, Germany

Bahram Ravani
bravani@ucdavis.edu
Department of Mechanical and Aerospace Engineering, University of California, Davis
Davis, CA, USA

Achim Ebert
achim.ebert@rptu.de
Human Computer Interaction Lab,
RPTU Kaiserslautern-Landau
Kaiserslautern, Germany

Figure 1: *Move'n'Hold Pro* provides direct and continuous object translation or rotation for MR-HHDs, MR-HMDs, and VR-HMDs; for example: (a) continuous translation using left- and right-thumb-touch with MR-HHD; (b) object selection by moving the red selection point to the manipulable object with MR-HMD; (c) rotation through direct mapping using left-thumb-touch with VR-HMD.

ABSTRACT

Extended Reality (XR) technologies are still lacking appropriate interaction methods that enable users to seamlessly switch between different XR devices and degrees of virtuality. Addressing this gap, we present *Move'n'Hold Pro* – a set of consistent object manipulation techniques that are available for Mixed Reality handheld displays (MR-HHDs) as well as for Mixed and Virtual Reality head-mounted displays (MR-/VR-HMDs). *Move'n'Hold Pro* extends MR- and VR-HMDs with a tablet controller that implements object manipulation methods proposed by latest research on MR-HHD-UIs. Thereby, users can combine tablet movement and peripheral touch to translate or rotate virtual objects through direct or continuous manipulations. In our evaluation, comparing *Move'n'Hold Pro* to a State of the Art system, *Move'n'Hold Pro* was rated as the preferred system and to be easier to relearn. Furthermore, *Move'n'Hold Pro* reduced cognitive efforts, improved usability, and provided more cross-device benefits.

CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality**; **Virtual reality**.

KEYWORDS

Augmented Reality, Mixed Reality, Virtual Reality, Extended Reality, Head-mounted Displays, Handheld displays, Spatial Interaction, Interaction Technique, User Interface

ACM Reference Format:
Vera M. Memmesheimer, Kai J. Klingshirn, Cindy Herold, Bahram Ravani, and Achim Ebert. 2024. Move'n'Hold Pro: Consistent Spatial Interaction

ECCE 2024, October 08–11, 2024, Paris, France

ACM ISBN 979-8-4007-1824-3/24/10
https://doi.org/10.1145/3673805.3673814

Techniques for Object Manipulation with Handheld and Head-mounted Displays in Extended Reality. In *European Conference on Cognitive Ergonomics (ECCE 2024), October 08–11, 2024, Paris, France.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3673805.3673814

1 INTRODUCTION

Extended Reality (XR) environments, encompassing different degrees of virtuality (i.e., Mixed and Virtual Reality (MR/VR)) as well as different devices (i.e., handheld and head-mounted displays (HHDs/HMDs)), have the potential to transform and improve entire working environments by relocating them to 3D. Yet, despite great enthusiasm, the widespread application of XR in real-world settings fails to materialize. We consider the lack of suitable interaction methods as one of the main obstacles to XR's broader acceptability. While previous research has focused extensively on solving usability issues of interaction techniques for single access points of XR (MR-HHDs, MR-HMDs, or VR-HMDs), little research has been conducted on interaction paradigms that are applicable across different devices and degrees of virtuality. However, seamless transitions between these access points are expected to be an essential requirement in real-world applications. We believe that effective spatial interaction requires a completely new interaction paradigm rather than transferring well-established 2D interaction paradigms to XR. Thereby, the initial effort required to learn a new interaction technique is likely to be compensated over time if the new technique is tailored specifically for spatial interaction. When all access points are equipped with the same interaction paradigm, we further expect users to benefit from learning effects while switching between different access points. Addressing this gap, we propose *Move'n'Hold Pro* – an entire set of consistent interaction techniques for object translation and rotation with MR-HHDs, MR-HMDs, and VR-HMDs. *Move'n'Hold Pro* extends an advanced interaction method for MR-HHDs called *Move'n'Hold* [11] that allows seamless combinations of direct and continuous object manipulation solely through device movement and peripheral touch. To this end, we revisit the design of *Move'n'Hold*, transfer it to a MR-HMD and a VR-HMD using tablet controllers, and compare *Move'n'Hold Pro* to a State of the Art system in a detailed user study.

2 BACKGROUND AND RELATED WORK

Previous research has outlined XR's wide range of applications. For instance, VR can provide detailed training environments, enable virtual prototyping, or remote collaboration. In MR, virtual augmentations can extend physically existing parts of a prototype or support communication between a local worker and a remote expert. Depending on the use case, either HHDs or HMDs may be favored. For instance as mentioned by Zhu and Grossman [20], HHDs are highly accessible due to their ubiquity and precise input options on high-resolution displays while they provide limited spatial in- and output options. HMDs enable spatial in- and output but lack well-known and precise interaction methods. Moreover, HMDs are less ubiquitous than HHDs and not yet qualified for industrial environments.

Well-established touch-based interaction techniques are inappropriate for MR-HHDs as they are prone to occlusion and fatigue when the HHD is held up with one hand while the other hand performs touch gestures. A detailed requirements analysis for MR-HHD-UIs can be found in [11]. As an alternative, device-based object manipulation techniques have been proposed which map the HHD's movement to objects (e.g., [10, 15]). While device-based manipulation allows holding the device with both hands, it is limited by arm and wrist movement restrictions. Addressing these limitations, Memmesheimer et al. [11] proposed *Move'n'Hold* – a novel interaction paradigm for object translation and rotation with MR-HHDs which offers an option for automated continuous object manipulation. Hence, it allows users to intuitively combine large and coarse movements with small and precise movements.

Out-of-the-box interaction techniques for HMDs such as mid-air gestures and controllers are often physically demanding, imprecise, or rely on external tracking. In this context, the integration of mobile devices like phones or tablets was deemed promising due to their ubiquity and precise input options such as multi-touch or device movement captured through IMU sensors. Furthermore, they can be employed as secondary displays. Previous research has explored the integration of smartphones (e.g., [6, 7, 14, 16, 18, 20]) and tablets (e.g., [5, 8, 9, 17]) with VR-HMDs (e.g., [6, 8, 16, 17, 19]) as well as with MR-HMDs (e.g., [5, 7, 9, 14, 18, 20]) for various purposes. In this paper, we focus on the integration of tablets to enhance object manipulation. Knierim et al. [7] combined a MR-HMD and a phone which can be used to perform object manipulations through different touch gestures. Whenever the user begins to translate an object, a reference coordinate system is set up. Translations along the x- or z-axis can be performed by single taps followed by a horizontal or vertical swipe while a double tap followed by a swipe enables y-axis translations. Rotations around the y-axis can be performed with multi-touch (i.e., two finger rotations on the screen). In Luo et al.'s [9] approach, a virtual object seen though a MR-HMD is attached to a tablet. The tablet's movement in space is mapped to the object such that its position and orientation are updated at the same time. Another approach of a MR-HMD extended with a phone was proposed by Unlu and Xiao [18]. In their application, the phone acts as a 6DOF input device as well as a 3D trackpad. To this end, a virtual plane is attached to the phone. Virtual objects that are located on this plane can be moved on the plane through touch gestures on the phone's display. In the application proposed by Kari and Holz [6], a phone connected to a VR-HMD can be used to adjust the position and orientation of two virtual hands that are located on a plane. The plane is generated according to the phone's position and orientation and constantly adapted based on the phone's movement. The virtual hands can then be controlled through touch input via the thumbs on the phone's left and right display side. Clutch during touch input and phone movement amplifications for computing the plane position enable large hand movements. Users can grasp objects by applying touch after hovering them with the virtual hand and place the objects by releasing touch. Further approaches involving VR-HMDs have been proposed by Surale et al. [17] who enabled object manipulation in VR through a tablet as well as by Zhao et al. [19] who extended the VR-HMD with a custom, tablet-like device that is able to provide haptic feedback. In both approaches, the tablet is visualized in the VR scene. In [19], the complete VR scene is rendered on the virtual tablet and frozen upon the initiation of object manipulation. As such, the user can select objects by touching them on the tablet's screen and then translate and rotate them through touch gestures. In [17], the tablet serves as a viewport. To select an object, users can hit the desired object with one corner of the tablet or move the tablet in front of the desired object and tap on the screen. Selected objects can be manipulated with touch gestures whereby different tablet orientations can be used to fix different axes.

While these interaction paradigms presented in previous research were deemed helpful in the context of HMDs, their design is not directly transferable to object manipulation with MR-HHDs. One-handed touch input is not suitable for MR-HHDs as it is prone to fatigue, occlusion issues, and tends to become tedious during complex multidimensional manipulations. Furthermore, the integration of continuous movement can reduce clutching during large manipulations. Hence, a consistent interaction paradigm that allows switching between XR's main access points without having to spend a lot of temporal and cognitive efforts to adapt to the new setup is still missing.

3 DESIGNING MOVE'N'HOLD PRO

To address the gap identified in Section 2, we extend the interaction paradigm *Move'n'Hold* [11] for HMDs and present *Move'n'Hold Pro* – a novel, unvarying set of object translation and rotation techniques for XR's main access points: MR-HHDs, MR-HMDs, and VR-HMDs. To this end, we divide the translation and rotation process into the following steps: First, the user (1) specifies the object to be manipulated, then (2) confirms the selection, and finally (3) translates or rotates the object.

Using *Move'n'Hold* for MR-HHDs as described in [11], (1) the desired object is specified by centering it in the HHD's screen (i.e., it turns green), (2) selection is then confirmed by left-thumb-touch, and (3) the object is manipulated by translating or rotating the HHD. As long as left-thumb-touch is active, the HHD's movement (i.e., its translation or rotation) is directly mapped to the object. Since such manipulations affect physical effort and are limited by arm and wrist movement restrictions, *Move'n'Hold* offers continuous movement: After some initial object manipulation (1+2+3), the user can (4) add right-thumb-touch to translate or rotate the object automatically (i.e., without moving the HHD) in the direction specified in step (3). This movement continues until right-thumb-touch is released, whereby the speed of movement can be set by the length of the initial movement. *Move'n'Hold Pro* implements *Move'n'Hold* for MR-HHDs as in [11] with two exceptions: We add a red selection point to the screen's center to support object selection and do not implement axis locking as Memmesheimer et al. [11] found that it was less helpful than expected.

Thus, *Move'n'Hold* involves two senses: vision (i.e., centering the desired object in the field of view) and touch (i.e., translating or rotating the device while applying peripheral touch). To transfer *Move'n'Hold* as seamlessly as possible to HMDs, we add the same red selection point to the HMD's scene and integrate a tablet controller. The desired object can then be (1) specified through vision (i.e., adjusting the field of view to move the red point) and manipulated (2+3+4) by combining touch and tablet movements. While in the MR-HHD setting (Fig. 1a) the tablet handles input and output and thus has to be held up high, it only handles input and can be held in any position and orientation in the HMD settings (Fig. 1b+c, 2a+b). Following [11, 12], we separate translation and rotation methods and did not implement axis locking.

4 IMPLEMENTATION

To compare *Move'n'Hold Pro* to a State of the Art system, we implemented translation and rotation methods for a MR-HHD (Apple iPad Pro, 11inch 3rd Gen.), a MR-HMD (Microsoft HoloLens 2), and a VR-HMD (HTC VIVE PRO). The 12 interaction techniques were developed in Unity using ARKit, Mixed Reality Toolkit, and XR Interaction Toolkit. An overview of the interaction methods provided by both systems is given in Table 1.

4.1 Move'n'Hold Pro Interaction Methods

Move'n'Hold Pro extends *Move'n'Hold* from [11] and transfers it to MR- and VR-HMDs as described in Section 3.

4.1.1 MR-HHD. *Move'n'Hold Pro* for MR-HHDs was implemented as described in [11]. We check for collisions between a ray shot from the device camera's center (i.e., the red selection point) and a manipulable object as long as no touch input is detected. If hit by the ray, the object turns green and the selection point disappears. Upon left-thumb-touch, we store the device's current position v_{ltt} and orientation q_{ltt} in space. While only left-thumb-touch is active, we then translate / rotate the object through direct mapping. Thus, we add the vector v_{move} / multiply the quaternion q_{move} that describes the device's translation / rotation in the last frame to the object's current position v_{obj} / by the object's current orientation q_{obj} (Eq. 1/2). When touch is added on the right display side, we compute v_{hold} / q_{hold} that describe the device's translation / rotation from v_{ltt} / q_{ltt} to the device's current position v_{rtt} / orientation q_{rtt} (Eq. 3/4). For continuous object manipulation, we then add v_{hold} / multiply q_{hold} to the object's current position v_{obj} / by the object's current orientation q_{obj} and perform a linear interpolation with *0.1* like in [11]. The object's position / orientation is then updated to the interpolated value v_{lerp} / q_{lerp} (Eq. 5/6). While left- and right-thumb-touch remain active, the object moves automatically by updating its position / orientation to v_{lerp} / q_{lerp} in every frame. When right-thumb-touch is released, the object stops moving automatically and the user can resume manipulation with direct mapping.

4.1.2 MR-HMD / VR-HMD. *Move'n'Hold Pro* for MR- and VR-HMDs consists of two apps running on an Apple iPad Pro (11inch, 4th Gen.) and the HMD (i.e., Microsoft HoloLens 2 or HTC VIVE PRO) which communicate through UnityWebRequests. To ensure that the tablet's and HMD's coordinate systems have the same orientation, both devices are started in a fixed position and orientation. Upon launch, the tablet app asks the HMD app to share the active manipulation mode (translation/rotation). Similar to the MR-HHD app, a red selection point is located in the center of the HMD's camera. As long as no touch is detected on the tablet, the HMD checks for collisions between a ray originating from the HMD (i.e., the red selection point) and manipulable objects. When a collision is detected, the selection point disappears and the object turns green. Upon touch detection, the tablet asks the HMD for currently selected objects. If an object was hit by the HMD's head-ray when left-thumb-touch was applied, the tablet starts to

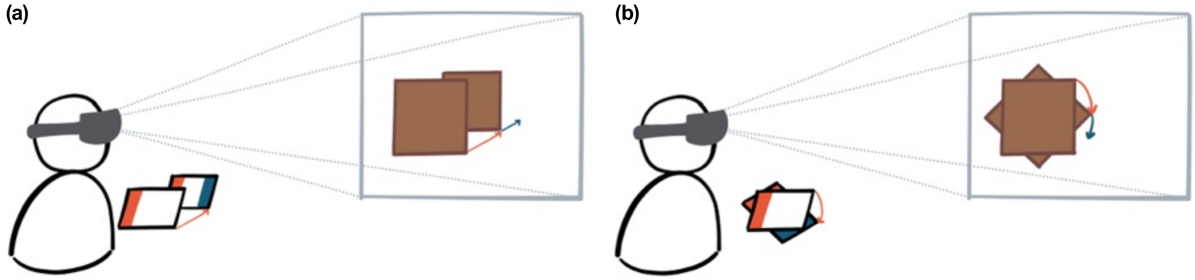

Figure 2: *Move'n'Hold Pro* for (a) translation and (b) rotation: When only left-thumb-touch is active (orange) the tablet's movement is directly mapped to the object. Adding right-thumb-touch (blue) starts automated object movement (i.e., without tablet movements).

cache its movement: v_{move}, q_{move}, and *hold* (a boolean which stores if right-thumb-touch is registered). If *hold* is true, v_{hold} / q_{hold} are calculated and cached as well. In every frame, the collected data is sent to the HMD. The HMD constantly updates internal variables storing *hold*, v_{hold} / q_{hold}, and adds v_{move} / q_{move} to two lists. If only left-thumb-touch is active, the selected object is manipulated (Eq. 1/2) by processing the list of vectors v_{move} / quaternions q_{move}. If left- and right-thumb-touch are active, the object's new position v_{lerp} / orientation q_{lerp} is computed by linear interpolations (Eq. 5/6) similar to the MR-HHD app described in Section 4.1.1. If left- and right-thumb-touch are released, the tablet informs the HMD accordingly, selection with the HMD's head-ray is activated again, object manipulation stops immediately, and the lists of translation vectors and rotation quaternions in the HMD app are cleared.

$$v_{\text{objNewDM}} = v_{\text{move}} + v_{\text{obj}} \quad (1)$$

$$q_{\text{objNewDM}} = q_{\text{move}} * q_{\text{obj}} \quad (2)$$

$$v_{\text{hold}} = v_{\text{rtt}} - v_{\text{ltt}} \quad (3)$$

$$q_{\text{hold}} = q_{\text{rtt}} * q_{\text{ltt}}^{-1} \quad (4)$$

$$v_{\text{objNewCM}} = v_{\text{lerp}} = Vector3.Lerp(v_{\text{obj}}, v_{\text{hold}} + v_{\text{obj}}, 0.1) \quad (5)$$

$$q_{\text{objNewCM}} = q_{\text{lerp}} = Quaternion.Lerp(q_{\text{obj}}, q_{\text{hold}} * q_{\text{obj}}, 0.1) \quad (6)$$

4.2 State of the Art Interaction Methods

For our State of the Art system (*SotA*) we chose to implement out-of-the-box techniques based on touch, gestures, and controllers which are most common in practical use. Similar to our implementation of the *Move'n'Hold Pro* interaction methods, we separate translation and rotation.

4.2.1 MR-HHD. Objects are selected and turn green via touch. The selected object can then be manipulated by dragging the finger on the screen. During translation tasks, the object is moved according to the finger's movement on an invisible plane which is parallel to tablet's orientation. During rotation tasks, the object is surrounded by an invisible sphere. By moving the finger on the screen the sphere is rotated around its center and the same rotation is applied to the object.

4.2.2 MR-HMD. Objects are manipulated with the standard mid-air gestures for Microsoft HoloLens 2 from the Mixed Reality Toolkit. The user can point at objects and perform a pinch gesture once a ray originating from the hand collides with the object. While performing the pinch gesture, the object follows the hand's translation or rotation.

4.2.3 VR-HMD. Object translations and rotations can be performed with the standard HTC VIVE PRO controller. We extract controller input, its position and orientation such that the user can select objects by pressing the controller's trigger button when a ray originating from the controller collides with the object. The controller's translation or rotation in space is then mapped to the object as long as the trigger button remains pressed.

5 USER STUDY

5.1 Experimental Design and Procedure

We compared *Move'n'Hold Pro* (Fig. 3a–c) to *SotA* (Fig. 3d–f) in a user study with 20 participants (10 male / 10 female; 20-37y/o). Some had previous experience with MR-HHDs (70%), MR-HMDs (40%), and VR-HMDs (60%). We considered three independent variables: system (*Move'n'Hold Pro*, *SotA*), device (MR-HHD, MR-HMD, VR-HMD), and task (translation, rotation), yielding 12 sessions (i.e., $system \times device \times task$). Half of the participants performed all tasks with *SotA* prior to *Move'n'Hold Pro* and vice versa. Using both systems, translation and rotation tasks were first completed with the MR-HHD. To investigate cross-device learnability, half of the participants continued with the VR-HMD and used the MR-HMD last and the other half used the HMDs in reverse order. As recommended in [11], translation tasks were performed prior to rotation tasks. The participants watched an explanatory video and performed a short training session before each session and provided difficulty ratings (DIFF_exp, DIFF_large) afterwards. After task completion with all devices using the first system, the system was assessed with the System Usability Scale (SUS) [2] and NASA TLX [13]. This procedure was repeated for the second system followed by final questions about both systems.

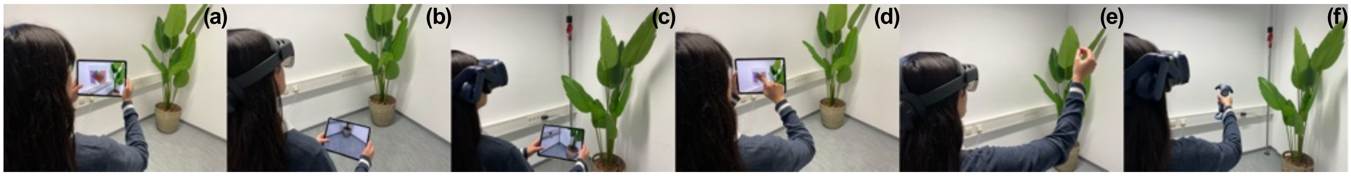

Figure 3: (a-c) *Move'n'Hold Pro* and (d-f) *SotA* for (a+d) MR-HHDs, (b+e) MR-HMDs, and (c+f) VR-HMDs.

	Move'n'Hold Pro *MR-HHD, MR-HMD, VR-HMD*	***SotA*** *MR-HHD*	*MR-HMD*	*VR-HMD*
selection	adjust the position or orientation of the HHD / HMD such that the red selection point in the display's center hits the object	touch the object on the HHD's screen	point at the object with the finger (hand gesture)	point at the object with the controller
translation or rotation	the tablet's translation / rotation is mapped to the object while left-thumb-touch is applied; continuous translation / rotation can be started by adding right-thumb-touch	the object is translated / rotated by dragging the finger on the screen	the object is translated / rotated based on hand movements while a pinch gesture is performed	the object is translated / rotated based on controller movements while the trigger button is pressed

Table 1: Selection, translation, and rotation methods as provided in *Move'n'Hold Pro* and *SotA*.

5.2 Tasks

In each of the 12 sessions, each participant completed 8 translation or 8 rotation tasks with a MR-HHD, MR-HMD, or a VR-HMD using *Move'n'Hold Pro* or *SotA* (i.e., 96 object manipulation tasks per participant). During translation tasks an opaque manipulable cube ($0.2m \times 0.2m \times 0.2m$) had to be moved into a semi-transparent target cube ($0.25m \times 0.25m \times 0.25m$) which was placed in the scene's center. When the manipulable cube was inside the target cube, the manipulable cube disappeared and the next manipulable cube appeared at a new position. We computed the initial positions of the manipulable cubes as follows: Starting from the target cube's position, the manipulable cube was translated 0.75m along the x-, y-, and z-axis either in the positive or negative direction, yielding eight different positions. During rotation tasks an opaque box ($0.1m \times 0.15m \times 0.1m$) was placed inside a semi-transparent box ($0.2m \times 0.3m \times 0.2m$) in the scene's center. Participants had to rotate the inner box such that the differently colored sides of the inner and outer box matched. When the inner box's orientation differed less than 4 degrees on each axis from the outer box, the next box appeared. Similar to translation tasks, the initial orientations were set by 40 deg rotations around the x-, y-, and z-axis either in the positive or negative direction, yielding 8 different orientations. To ensure the comparability of task completion times, a simple task (i.e., a 0.75m translation along / 40 deg rotation around the x-axis) had to be completed before the first task started. Throughout the experiment, the participants could move around in a limited area ($3m \times 3m$).

5.3 Hypotheses and Measuring Instruments

We compared *SotA* and *Move'n'Hold Pro* based on hypotheses H1-H9 with respect to temporal and cognitive effort, usability, and user preferences with the following measuring instruments. We measured task completion times (TCTs) (i.e., the time span between the appearance and disappearance of each virtual box). For each interaction technique we collected two difficulty ratings DIFF_exp and DIFF_large from 1 (very easy) to 5 (very difficult). Thereby, DIFF_exp corresponds to the perceived difficulty while performing the tasks in our study and DIFF_large corresponds to the expected difficulty while performing larger manipulations (i.e., longer translations / rotations). Furthermore, we computed NASA TLX [13] and SUS [2] scores for both systems and asked the participants about the system that they (Q1) prefer, think provides (Q2) higher accuracy, (Q3) more cross-device benefits, and expect to be (Q4) easier to relearn.

- **H1** Overall, translation and rotation tasks can be completed faster with *Move'n'Hold Pro* than with *SotA*.
- **H2** Participants who use the MR-HMD after the VR-HMD, complete (a) translation and (b) rotation tasks with the MR-HMD faster than participants who use the MR-HMD first. Participants who use the VR-HMD after the MR-HMD, complete (c) translation and (d) rotation tasks with the VR-HMD faster than participants who use the VR-HMD first. These improvements are higher for *Move'n'Hold Pro* compared to *SotA*.
- **H3** The perceived difficulty of performing object (a) translations and (b) rotations in our study is lower with *Move'n'Hold Pro* than with *SotA*. The expected difficulty of performing longer (c) translations and (d) rotations with *Move'n'Hold Pro* is lower than with *SotA*.
- **H4** The NASA TLX scores for (a) translation and (b) rotation tasks are lower for *Move'n'Hold Pro* than for *SotA*.
- **H5** The SUS scores for *Move'n'Hold Pro* are higher than for *SotA*.

- **H6** If their job would require them to use and switch between MR-HHDs, MR-HMDs, and VR-HMDs, the participants will prefer to use *Move'n'Hold Pro* instead of *SotA*.
- **H7** The participants think relearning *Move'n'Hold Pro* will be easier than relearning *SotA*.
- **H8** If they had to complete similar tasks as accurately as possible, the participants think they could achieve the best results using *Move'n'Hold Pro* instead of *SotA*.
- **H9** The participants think that *Move'n'Hold Pro* provides more cross-device benefits than *SotA*.

6 RESULTS

The mean sum of TCTs for completing all translation and rotation tasks with all devices was only slightly lower with *Move'n'Hold Pro* than with *SotA* (see Fig. 5e, **H1**). To examine cross-device learnability, the participants were divided into two groups: *GroupA* used the MR-HMD after the VR-HMD and *GroupB* used the VR-HMD after the MR-HMD. When comparing the mean TCTs of these groups for single translation and rotation tasks, we found that *Move'n'Hold Pro* provided higher overall improvements than *SotA* when moving from one HMD to the other. When using the MR-HMD with *SotA*, *GroupA* was slower than *GroupB* – resulting in a negative value in the chart showing the cross-device improvement. Considering the MR-HMD, *GroupA* had lower TCTs when translating (Fig. 4a, **H2a**) and rotating (Fig. 4b, **H2b**) objects than *GroupB*. For the VR-HMD, *GroupB* was faster than *GroupA* using *SotA* and *Move'n'Hold Pro*. Thereby, higher improvements were observed for *Move'n'Hold Pro* during translation (Fig. 4c, **H2c**) and for *SotA* during rotation (Fig. 4d, **H2d**).

As shown in Fig. 5c, *Move'n'Hold Pro* was perceived slightly more difficult than *SotA* for translation (**H3a**) but *SotA* was rated substantially more difficult than *Move'n'Hold Pro* for rotation (**H3b**). Moreover, *Move'n'Hold Pro* is expected to be much easier than *SotA* during large translations and rotations (Fig. 5d, **H3c+d**). Regarding the NASA TLX, *Move'n'Hold Pro* outperformed *SotA* in the translation and rotation tasks (Fig. 5a, **H4a+b**). Similar to DIFF_exp and DIFF_large, *Move'n'Hold Pro* substantially decreased the workload experienced during rotation compared to *SotA*.

Usability measured by the SUS score for *Move'n'Hold Pro* (80.88) was higher than for *SotA* (66.13) (Fig. 5b, **H5**). Fig 6 shows that the majority chose *Move'n'Hold Pro* as their preferred system (**H6**), thinks that it will be easier to relearn (**H7**), provides more accuracy (**H8**) and cross-device benefits (i.e., benefits gained from using one device when switching to another) (**H9**).

7 DISCUSSION

Move'n'Hold Pro only slightly reduced temporal efforts compared to *SotA* in our study. However, based on the cross-device improvements (Fig. 4), we expect temporal efforts for *Move'n'Hold Pro* to decrease even further when the different devices are used more extensively and users have to switch between them more often. Furthermore, *Move'n'Hold Pro* substantially decreased cognitive efforts, especially while rotating objects. Considering Grier's [4] meta analysis, the total weighted NASA TLX score obtained for *Move'n'Hold Pro* (20.88) is lower than 90% of the UIs reviewed while for *SotA* (33.73) it is only lower than 75% of the UIs reviewed. Based on our difficulty ratings, we expect *Move'n'Hold Pro* to be particularly advantageous for large manipulations. Moreover, we conclude that *Move'n'Hold Pro* provides higher usability than *SotA* based on the SUS ratings and the final questionnaire. Considering the adjective ratings for SUS in [1], *Move'n'Hold Pro*'s SUS score indicates *Good* to *Excellent* user-friendliness while *SotA*'s SUS score only indicates *OK* to *Good* user-friendliness.

While using *Move'n'Hold Pro*, participants did not only apply right-thumb-touch to perform long and continuous manipulations but also repeatedly applied / released right-thumb-touch to perform short automated movements. Furthermore, participants used the smart rotation strategy which has already been observed in [11]: Aligning the tablet's front side with the manipulable box's front side and then moving the tablet towards the target box's front side while applying touch allowed the participants to intuitively perform complex rotations even without looking at the tablet (i.e., when using the MR-HMD or VR-HMD). In the last sessions of *Move'n'Hold Pro* many participants reported that they already know how to use *Move'n'Hold Pro* and do not need the explanatory video. To maintain comparability, the participants still watched the videos. Nevertheless, we consider these comments as a confirmation that *Move'n'Hold Pro* indeed enables seamless transitions from one of XR's access points to another.

While our approach does not allow hands-free interaction, we do not consider this to be a disadvantage as other researchers have shown that the integration of mobile devices provides versatile advantages such as offering a secondary output modality or input through non-spatial UIs [5, 7, 9, 14, 16, 18, 20] as well as a window towards reality when immersed in VR [3]. Furthermore, our study also revealed the limitations of hands-free input: Especially while performing object rotations with mid-air gestures the participants experienced difficulties due to movement restrictions in the wrist. Similar issues were observed when using the controller, however they were not as severe as with the gestures. *Move'n'Hold Pro* addresses this issue through the integration of continuous manipulation. Apart from mid-air gestures, using other hands-free input modalities such as gaze or speech for 3D manipulation is even more complex. Therefore, we believe that for 3D manipulation, the integration of a tablet controller as provided in *Move'n'Hold Pro* is more suitable than hands-free input.

8 CONCLUSIONS AND FUTURE WORK

In this paper, we presented a holistic XR interaction technique considering different XR devices and degrees of virtuality. To this end, we developed *Move'n'Hold Pro* – a set of object translation and rotation techniques that is consistent for MR-HHDs, MR-HMDs, and VR-HMDs. With *Move'n'Hold Pro* we seek to make it easier for future users to seamlessly switch between XR's main access points and reduce cognitive and temporal efforts required to re-adapt to the system. The development of *Move'n'Hold Pro* builds up on latest research on MR-HHD-UIs such that using all three devices, virtual objects can be manipulated based on a tablet's movement and peripheral touch. In a first evaluation comparing *Move'n'Hold Pro* to a State of the Art system, *Move'n'Hold Pro* was rated as the preferred system and to be easier to relearn. It reduced cognitive efforts, improved usability, and provided more cross-device benefits.

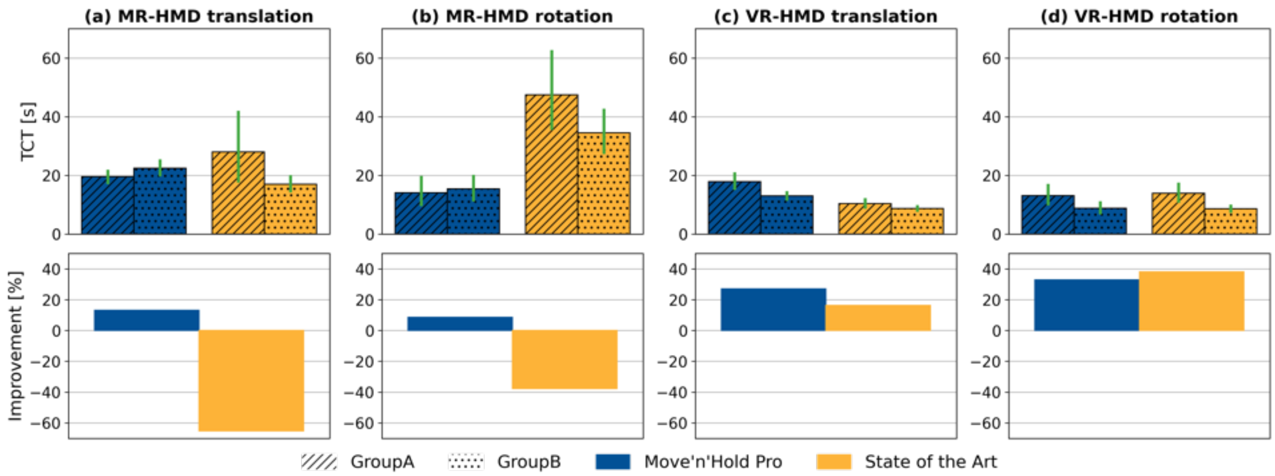

Figure 4: First row: Mean TCTs and 95% confidence intervals for single translation / rotation tasks with the MR-HMD / VR-HMD by *GroupA* (VR-HMD before MR-HMD) / *GroupB* (MR-HMD before VR-HMD). Second row: Improvement of the mean TCTs from (a+b) *GroupB* to *GroupA*, (c+d) *GroupA* to *GroupB*.

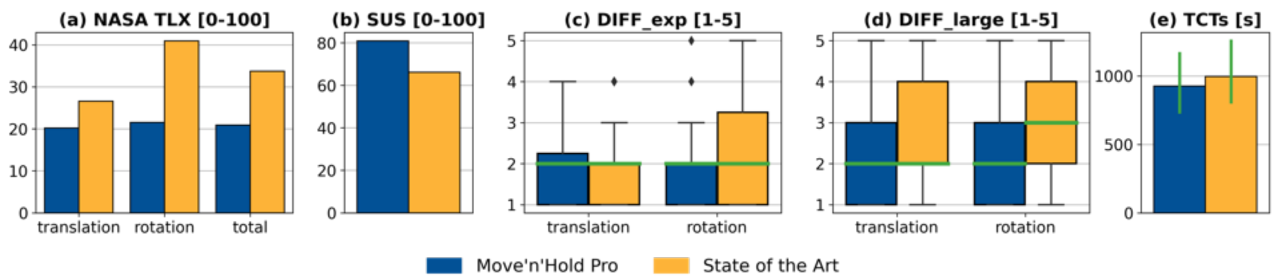

Figure 5: (a) mean weighted NASA TLX scores [0 = best; 100 = worst]; (b) mean SUS scores [0 = worst; 100 = best]; (c+d) perceived difficulty [1=very easy; 5 = very difficult]; (e) mean total TCTs [s] of all translation + rotation tasks and 95% confidence intervals.

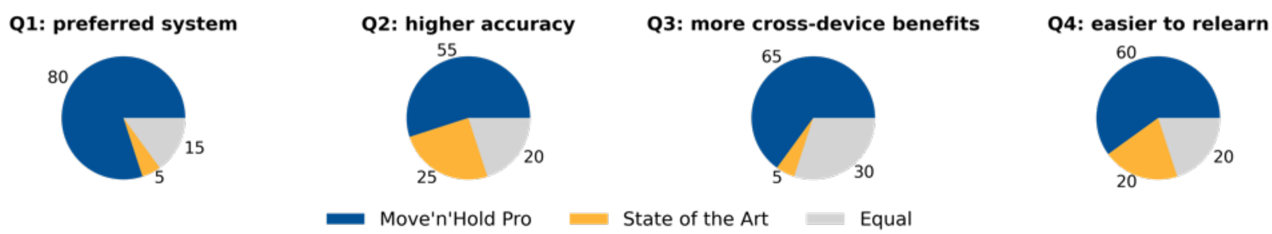

Figure 6: Answers to the final questions comparing *Move'n'Hold Pro* and *SotA* in [%].

Here, we present the integration of a tablet controller for two HMDs: the Microsoft HoloLens 2 and HTC VIVE PRO. However, based on the implementation described above, *Move'n'Hold Pro* can also be extended to other HMDs which allow receiving and sending http requests. Furthermore, the virtual boxes that were manipulated in our study can be easily replaced with multiple virtual objects like furniture, machines, or components of virtual prototypes in various domains such as the automotive or aerospace industry. In future work, we are planning to conduct an extended evaluation with a larger sample size and to apply *Move'n'Hold Pro* to different complex use cases such as factory layout planning. We expect planning engineers to benefit from *Move'n'Hold Pro*'s scalability throughout the planning process: They may easily switch between VR-HMDs in early stages of the planning process and MR-HMDs or MR-HHDs in later stages when physically existing factory parts are augmented with virtual components that can be translated and rotated using *Move'n'Hold Pro*.

ACKNOWLEDGMENTS

We thank H. Subbaraj for supporting the data analysis as well as K. Mannweiler for supporting the implementation. This research was funded by the Innovative University Initiative (Bund-Länder-Initiative Innovative Hochschule, BMBF / GWK, 13IHS254B) at RPTU Kaiserslautern-Landau. Furthermore, it was supported in part by NASA through the HOME (Habitats Optimized for Missions of Explorations) Space Technology Research Institute at UC-Davis.

REFERENCES

[1] Aaron Bangor, Philip T. Kortum, and James T. Miller. 2008. An Empirical Evaluation of the System Usability Scale. *International Journal of Human–Computer Interaction* 24, 6 (2008), 574–594. https://doi.org/10.1080/10447310802205776

[2] John Brooke. 1996. *SUS: A quick and dirty usability scale.* Taylor and Francis, London, 189–194.

[3] Ceenu George, An Ngo Tien, and Heinrich Hussmann. 2020. Seamless, Bi-directional Transitions along the Reality-Virtuality Continuum: A Conceptualization and Prototype Exploration. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 412–424. https://doi.org/10.1109/ISMAR50242.2020.00067

[4] Rebecca A. Grier. 2015. How High is High? A Meta-Analysis of NASA-TLX Global Workload Scores. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 59, 1 (2015), 1727–1731. https://doi.org/10.1177/1541931215591373

[5] Sebastian Hubenschmid, Johannes Zagermann, Simon Butscher, and Harald Reiterer. 2021. STREAM: Exploring the Combination of Spatially-Aware Tablets with Augmented Reality Head-Mounted Displays for Immersive Analytics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. ACM, New York, NY, USA, Article 469, 14 pages. https://doi.org/10.1145/3411764.3445298

[6] Mohamed Kari and Christian Holz. 2023. 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) *(CHI '23)*. ACM, New York, NY, USA, Article 528, 15 pages. https://doi.org/10.1145/3544548.3580677

[7] Pascal Knierim, Dimitri Hein, Albrecht Schmidt, and Thomas Kosch. 2021. The SmARtphone Controller: Leveraging Smartphones as Input and Output Modality for Improved Interaction within Mobile Augmented Reality Environments. *i-com* 20, 1 (2021), 49–61. https://doi.org/10.1515/icom-2021-0003

[8] Khanh-Duy Le, Tanh Quang Tran, Karol Chlasta, Krzysztof Krejtz, Morten Fjeld, and Andreas Kunz. 2021. VXSlate: Exploring Combination of Head Movements and Mobile Touch for Large Virtual Display Interaction. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference* (Virtual Event, USA) *(DIS '21)*. ACM, New York, NY, USA, 283–297. https://doi.org/10.1145/3461778.3462076

[9] Weizhou Luo, Eva Goebel, Patrick Reipschläger, Mats Ole Ellenberg, and Raimund Dachselt. 2021. Exploring and Slicing Volumetric Medical Data in Augmented Reality Using a Spatially-Aware Mobile Device. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 334–339. https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00076

[10] Asier Marzo, Benoît Bossavit, and Martin Hachet. 2014. Combining multi-touch input and device movement for 3D manipulations in mobile augmented reality environments. In *Proceedings of the 2nd ACM Symposium on Spatial User Interaction* (Honolulu, HI, USA) *(SUI '14)*. ACM, New York, NY, USA, 13–16. https://doi.org/10.1145/2659766.2659775

[11] Vera Marie Memmesheimer, Kai Jonas Klingshirn, Bahram Ravani, and Achim Ebert. 2023. Move'n'Hold: Scalable Device-Based Interaction for Mixed Reality Handheld Displays. In *Proceedings of the European Conference on Cognitive Ergonomics 2023* (Swansea, United Kingdom) *(ECCE '23)*. ACM, New York, NY, USA, Article 13, 8 pages. https://doi.org/10.1145/3605655.3605656

[12] Daniel Mendes, Maurício Sousa, Rodrigo Lorena, Alfredo Ferreira, and Joaquim Jorge. 2017. 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* (Gothenburg, Sweden) *(VRST '17)*. ACM, New York, NY, USA, Article 27, 8 pages. https://doi.org/10.1145/3139131.3139157

[13] NASA TLX Paper and Pencil Version Instruction Manual. 2024. https://humansystems.arc.nasa.gov/groups/tlx/tlxpaperpencil.php accessed March 17, 2024.

[14] Jie Ren, Yueting Weng, Chengchi Zhou, Chun Yu, and Yuanchun Shi. 2020. Understanding Window Management Interactions in AR Headset + Smartphone Interface. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. ACM, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382812

[15] Ali Samini and Karljohan Lundin Palmerius. 2016. A study on improving close and distant device movement pose manipulation for hand-held augmented reality. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology* (Munich, Germany) *(VRST '16)*. ACM, New York, NY, USA, 121–128. https://doi.org/10.1145/2993369.2993380

[16] Helmut Schrom-Feiertag, Georg Regal, Jaison Puthenkalam, and Stefan Suette. 2021. Immersive Experience Prototyping: Using Mixed Reality to Integrate Real Devices in Virtual Simulated Contexts to Prototype Experiences with Mobile Apps. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 75–81. https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00025

[17] Hemant Bhaskar Surale, Aakar Gupta, Mark Hancock, and Daniel Vogel. 2019. TabletInVR: Exploring the Design Space for Using a Multi-Touch Tablet in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300243

[18] Arda Ege Unlu and Robert Xiao. 2021. PAIR: Phone as an Augmented Immersive Reality Controller. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology* (Osaka, Japan) *(VRST '21)*. ACM, New York, NY, USA, Article 27, 6 pages. https://doi.org/10.1145/3489849.3489878

[19] Lu Zhao, Yue Liu, Dejiang Ye, Zhuoluo Ma, and Weitao Song. 2020. Implementation and Evaluation of Touch-based Interaction Using Electrovibration Haptic Feedback in Virtual Environments. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 239–247. https://doi.org/10.1109/VR46266.2020.00043

[20] Fengyuan Zhu and Tovi Grossman. 2020. BISHARE: Exploring Bidirectional Interactions Between Smartphones and Head-Mounted Augmented Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. ACM, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376233