Wall-based Space Manipulation Technique for Efficient Placement of Distant Objects in Augmented Reality

Han Joo Chae
Seoul National University
Seoul, Republic of Korea
hanjoo@hcil.snu.ac.kr

Jeong-in Hwang
Seoul National University
Seoul, Republic of Korea
jihwang@hcil.snu.ac.kr

Jinwook Seo
Seoul National University
Seoul, Republic of Korea
jseo@snu.ac.kr

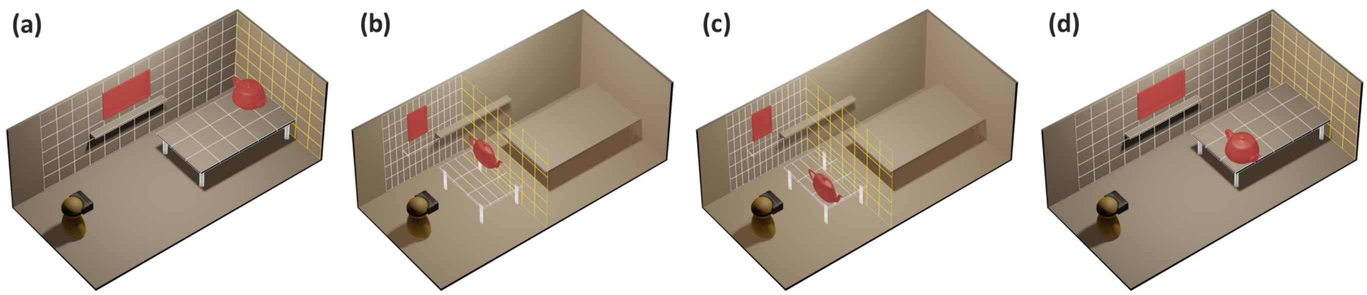

Figure 1. An overview of a basic usage scenario of WSM where virtual objects (a kettle and a wall frame in red) are placed in a room. (a) The *wall planes* with gridlines and surface meshes are overlaid on top of the detected walls and other real objects, respectively. (b) A user identifies a kettle far away and drags a *wall plane* toward oneself to squeeze the space between the user and the wall. (c) The user moves the kettle to the closer edge of the table in the reduced space. (d) As the user double-taps the *wall plane*, the space returns to the normal scale, and all the objects are correctly positioned in the real space as intended.

ABSTRACT

We present a wall-based space manipulation (WSM) technique that enables users to efficiently select and move distant objects by dynamically squeezing their surrounding space in augmented reality. Users can bring a target object closer by dragging a solid plane behind the object and squeezing the space between them and the plane so that they can select and move the object more delicately and efficiently. We furthermore discuss the unique design challenges of WSM, including the dimension of space reduction and the recognition of the reduced space in relation to the real space. We conducted a user evaluation to verify how WSM improves the performance of the hand-centered object manipulation technique on the HoloLens for moving near objects far away and vice versa. The results indicate that WSM overall performed consistently well and significantly improved efficiency while alleviating arm fatigue.

CCS Concepts

•**Human-centered computing** → **Mixed / augmented reality;** *Gestural input; User interface design;* Usability testing;

UIST 2018, October 14–17, 2018, Berlin, Germany.

ACM ISBN 978-1-4503-5948-1/18/10 ...$15.00.
http://dx.doi.org/10.1145/3242587.3242631

Author Keywords

3-D interaction technique; space manipulation; distant object placement; augmented reality.

INTRODUCTION

Recent advances in augmented reality (AR) technology have provided people with more realistic and immersive experiences by seamlessly blending virtual objects into their real surroundings and allowing natural interaction methods such as direct manipulation using hands. With AR, various virtual objects can be placed and manipulated within the real environment, and it is actively being used in diverse fields, from professional usage (e.g., architecture and medicine) to personal usage (e.g., placing furniture at home prior to purchasing).

However, despite their practical applications, interaction techniques still lack support for distant manipulation. In particular, moving nearby objects far away or distant objects closer is often difficult and burdensome as many interaction techniques require people to repeat similar gestures to move forward and backward (along the *z*-axis) multiple times. Moreover, given the limitations of human visual perception and rendering technologies, recognizing exact positions of virtual objects than those of real ones is difficult [10] and even worse when the objects are distant from the user [2, 22]. For instance, even on one of the most advanced AR head-mounted displays (HMDs), Microsoft HoloLens, the positions of objects are often misunderstood (Figure 2). Though the inaccuracy in the HoloLens' tracking algorithm (SLAM) might also have contributed to the incorrect perception, the perceptual issues with distant objects will remain even with a more precise algorithm.

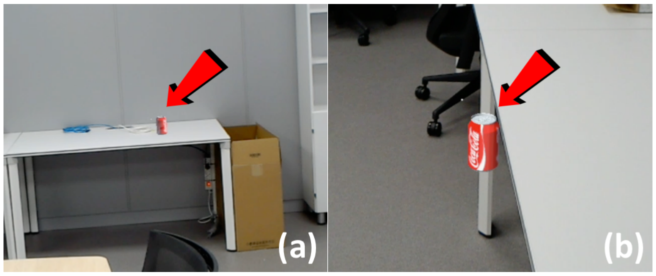

Figure 2. On the HoloLens, (a) a soda can that seems to be placed on the table from far away (b) is actually off the table if seen from nearby.

While remote manipulation has been widely studied in the field of virtual reality (VR), research regarding the difficulty of moving objects forward and backward in AR is limited. Unlike VR, where people can fly around the VR world using controllers while sitting on a chair, in AR, people need to physically move around the space to change perspectives, and an interaction with virtual objects could be affected by seeing the real environment [9]. Although efforts have been exerted to improve interactions for selecting and controlling virtual objects in the real environment [5, 6, 17, 20], they fall short and remain incomplete for the domain of effectively moving and placing objects across a longer distance along the *z*-axis.

In this paper, we suggest a wall-based space manipulation (WSM) technique that can be added upon various existing interaction techniques and enables people to dynamically scale down the surrounding space for efficient virtual-object placement in an indoor environment using AR. With WSM, people can move virtual objects with smaller motions in the squeezed space between them and the wall behind the objects (Figure 1). We evaluated our technique by comparing performances between the cases where WSM was used or not. The results indicate that the participants were able to move virtual objects more consistently and efficiently with less arm fatigue when WSM was used.

RELATED WORK

Since the conception of VR, the remote selection and manipulation of virtual objects have been widely studied. In this section, we discuss various interaction techniques of prior work in three-dimensional space—both in VR and AR as their interaction methods are very similar—focusing on moving target objects for a longer distance along the *z*-axis.

Interaction techniques in VR can be divided into two different metaphors based on the user's perspective: exocentric and egocentric [19]. In an exocentric metaphor, a scene or object is seen from the outside, while in egocentric, it is seen from the inside like an immersive VR. World-In-Miniature [21] uses an exocentric metaphor where the user scales down a 3-D world into a miniature and manipulates tiny objects. Although this technique allows efficient large-scale manipulation, accurate selection and fine-grained manipulation can be difficult [15]. Furthermore, having an extra miniature can clutter the display [7], and the technique encompasses an overhead of looking at the miniature and the real world alternately.

In the egocentric metaphor, interaction techniques to select and manipulate remote objects can be classified into two categories: arm extension and ray casting techniques [3]. As an arm extension, the go-go technique [18] removed the restriction of the limited interaction range of the original virtual-hand technique [15] by nonlinearly mapping real and virtual hands. Although such techniques allowed a significant extension of the user's reach, selecting and placing distant objects remained problematic; a small motion of the user's hand led to a large translation in the virtual hand [7], and recognizing the exact positions of faraway objects was still difficult.

Ray casting [14], which attaches an intersecting object to the ray extended from the user's hand, made distant object selection easy and ray casting from the eye (RCE) [1] later improved the original ray casting by eliminating the eye-hand mismatch problems. Ray casting is effective in object selection, though inefficient for controlling the object's distance from the user as it needs to be combined with other methods, such as the "fishing reel" metaphor [3]. Extender grab [15] and direct HOMER [3] techniques took hybrid approaches by allowing the user to select an object with ray casting and manipulate the selected object using their hands instead of ray attachment. Particularly, the direct HOMER technique supported distant manipulation in a way that moving the hand twice as far from the body places the object twice as far away. Modifications to existing hybrid techniques suited for their domain have also been studied; for instance, scaled HOMER [23] added velocity-based scaling to the HOMER technique to improve the precision of the distant 3-D manipulation, and WeARHands [6] and BoostHand [8] adapted HOMER and go-go techniques, respectively, to support bare-hand interaction in remote AR scenarios. Though the hybrid techniques combining ray casting and hand-centered manipulation are effective and commonly used in AR, they still hinge on human-perception limitation for handling distant objects. Physical controllers (e.g., joystick and mouse) can also be used together with ray casting as in the indirect HOMER [3] technique, but they lack natural mappings and can make placing an object at some arbitrary position and orientation difficult [14].

Our WSM technique dynamically reduces the world for efficient placement of distant objects and is closely related to Scaled-world grab (SWG) [15], which isotropically scaled down the VR world for remote manipulation. However, since SWG required the user to select an object to initiate the scaling, the problem of recognizing and selecting small and faraway objects remained unsolved. Moreover, SWG's scaling factor, which was fixed to the ratio of the distance of the object to the distance of the user's hand, was not effective for moving near objects far away or distant objects closer; selecting near objects did not sufficiently scaled down the world and faraway objects reduced the world too small for fine-grained control. In contrast, WSM exploits large walls for easy world scaling and enables the user to efficiently move objects at any location forward and backward for a longer distance.

Attempts have been made to add handy functionality to enhance existing techniques rather than propose a new metaphor of interaction. SnapToReality [17] added a sophisticated feature that automatically snaps virtual objects onto the edges and surfaces of real objects in the environment to the RCE technique to assist efficient and accurate placement. Projective Windows [11] allowed the user to easily scale and snap 2D windows on planar surfaces using hand gestures. However, both techniques do not help the user place virtual objects on arbitrary locations where no real constraints are available. In addition, physical-constraint detection is prone to error, especially for small and faraway objects. Our WSM technique, which fully utilizes the advantages of generally large and well-detected walls, can be added upon existing interaction techniques so that the user can select and move a virtual object to any arbitrary location more efficiently.

WALL-BASED SPACE MANIPULATION

The WSM technique utilizes a large surface (i.e., a wall) that is easily detected on portable devices in real time to squeeze the space between a user and the wall for efficient selection and placement of objects. Since the real space cannot actually be reduced and walls cannot be physically moved, we created interactive *wall planes* overlaid on top of each detected wall so that the user can drag them instead of the unmovable real walls. WSM adds upon many existing interaction techniques in AR to address their two major problems: (1) difficulties in selection and recognition when objects are distant from the user, and (2) inefficient repositioning along the z-axis.

Overview of the Technique

To give an overview of the basic concept of WSM, we describe a simple usage scenario where a user moves a virtual object placed far away to another location (Figure 1). For clarity, we assume the user uses the direct HOMER technique on the HoloLens where they point at the object using the cursor on the center of the HoloLens, use the air tap [12] gesture to select, and drag the object by moving their hands.

1. The user enables WSM via a voice command, *"Turn on space manipulation."* Room boundaries (i.e., walls, floor, and ceiling) are detected as the user scans the room. Then, *wall planes* are created for the detected walls and surface meshes of the real objects (e.g., a table and a shelf) except for the boundaries are overlaid onto the scene.
2. The user identifies the virtual object far away and drags a *wall plane* toward oneself to squeeze the space between the wall and the user. All the *wall planes*, virtual objects, and surface meshes are moved and scaled accordingly.
3. In the reduced space, the user drags the virtual object to the desired location with a simple and short hand movement.
4. As the user double-taps (i.e., air tapping twice while pointing) on the *wall plane*, the space returns to the normal scale, and all the objects are correctly positioned in the real space. The user disables WSM via a voice command, *"Turn off space manipulation."*

Design Decision and Consideration

Why Walls?

Indoor places are generally surrounded by rectangular walls that are flat and perpendicular to floors. While detecting and classifying many complicated objects on the fly is difficult, walls are relatively easy to detect even on mobile devices in real time. On the HoloLens, our prototype takes less than ten seconds to detect walls and create *wall planes*. The large walls are also easily selected and manipulated from distant locations in general and boundaries like walls greatly help users feel and recognize the reduced space. Moreover, walls are robust from occlusion since empty wall areas can be identified even in a crowded room. That is, WSM can be used even when a small—but large enough to be detected as a wall—part of a wall is visible. In an open space without walls, WSM could utilize other flat surfaces or detectable landmarks, such as fences or pillars, to reduce the space; however, because of all the wall's benefits described above, we focus on indoor environments where walls are available.

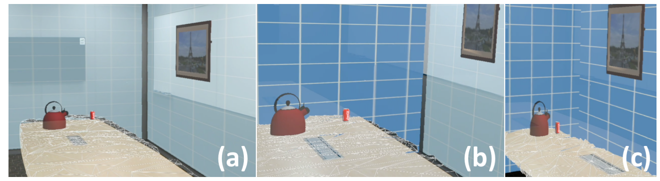

Figure 3. When (a) the room is in the normal scale, (b) is squeezed along the surface normal of one *wall plane* (in darker blue), and (c) is squeezed using two *wall planes*.

Dimension of Space Reduction

WSM scales space along the surface normal of the selected wall. Though scaling across a single dimension might break the aspect ratios of objects, one-dimensional scaling is advantageous since it assists the user for forward and backward movement without additional complications. Also, WSM's one-dimensional scaling keeps the shapes and sizes of objects along the x-y dimensions as they are; the objects might look thinner depth-wise but mostly the same from the user's perspective, thereby preventing them from being too small to manipulate. In addition, as the inaccuracy of depth judgment [16] can be alleviated by reducing the depth, users can have more detailed control within the squeezed space. Furthermore, as the space can be scaled using multiple *wall planes* together, the user can decide across which dimensions they would want to scale the space. For instance, if an object is located at the front-right corner of the room, the user can first drag the *wall plane* in the front and then the one on the right so that the space is reduced across both dimensions (Figure 3c).

Alternative Approaches

Isotropic scaling used in World-In-Miniature [21] and Scaled-world grab [15] allows large-scale manipulation without distortion. However, they make distant objects too small for fine-grained manipulation while WSM prevents objects from being too small by scaling the space along a single dimension and letting the user dynamically adjust the amount of scaling. Shifting the space instead of scaling is another alternative as it does not reduce the object size or create distortion, but some part of the space initially between the user and the wall would have to be cropped or located behind the user when the space is shifted. This makes placing objects in certain locations more difficult or even impossible. In contrast, WSM utilizes the entire space for efficient placement. Snapping features like SnapToReality [17] and Projective Windows [11] are handy to place objects based on the environment's physical constraints (e.g., surfaces and edges). As WSM supports the user to place objects even on areas where the constraints are limited, snapping and WSM can complement each other when they are combined. This is one of the main reasons that we designed WSM to work with other existing techniques so that it can provide further advantages while retaining the benefits of the existing techniques.

Understanding the Reduced Space

For the user to shrink the space and interact with objects within the reduced space, they first need to visually understand how much the space is reduced, and then recognize both the locations of the objects at interest and the target positions to which the objects are desired to be moved. Recognizing how each position in the reduced space corresponds to the position in the real space is particularly crucial.

We thus provide visual feedback while using WSM for this matter. The user can recognize space reduction by observing changes in the positions and scales of surrounding *wall planes*, virtual objects, and surface meshes of the real objects (Figure 3). We drew gridlines on the *wall planes* to better instruct on the changes in space size; the user can notice the space reduction by observing the narrower grid spacing of the side *wall planes*. Furthermore, the surface meshes can act as reference points from reduced space to real space. For instance, if the user wants to place an object on a table they can simply identify the surface mesh of the table from the reduced space and place the object on top of the table's mesh.

Implementation

We applied our WSM technique onto the direct HOMER [3] technique, one of the most popular in AR interaction, to see how WSM can improve the performance of moving objects along the z-axis. We implemented our current prototype on the HoloLens using the Mixed Reality Toolkit (MRTK) [13].

Room Boundary Detection and Wall Plane Creation

As a room is being scanned, triangular meshes are created on the detected surfaces, and flat surfaces are identified by the spatial mapping of MRTK. A ceiling and floor are classified by the maximum horizontal surfaces above and below the user's head, respectively, and any vertical surfaces greater than three square meters in area are classified as walls. *Wall planes* with gridlines are then created and overlaid on top of the detected walls. All the surface meshes classified as room boundaries were removed from the scene so that the user can focus more on the objects within the boundaries.

Manipulating Space

The user can scale the space as they want by dragging one of the surrounding *wall planes*. The sensitivity of the *wall plane* while dragging is determined by the distance between the initial position of the *wall plane* and user (D_{w_0,u_0}) as the position of the *wall plane* (P_w) is updated by the following equation:

$$P_w = P_{w_0} + \hat{n}_w \cdot \frac{D_{w_0,u_0}}{R_{max}} \cdot D_{h,h_0} \tag{1}$$

where P_{w_0} is the initial position of the *wall plane*; $\hat{n}_w$ is the unit normal vector of the *wall plane*; D_{h,h_0} is the distance between the current and initial position of the user's hand; and R_{max} is the user's maximum hand-reach. In our prototype, we set R_{max} to 0.6 meters, found within the comfortable interaction range. Note that we do not consider the case where the *wall plane* is pushed farther away from its initial position since WSM is primarily designed to reduce the space.

From the above equation, $\frac{D_{w_0,u_0}}{R_{max}}$ allows the user to drag any *wall plane* all the way toward them in a single drag, regardless of its distance from the user. Though this might make the *wall planes* overly sensitive when they are very far, it would not damage the efficacy of the technique since the space scaling does not require precise adjustments and the user only needs to drag the *wall planes* to somewhere near them. Returning the reduced space back to the normal scale is as easy as double-tapping on the *wall plane*.

Changes in Position and Scale

As a user scales the space by dragging a *wall plane*, the positions and scales of all other *wall planes*, virtual objects, and surface meshes are updated accordingly. WSM ensures that everything is moved and scaled in relation to the *wall plane* movement, but not affected by any change in the user's position or orientation during the space manipulation. Each item is scaled by $\frac{D_{w,u_0}}{D_{w_0,u_0}}$ along $\hat{n}_w$, where D_{w,u_0} is the current distance of the *wall plane* from the initial position of the user. The position of each item is determined by the following equations:

$$P_i = P_{i_0} + proj_{\hat{n}_w}(P_w - P_{i_0}) + \hat{n}_w \cdot D_{i_0,w_0} \cdot \frac{D_{w,u_0}}{D_{w_0,u_0}} \tag{2}$$

where P_{i_0} is the initial position of the item; $proj_{\hat{n}_w}(P_w - P_{i_0})$ is projecting the vector from P_{i_0} to P_w onto $\hat{n}_w$; and D_{i_0,w_0} is the distance between the initial position of the item and wall.

In addition, the user can freely move the virtual objects in the reduced space as they do in the normal space; they can even move around the reduced space while dragging objects. The only difference is that small changes in the position of the objects along the z-axis within the reduced space will be largely amplified when the space returns to the normal scale.

USER EVALUATION

We conducted a controlled user study to verify whether WSM could improve the direct HOMER technique for moving nearby objects far away and vice versa, and whether WSM could reduce total arm movement and alleviate arm fatigue.

Participants

We recruited twelve participants (six females) from a local university between ages twenty and thirty-four ($\mu = 23.36, \sigma = 3.82$), and they all had normal or corrected-to-normal vision. Three participants had previous experience with AR HMDs.

Study Design

We used a within-subjects design and a three-by-three-meter wall located eleven meters away from the participant's position for space manipulation (Figure 4a). The participants were given tasks of moving and aligning a source to a target (Figure 4b). The three independent variables were WSM state

Figure 4. (a) A user performs a task standing eleven meters apart from the wall. (b) A screenshot of the user's view with a *wall plane*, a source (red cube), and a target (green cube).

(whether WSM is enabled or disabled), direction (whether to move a nearby source to a far target or a far source to a nearby target), and source and target sizes [0.3 meters (small), 0.4 meters (medium), and 0.5 meters (large) in the length of each edge]. While sizes varied, those of the target and source were always equal. Within the eleven meters between the user and the wall, the source and target positions were randomly generated with the fixed distance of eight meters along the z-axis. This ensured seeing the WSM effect primarily toward the z-axis.

We made a set of tasks consisting of twelve trials (two directions $\times$ three sizes $\times$ two repetitions) and the participants performed two sets of tasks for each WSM state. That is, each participant performed forty-eight trials (two states $\times$ two sets $\times$ twelve trials). The participants proceeded to the next trial only when they successfully moved the source to the target and a trial was successful when the centroids of the two cubes were closer than half the length of their edges. When WSM was disabled, the participants directly dragged the source to the target, and to ensure they only rely on their visual perception, the success of a trial was notified only when the dragging was completed. When WSM was enabled, the participants were required to use the space manipulation, and the success of a trial was determined when the space returned to the normal scale. This ensured that each WSM-enabled trial included the overhead of dragging and returning the *wall plane*.

Procedure

After the participants signed the consent form, we explained how to use our interface with the HoloLens, and they were allowed to practice as much as they felt comfortable. The orders within each task set were randomized, and the order of the WSM states was counterbalanced to avoid any learning effect. The participants performed two sets of tasks for each given WSM state and were told to perform each task as fast and accurately as possible. If needed, they were allowed to take a break between sets. We measured the time and total distances of their hands traveled during each task using the hand tracking of the HoloLens. The participants filled out SUS [4] for both WSM states and we interviewed them after all the tasks were completed.

Results

We present the quantitative results of the time taken to complete each task and the distances of the participants' hands traveled during each task. Though we recruited twelve participants, only eleven completed the experiment since one was unable to wear the HoloLens properly to see its display. Thus, we analyze and report the eleven participants' results (five participants performed the task first with WSM enabled and then disabled, while the other six did in the reverse order).

Time to Complete Each Task

The results of the three-way RM ANOVA show that there were significant effects of the WSM state ($F(1,10) = 7.31, p = 0.022$), the direction ($F(1,10) = 48.72, p < 0.001$), and the cube size ($F(1.3,12.8) = 9.18, p = 0.007$ with Greenhouse-Geisser correction) upon completion time. Pairwise comparisons with the Bonferroni correction revealed that the participants completed the tasks significantly faster with WSM enabled ($\mu = 15.38, \sigma = 1.64$) than disabled ($\mu = 17.09, \sigma = 14.37$); direction far-to-near ($\mu = 14.31, \sigma = 1.20$) than near-to-far ($\mu = 18.16, \sigma = 1.64$); and large cube sizes ($\mu = 14.95, \sigma = 1.30$) than other sizes (small: $\mu = 17.55, \sigma = 1.80$; medium: $\mu = 16.21, \sigma = 1.19$). In addition, there were significant interactions between state and direction ($F(1,10) = 50.71, p < 0.001$), state and size ($F(2,20) = 7.32, p = 0.004$), and direction and size ($F(2,20) = 6.43, p = 0.007$) upon completion time. The estimated marginal means showed that the participants took significantly longer when (1) WSM was disabled and the direction was near-to-far, (2) WSM was disabled and the cube was small, and (3) the direction was near-to-far and the cube was small (Figure 5 left).

Distance of the Hand Traveled During Each Task

The distances of the participants' hands traveled during each task were also analyzed using the three-way RM ANOVA and there were significant effects of the WSM state ($F(1,10) = 17.53, p = 0.002$), and the direction ($F(1,10) = 2.96, p < 0.001$). Pairwise comparisons with the Bonferroni correction revealed that the participants moved their hands significantly less with WSM enabled ($\mu = 2.17m, \sigma = 0.13$) than disabled ($\mu = 2.60m, \sigma = 0.12$), and direction far-to-near ($\mu = 2.24m, \sigma = 0.11$) than near-to-far ($\mu = 2.54m, \sigma = 0.13$). Significant interactions existed between state and direction ($F(1,10) = 31.97, p < 0.001$) and among all three independent variables ($F(2,20) = 4.32, p = 0.028$). The estimated marginal means showed that the participants moved their hands significantly more when WSM was disabled and the direction was near-to-far. Moreover, when the direction was near-to-far and WSM was disabled, the participants moved their hands significantly more even for the large cubes (Figure 5 right).

Subjective Assessment

We measured SUS for both WSM states as a subjective assessment. The average SUS score was better with WSM enabled ($\mu = 77.27, \sigma = 15.99$) than disabled ($\mu = 71.36, \sigma = 15.75$), but the difference was not statistically significant.

DISCUSSION

In this section, we discuss the user-evaluation results in relation to our observations as well as the comments received from the participants during the interview. We also discuss the flexibility and applicability of WSM.

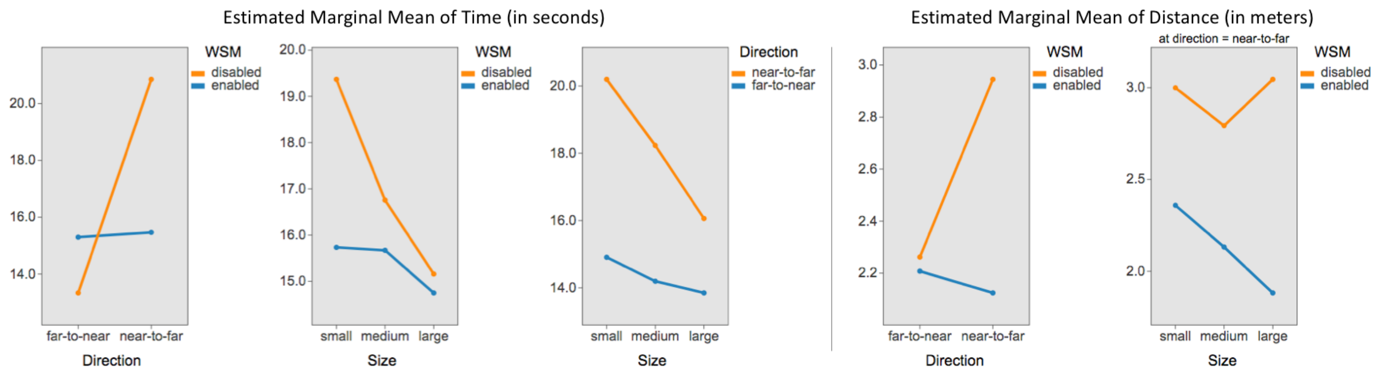

Figure 5. Left: the plots of the estimated marginal means of the time (state of WSM × direction, state of WSM × size, and direction × size). Right: the plots of the distance (state of WSM × direction, state of WSM × size at direction = near-to-far). All of the above showed significant interactions.

WSM Improved Performance with Less Arm Fatigue

The results of the completion time and hand-travel distance indicate that when WSM is applied to the direct HOMER technique, the user moved the objects forward and backward more efficiently with less arm fatigue.

> *I think it can be used easily for controlling far objects in detail. (P3)*

> *I was able to move my hand less (with WSM)...It was more convenient. (P1)*

The results also show that WSM was more efficient despite its inherent overhead that requires the user to drag a *wall plane* and then put it back with a double-tap, respectively, before and after repositioning. That is, with WSM, the problems of inefficient repositioning along the z-axis and the difficulties in selecting and recognizing distant objects were effectively mitigated.

WSM Enhanced Overall Consistency

The results of the significant interactions indicate that the direct HOMER technique without WSM significantly suffered when the objects were small or the task was moving nearby objects to far locations. It could have been difficult for the user to precisely place the objects from far away because the sensitivity of the direct HOMER technique increased as the objects moved away from the user, and the perceptual limitation made small and distant objects even more difficult to be recognized. By contrast, when WSM was applied, the performance was consistently high overall. That is, by adding WSM to the direct HOMER technique, not only did the efficiency of the technique on distant manipulation significantly increase, but also its robustness against the variations on object size, movement direction, and user-object distance was considerably enhanced.

> *I didn't feel much about the size difference using the wall (wall plane), but it was difficult to control precisely without it. (P4)*

One interesting behavior we observed during the experiment was that the participants often grabbed relatively near objects with their arms almost stretched, thereby making it difficult for them to move objects farther away. This partially explains the lower performance of the participants for moving near objects far away during the experiment and we interpret this as natural human behavior that the participants, without noticing, stretched their arms to reach for virtual objects as with real objects. Thus, WSM aligns well with such behavior since it only requires the user to pull and double-tap to bring and return the *wall plane*, respectively, without any pushing gesture.

Wall Plane Was Perceptually Advantageous

In addition to the results of the experiment, we were able to gain feedback from the participants about how they felt when they interacted with the *wall planes* to manipulate the space. All participants well understood the concept of WSM and the *wall planes* help them recognize the positions and distances of the objects within the reduced space.

> *It was convenient because I could reference the wall (wall plane) to sense the space. (P12)*

> *(The wall plane) helped me feel the depth. (P2)*

Also, most participants agreed on the advantage of the large surface area of the *wall plane*, as P10 mentioned, *"Selecting the wall wasn't hard because it was large."* Though one participant (P3) said he might not use WSM everyday due to the extra efforts to move the *wall plane*, all the participants agreed that WSM was convenient and effective overall.

WSM Is Flexible and Widely Applicable

WSM is designed to provide high degree of freedom in terms of its applicability. Although we implemented our prototype on the HoloLens, WSM is not limited to the direct HOMER technique or the HoloLens and is intended to be applied to various interaction techniques and devices in a non-intrusive way to support remote placement. That is, while the concept of the *wall plane* is used to manipulate space, methods to point, select and drag are dependent on whichever technique and device the user chooses based on their application and usage scenarios. Moreover, the *wall plane* does not even appear unless WSM is explicitly enabled by the user and the method to toggle WSM can also be decided by them. For instance, on a mobile device, the user might point at an object using the rear camera, press a button on the screen to enable WSM, and use touch gestures to drag the *wall plane*.

Reducing the space inevitably reduces the sizes of objects; however, it may not be prominent from the user's perspective since the objects come closer while they get smaller, and we did not experience any particular issues using our prototype. Paper-thin or extremely small objects might not be suitable for WSM, but they would be difficult to be manipulated on any 3-D interface. WSM let the user to flexibly decide on the direction and number of *wall planes* to be used for space reduction depending on their requirements as well as the shape, size and location of each object. Therefore, WSM can be integrated with various interaction techniques and devices and selectively be used by the user based on their application and circumstances.

LIMITATION AND FUTURE WORK

Though the user study showed the efficacy and advantages of WSM, some apparent design limitations still exist. Currently, WSM only supports translation without rotation and scaling. Rotating and scaling objects in the squeezed space would definitely be interesting. In addition, the complex surface meshes were occasionally too difficult to be understood and could potentially cause visual clutter, one of the major problems in AR, in more crowded scenes. Simplifying and selectively displaying meshes or removing unimportant real-world objects from the scene as described in SceneCtrl [24] would be beneficial to alleviate such problems. Classic and well-studied methods, such as adjusting object layouts, texture patterns, color opponency, and illumination, could also be used to mitigate the issues of the cluttered scene [10]. Moreover, adding the textures of the corresponding objects would be another promising direction to improve WSM in the future.

In terms of user study, we only evaluated WSM using a single *wall plane* given that the target positions were already known. Since recognizing the target locations in the reduced space while moving the objects and scaling the space using multiple *wall planes* are also crucial, we plan to conduct another experiment that includes such part in the near future. Furthermore, some insights obtained from the study might be too specific to the direct HOMER technique. We hope to empirically evaluate how WSM could improve other techniques in the future, such as the indirect HOMER technique, which uses external controllers, and touch gestures in mobile AR applications.

CONCLUSION

We presented that WSM enables the user to dynamically scale down the surrounding space for the efficient placement of virtual objects in an indoor environment using AR. Our prototype extracts room boundaries to create interactive *wall planes* and generates surface meshes of real objects as reference points from reduced space to real space. We discussed a set of design considerations and insights that were unique to the remote AR interaction. Finally, our user evaluation showed that WSM performed consistently well and significantly improved the efficiency of the direct HOMER technique for moving objects forward and backward while alleviating arm fatigue.

ACKNOWLEDGMENTS

This work was supported by Samsung Electronics Co., Ltd.

REFERENCES

1. Ferran Argelaguet, Carlos Andujar, and Ramon Trueba. 2008. Overcoming Eye-hand Visibility Mismatch in 3D Pointing Selection. In *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology (VRST '08)*. ACM, New York, NY, USA, 43–46. DOI: http://dx.doi.org/10.1145/1450579.1450588
2. Claudia Armbrüster, Marc Wolter, Torsten Kuhlen, Will Spijkers, and Bruno Fimm. 2008. Depth perception in virtual reality: distance estimations in peri-and extrapersonal space. *Cyberpsychology & Behavior* 11, 1 (2008), 9–15.
3. Doug A. Bowman and Larry F. Hodges. 1997. An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics (I3D '97)*. ACM, New York, NY, USA, 35–ff. DOI:http://dx.doi.org/10.1145/253284.253301
4. John Brooke and others. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
5. Volkert Buchmann, Stephen Violich, Mark Billinghurst, and Andy Cockburn. 2004. FingARtips: Gesture Based Direct Manipulation in Augmented Reality. In *Proceedings of the 2Nd International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia (GRAPHITE '04)*. ACM, New York, NY, USA, 212–221. DOI: http://dx.doi.org/10.1145/988834.988871
6. Taejin Ha, Steven Feiner, and Woontack Woo. 2014. WeARHand: Head-worn, RGB-D camera-based, bare-hand user interface with visually enhanced depth perception. In *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on*. IEEE, 219–228.
7. Don Hannema. 2001. Interaction in virtual reality. *Interaction in Virtual Reality* (2001).
8. Whie Jung, Woojin Cho, Hayun Kim, and Woontack Woo. 2017. BoostHand: Distance-free Object Manipulation System with Switchable Non-linear Mapping for Augmented Reality Classrooms. In *Mixed and Augmented Reality (ISMAR-Adjunct), 2017 IEEE International Symposium on*. IEEE, 321–325.
9. Max Krichenbauer, Goshiro Yamamoto, Takafumi Taketomi, Christian Sandor, and Hirokazu Kato. 2017. Augmented Reality vs Virtual Reality for 3D Object Manipulation. *IEEE Transactions on Visualization and Computer Graphics* (2017).
10. Ernst Kruijff, J Edward Swan, and Steven Feiner. 2010. Perceptual issues in augmented reality revisited. In *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on*. IEEE, 3–12.
11. Joon Hyub Lee, Sang-Gyun An, Yongkwan Kim, and Seok-Hyung Bae. 2018. Projective Windows: Bringing Windows in Space to the Fingertip. In *Proceedings of the*

2018 CHI Conference on Human Factors in Computing Systems (CHI '18). ACM, New York, NY, USA, Article 218, 8 pages. DOI: http://dx.doi.org/10.1145/3173574.3173792

12. Microsoft. 2018a. Gestures - Mixed Reality. (2018). Retrieved April 3, 2018 from https://docs.microsoft.com/en-us/windows/mixed-reality/gestures.
13. Microsoft. 2018b. MixedRealityToolkit-Unity. (2018). Retrieved March 29, 2018 from https://github.com/Microsoft/MixedRealityToolkit-Unity.
14. Mark R Mine. 1995. Virtual environment interaction techniques. *UNC Chapel Hill CS Dept* (1995).
15. Mark R. Mine, Frederick P. Brooks, Jr., and Carlo H. Sequin. 1997. Moving Objects in Space: Exploiting Proprioception in Virtual-environment Interaction. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '97)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 19–26. DOI: http://dx.doi.org/10.1145/258734.258747
16. Tamara Munzner. 2014. *Visualization analysis and design*. AK Peters/CRC Press.
17. Benjamin Nuernberger, Eyal Ofek, Hrvoje Benko, and Andrew D. Wilson. 2016. SnapToReality: Aligning Augmented Reality to the Real World. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1233–1244. DOI: http://dx.doi.org/10.1145/2858036.2858250
18. Ivan Poupyrev, Mark Billinghurst, Suzanne Weghorst, and Tadao Ichikawa. 1996. The Go-go Interaction Technique: Non-linear Mapping for Direct Manipulation in VR. In *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology (UIST '96)*. ACM, New York, NY, USA, 79–80. DOI: http://dx.doi.org/10.1145/237091.237102
19. Ivan Poupyrev, T Ichikawa, S Weghorst, and M Billinghurst. 1998. Egocentric object manipulation in virtual environments: empirical evaluation of interaction techniques. In *Computer graphics forum*, Vol. 17. Wiley Online Library, 41–52.
20. Dong Woo Seo and Jae Yeol Lee. 2013. Direct hand touchable interactions in augmented reality environments for natural and intuitive user experiences. *Expert Systems with Applications* 40, 9 (2013), 3784–3793.
21. Richard Stoakley, Matthew J. Conway, and Randy Pausch. 1995. Virtual Reality on a WIM: Interactive Worlds in Miniature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '95)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 265–272. DOI: http://dx.doi.org/10.1145/223904.223938
22. R. Troy Surdick, Elizabeth T. Davis, Robert A. King, and Larry F. Hodges. 1997. The Perception of Distance in Simulated Visual Displays:A Comparison of the Effectiveness and Accuracy of Multiple Depth Cues Across Viewing Distances. *Presence: Teleoperators and Virtual Environments* 6, 5 (1997), 513–531. DOI: http://dx.doi.org/10.1162/pres.1997.6.5.513
23. Curtis Wilkes and Doug A. Bowman. 2008. Advantages of Velocity-based Scaling for Distant 3D Manipulation. In *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology (VRST '08)*. ACM, New York, NY, USA, 23–29. DOI: http://dx.doi.org/10.1145/1450579.1450585
24. Ya-Ting Yue, Yong-Liang Yang, Gang Ren, and Wenping Wang. 2017. SceneCtrl: Mixed Reality Enhancement via Efficient Scene Editing. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 427–436. DOI: http://dx.doi.org/10.1145/3126594.3126601