Comparison of Device-Based, One and Two-Handed 7DOF Manipulation Techniques

Cite As:

Jinbo Feng; Isaac Cho, Zachary Wartell. Comparison of Device-Based, One and Two-Handed 7DOF Manipulation Techniques Proceedings of the 3rd ACM Symposium on Spatial User Interaction, ACM, August 2015, pages 2-9, [doi: 10.1145/2788940.2788942]

BibTex:

```
@InProceedings{Feng2015,
  author    = {Feng, Jinbo and Cho, Isaac and Wartell, Zachary},
  title     = {Comparison of Device-Based, One and Two-Handed 7DOF Manipulation Techniques},
  booktitle = {Proceedings of the 3rd ACM Symposium on Spatial User Interaction},
  year      = {2015},
  series    = {SUI '15},
  publisher = {ACM},
  location  = {Los Angeles, California, USA},
  isbn      = {978-1-4503-3703-8},
  pages     = {2--9},
  doi       = {10.1145/2788940.2788942},
  url       = { http://doi.acm.org/10.1145/2788940.2788942 },
  acmid     = {2788942},
  address   = {New York, NY, USA},
  groups    = {multi-scale_UI},
  keywords  = {3d vision, 7 degree-of-freedom, fish tank vr, virtual object manipulation, virtual reality},
  numpages  = {8},
}
```

Comparison of Device-Based, One and Two-Handed 7DOF Manipulation Techniques

Jinbo Feng
Charlotte Visualization Center
UNC Charlotte
9201 University Blvd.
Charlotte, NC, USA
jfeng3@uncc.edu

Isaac Cho
Charlotte Visualization Center
UNC Charlotte
9201 University Blvd.
Charlotte, NC, USA
icho1@uncc.edu

Zachary Wartell
Charlotte Visualization Center
UNC Charlotte
9201 University Blvd.
Charlotte, NC, USA
zwartell@uncc.edu

ABSTRACT

We evaluate three bimanual 7 degree-of-freedom (7DOF) object manipulation techniques that use a pair of precision grasped isotonic devices called buttonballs. 7DOF manipulation means changing position, orientation and scale. We compare the techniques in a (stereo) Fish-tank Virtual Reality (VR) system. The user study displays multiple randomly located boxes of different sizes and the user must dock (i.e. align) each target box with an objective box at the screen center. Comparing task completion times shows that in cases where target and objective boxes are the same size, all three techniques perform equivalently. When the sizes differ--requiring a scale change--two of the technique's, Spindle+Wheel and a minor variant of Grab-and-Scale perform similarly, and are both faster than the third technique, One-Hand+Scale. We compare and contrast our results with other work including free-hand versus held device input and also with 7DOF object manipulation versus 7DOF view manipulation.

Categories and Subject Descriptors

H.5.2 [**Information interfaces and presentation**]: User Interfaces –*Method devices and strategies.*

General Terms

Performance, Design, Human Factors

Keywords

3D vision, 7 Degree-of-Freedom, virtual reality, Fish tank VR, virtual object manipulation

1. INTRODUCTION

Many interaction techniques have been developed for 3D manipulation and navigation [1]. Among others, these involve single and bi-manual 2D input devices, multi-touch, 6 degree-of-freedom (6DOF) isotonic tracked held devices (or "props") and 3D tracked hands and fingers using various technologies. Common 3D interactions are 6DOF manipulation and navigation. However, 7 degree-of-freedom (7DOF) interaction is important as well. For object manipulation, this means including position, orientation plus scale. For navigation, this means treating view scale as a separate 7th DOF [15]. The latter becomes important in multi-scale virtual environments that use technology that makes differing view scales discernible, in which case the view scale parameter value can strongly effect usability [22] [15].

For 3D user interfaces, besides the various application requirements that influence the choice of input devices, a key issue is the mapping from the input device's DOF's to the manipulated object's DOF's. Depending on the device technology and mapping design, this might allow all object 7DOF's to be manipulated simultaneously or might allow only a subset of the 7DOF's to be manipulated at a given time using different input modes.

In a docking task, the user must align a target 3D object with an objective object [25]. A common object shape is a tetrahedron. In 6DOF docking the target and object are the same size, while in the 7DOF docking they differ in size.

For a 6DOF docking task Masliah et al. [10] find that users tend to allocate their control to the rotational and translational DOF's separately and switch control between the rotating and translating. With training, allocation of control within the translational and rotational subsets increases at a faster rate than across all 6 DOF's together. Their results suggest that the simultaneous manipulation of all DOF's does not necessary lead to the fastest completion time.

This paper evaluates 3 input techniques using isotonic 6DOF devices [25] for a 7DOF object manipulation task, i.e. manipulating the Euclidean DOF's, xyz+yaw,pitch,roll, plus scale. Our goal is to compare a technique that always engages all 7DOF against a technique where the user can choose whether to engage only pose or to engage pose and scale together. We refer to the former as "simultaneous 7DOF" and the latter as "separated 6DOF+Scale". Our study is performed in a stereo Fish-Tank virtual reality (VR) [23] (*also* Desktop VR [6]) environment with precision-grasped, 6DOF button balls. We perform a user study with 12 participants comparing the performance among the following three techniques:

- One-Hand+Scale [23] [15] – an uni-manual, separated 6DOF with Scale technique
- Spindle+Wheel [4] – a bi-manual, simultaneous 7DOF technique
- "One-Handed with Two-Handed Scaling" [5] – a bi-manual, separated 6DOF+Scale technique. (Note this technique is a minor variant of Grab-and-Scale [5]).

Our choice of these 3 techniques is discussed in Section 2.

The study's manipulation tasks include conditions that both require and do not require scale adjustment. Overall, we find that when users do not have to change scale, all three techniques performed equivalently. If users do have to scale, Spindle+Wheel and Grab-and-Scale perform similarly, but better than One-Hand+Scale. Some of these results are consistent with our detailed hypothesis and prior related work [9] [12] [4]; others are not. We compare and contrast our results to the others. We discuss result differences between 7DOF object manipulation versus 7DOF travel (i.e. view manipulation) and discuss free-hand [12] versus held device input and conclude with some guidelines.

2. BACKGROUND

7DOF interaction can control object manipulation, i.e. location (*x*,*y*,*z*), orientation (yaw, roll, pitch) and scale. 7DOF interaction also occurs in multi-scale virtual environments where the view scale factor is discernible as a 7^{th} degree-of-freedom due to stereo parallax, head-coupled display's motion parallax and direct 3D manipulation [24] [22] [15]. Extensive multi-scale VE's require dynamic adjustment to this 7^{th} view DOF. This study is limited to a 7DOF object manipulation task rather than a 7DOF travel task. But any position control 7DOF object manipulation technique can be adapted to 7DOF travel with the scene-in-hand metaphor [2].

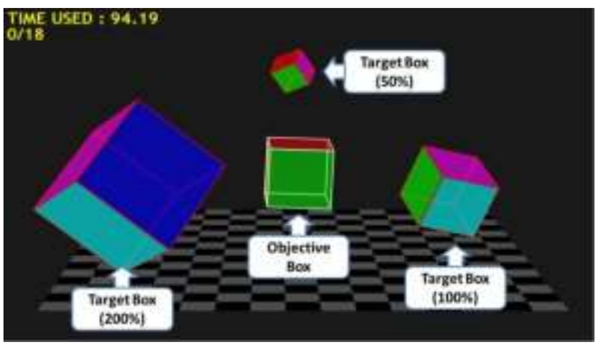

Figure 1: Screen capture of virtual environment displayed on desktop VR system in the experiment. The white frame objective box locates in the screen center with red frame target boxes around

A standard protocol for comparing object manipulation techniques is having users repeatedly perform a "docking" task [25]. As brief review, we introduce our experiment's particular 7DOF docking protocol. (Section 3 fully details our protocol and system hardware). Our virtual environment contains target boxes with 3 different fixed sizes at random locations in each trial. An objective box appears at fixed size at the screen center a random orientation in each trial (Figure 1). The user's task is to select each target box and manipulate it to align with the objective box. Like colored faces must be matched.

A B

Figure 2: (A) Button balls and (B) corresponding spherical cursors; left cursor is blue and right is pink.

In our experiment, the user holds isotonic input devices, in particular 6DOF tracked buttonballs [16] [25] [20] (Figure 2A). Each buttonball has a corresponding (stereoscopic) 3D cursor in the virtual environment (Figure 2B).

Next, we review prior interaction techniques suitable for 7DOF manipulation and present a simplified taxonomy. Our experiment only compares position control techniques therefore the review is limited to position control. It is outside our scope to review the pluses and minuses of position, rate, and acceleration control [2]. Our review is organized based on the 3 interaction techniques that we use in our experiment.

When discussing bimanual techniques, we use Guiard's Kinematic Chain Theory terminology [7]. For brevity, the rest of this document assumes the user is right handed. In the actual user studies, the roles of the cursors would be reversed for a left handed user.

2.1 Hand+Fingers vs Held Devices

Isotonic 6DOF input can be controlled by either 6DOF held devices or by hand and fingers tracking (abbreviated *hand+fingers*) [2]. Hand and finger tracking can be implemented in multiple ways. 6DOF tracked "data gloves" track the hand's 6DOF pose and detect each individual finger's phalange's pose. 6DOF tracked pinch gloves track the hand's 6DOF pose and detect combinations of finger pinches. Cordless marker based tracking is possible [13] as well as "free hand" tracking that requires no markers [21].

Hand+fingers and held devices each have advantages and disadvantages. The ideal hand+finger tracking implementation is a free-hand one and the ideal held device tracking is completely untethered (no wires) with no obtrusive markers sticking out of the device on stalks, etc. With current technology, achieving these ideals involve compromises on tracking precision, tracking volume, and occlusion.

Moehring and Froehlich [13] present an extensive study comparing a close-to-ideal free-hand implementation to a close-to-ideal held-device implementation used for several uni-manual docking tasks. The same tracking technology by A.R.T. GmbH is used over all conditions. Their general conclusion is:

> "finger based interaction is generally preferred if the functionality and ergonomics of manually manipulated virtual artifacts has to be assessed. However, controller-based interaction is often faster and more robust"

The passive-haptic feedback of held devices appears beneficial. Of course, some application domains require free hand 3D input such as surgeons in an operating room.

2.2 A Simplified Taxonomy

There are many uni-manual and bi-manual techniques suitable for 7DOF manipulation. Many have been formally evaluated. We will compare results of our study to the result of other studies. To help compare and contrast the range of techniques, Table 1 (next page) presents a taxonomy. Techniques grouped in a single row use the same mapping of input DOF's to the manipulated object DOF's. The *Original* column indicates the earliest usage of the mapping. The *H* column indicates the number of hands used. The *DOF* column indicates the number of DOF's available, denoted '*e*+*s*' where *e* are the number Euclidean DOF's and *s* indicates number of scale DOF's. The *Free-Hand* and *Held Device* column separate techniques that use free-hand tracking from those that use held devices. The *C.* column encodes various classifications (*C.* abbreviates "**C**lassification"). '–' means the technique cannot simultaneously adjust scale and pose, while + means the technique can simulatenously adjust pose and scale. '*' means a single button/pinch engages all supported DOF's at once. '**' means that a two button/pinch's in sequence are needed to engage pose and scale. The first button/pinch engages pose and a second button/pinch engages scale (while the pose adjustment remains engaged). Potentially, by requiring two button/pinch's the '+,**' class techniques may operate slightly slower than the '+,*' class techniques. 'A' vs 'S' indicate whether a bimanual technique is symmetric or asymmetric [7]. As we review prior work, we will discuss each row of the table.

2.3 One-Handed with Scale (OH)

	Original	H	DOF	C.	Free-Hand	Held Device
1	One-Hand w/ Scale [15] [23]	1	6,1	-,**		**OH**
2	5DOF+Scale [9] (glove)	2	5+1	+,* S	Handle-Bar [17]	TC[1] [20] Spindle [4] [14][2]
3	Spindle+Wheel [4]	2	6+1	+,* ≈S		**Spindle+Wheel**
4	Grab-and-Scale [5] (glove)	2	6+1	+,** A	6DOF-Hand [12]	**OTS** HIM [20]
5	Air TRS [12] (free hand)	2	5+1	+,** A		
6	3DOF-Hand [12] (free hand)	2	6+1	+,** A		

Legend:

+ can scale simultaneously, – cannot scale simultaneously
***** scale always enabled, ****** scale engaged with 2nd pinch or button
S symmetric, A asymmetric (Guiard classification **[7]**)
Conditions in our experiment are **Bold**

Table 1: Rows are techniques with the same DOF mapping. However, within a row the input technology and scale engagement conditions may vary.

Our experiment includes a one-handed isotonic technique, called "One-Handed with Scale technique" (abbreviated OH). OH works as follows. For translation and rotation, the user presses and holds the selection button after placing the cursor inside the target box [23] [15]. Then the box movement has been attached to the cursor and with the cursor center as the rotation center. To scale the target box, the user places the cursor inside the box and presses and holds a second scaling button. To scale up and down, the user moves the cursor hand toward or away from the screen. Scaling is controlled by rate control. Prior work suggests for this one-hand 7DOF technique users prefer rate control for scale compared to position control [4].

Within the taxonomy, the One-Hand with Scale (OH) appears in Table 1, row 1. 6DOF pose and scale cannot perform simultaneously and use two buttons, hence OH is '6,1' and '-,**'. OH is not bimanual and has no symmetric/asymmetric Guiard classification.

[1] TC ('Two-Corners') is actually 5+3 (separate xyz scales), but a 5+1 variant is possible

[2] Original Spindle uses power grip SpaceGrips [13]

2.4 Spindle+Wheel (S+W)

Our experiment's first two-handed manipulation condition is Spindle+Wheel [4]. Spindle+Wheel extends prior work [9] [5] [14] by allowing all 7DOF to be manipulated simultaneously with one button.

Figure 3A shows the Spindle+Wheel visual feedback. A thin orange cylinder, the "spindle", is drawn between the two cursors [14] with a small red sphere at the mid-point. The "wheel" [4] is a disc on the right cursor indicating the plane of rotation for the pitch rotation.

To select an object, the user places the spindle's center inside the target object and presses and holds the select button on the left button ball. This engages object manipulation [14] which works as follows and is illustrated in Figure 3B. Rotating one hand about the other while keeping their distance constant, rotates the selected box in yaw and roll, moving the hands closer or farther apart scales the box, while translating the hands rigidly translates the object [9]. Figure 3B illustrates the difference between Spindle+Wheel and the earlier Mapes and Moshell's 5DOF+Scale technique (re-used by Spindle [14]). The 5DOF+Scale technique provides all the illustrated DOF's except the ability to pitch -- the green arc with arrow in Figure 3B. Spindle+Wheel supports pitch by spinning or twisting the right button ball with the fingers around wheel axis. This rotates the selected object around the spindle axis [4].

Cho and Wartell demonstrate that Spindle+Wheel yields faster completion times than the 5DOF+Scale approach for a 7DOF docking task [4]. Spindle+Wheel is "mostly" symmetric in that all but the pitch DOF are bimanually symmetric.

A

B

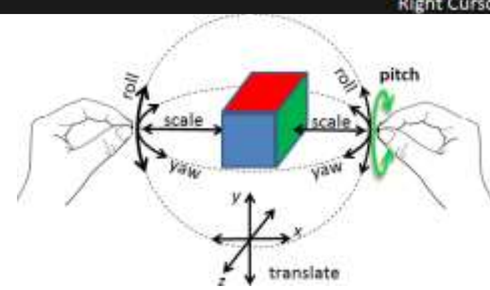

Figure 3: Spindle+Wheel condition. (A) Spindle+Wheel's visual feedback. (B) Bimanual DOF's of Mapes and Moshell "5DOF+Scale" technique [9] versus "Spindle+Wheel" [4] which adds pitch (green).

5DOF+Scale and Spindle appear in Table 1, row 2. All row 2 techniques are 5+1 DOF, engage all DOF's at once with one button/pinch ('+,*') and are symmetric ('S'). Spindle+Wheel in row 3. Song et al. [17] present a free-hand "Handle-Bar". By our observation Handle-bar replicates 5DOF+Scale [9] with nearly identical visual feedback of Spindle [14]. (Song et al. do not mention the replication).

2.5 One-Handed with Two-Hand Scaling (OTS)

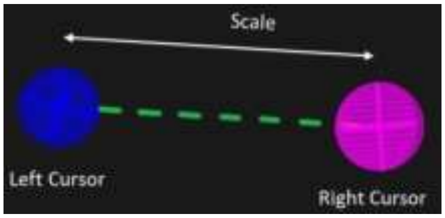

Figure 4: Graphical representation in scaling manipulation

Masliah et al.'s [10] result suggests a possible advantage to separating scale control into a separate second mode. This motivates evaluating a third technique called "One-Handed with Two-Hand Scaling" (OTS). In OTS, the left hand button ball operates the same as the One-Hand technique (OH) for translation and rotation by pressing a left cursor button. Additionally pressing a right cursor button engages scaling. In OTS, scale-mode displays a dotted green line between the two cursors (Figure 4). The scale is adjusted based on any ensuing change in distance between the cursors; this is similar to Spindle+Wheel.

The DOF mapping of OTS is the same a Culter et. al's bi-manual "Grab-and-Scale" (Table 1, row 4), but Grab-and-Scale uses tracked pinch gloves while OTS uses buttonballs. Also, unlike Grab-and-Scale, OTS also includes the deliberate translation offset between input device and cursor to counter fatigue [16]. OTS also essentially uses the same 7DOF mapping as the Hand-in-Middle (HIM) technique [20] and uses equivalent input devices, however, the target task differs.

We will compare and contrast our study's results to Mendes et al.'s [12] user study. So we briefly review their techniques and relate them to the ones we evaluate. Mendes et al. [12] present a bimanual free-hand technique called "6-DOF Hand" (Table 1, row 4). By our observation, 6-DOF Hand is the same as Grab-and-Scale but uses free hand tracking instead of tracked gloves. Hence, 6-DOF Hand and OTS differ in that OTS has an offset and uses buttonballs and two buttons, while 6-DOF-Hand is free-hand with two pinch gestures but no offset.

Mendes et al. [12] also present "Air TRS" (Table 1 row 5). Like 5DOF+Scale, Air TRS controls 5+1 DOF's, xyz-yaw-roll+scale, but while 5DOF+Scale is symmetric, Air TRS is asymmetric. In particular the center of scale and rotation is always the left hand. Air TRS appears to be a hybrid of 5DOF+Scale and Grab-and-Scale.

Mendes et al. also present 3DOF-Hand. It provides simultaneous 7DOF. Scaling works as in 5DOF+Scale. Translation is controlled by the initiating (left) hand but rotation is controlled by mapping the secondary hand's orientation directly to the selected object (rotation gain is 1).

3. EXPERIMENT

Section 2 introduced our experiment's protocol as part of a review of docking experimental protocols. Now we present our experimental design in detail.

The virtual environment is displayed in a stereoscopic, Fish-tank VR configuration [23]. The environment has a checker-board ground-plane (Figure 1). It is 40 cm square with half appearing behind the display surface and half appearing in front. In the center of the screen is a translucent box, the Objective Box, of fixed size and at a random orientation per trial. Each face has a different color. This cube's pose remains stationary relative to the display screen during target box manipulation. At each trial, three target boxes with 50%, 100% and 200% of the objective box's size appear at random locations and orientations on the ground-plane. The user must select the target boxes one by one and align the target cube with the objective cube. Like colored faces must match. This requires object rotation, translation, and scaling to match the sizes.

When the distance between the target cube's corresponding vertices is within a tolerance (0.84 cm) of the objective cube's vertices, the frame of the cube turns green. If the target box is selected and kept green for 0.8 seconds [25], it will disappear and a success sound will be played indicating one docking operation is complete. The user then proceeds to dock the next target box. After docking the three target boxes, one trial has been completed and the system automatically generates three new target boxes for the next trial.

Once for each user (at the start of her session), the user holds the button balls and rests her elbows on the chair's arms and the experimenter sets a translational offset that places the 3D cursors in the center of the screen. This is designed to maximize the degree to which the user rests her elbows during interaction [16].

As mentioned, we evaluate three technique conditions with buttonballs devices: One-Hand with Scale (OH), Spindle+Wheel (S+W) and One-Hand with Two-Handed Scaling (OTS) crossed with 3 box sizes (50%, 100% and 200%). We also ported Spindle+Wheel to a free-hand implementation but the tracking results of commodity systems were not robust enough to perform a useful free-hand versus held-device comparison. (We discuss the specific difficulties encountered in Section 5.4).

Our experiment is within subjects. Participants perform all 3 techniques and encounter all box sizes. For each technique condition, there are two blocks: a training block followed by a 6 trials of experiment block (18 dockings total). Presentation order of manipulation technique condition was counter-balanced between participants. All participants successfully finished the study in 80 minutes.

The system hardware uses Nvidia 3D Vision glasses with Nvidia Quadro 2000 and a 120Hz 22" LCD monitor. The position of button balls and user's head are tracked by a Polhemus Fastrak. Software is written in OpenSceneGraph, VRPN [19] and an in-house C++ integration API [18] .

Twelve unpaid students from the Computer Science and Computer Engineering departments with little or no experience using 3D computer graphic applications participated in the study. Participants were required to tell the distance differences of different boxes in the scene and distinguish the colors of the box frames and faces.

Our hypotheses are:

- **H1:** For 100% box size, OTS and OH will outperform S+W because S+W always engages scale and for the 100% box size which adds extra mental and physical effort to avoid changing the box size.

- **H2:** For the 50% and 200%, OTS and S+W will perform faster than OH because they allow simultaneous control of scale, while OH requires switching between 6DOF and Scale mode.
- **H3:** For the 50% and 200%, S+W will perform faster than OTS because it only requires engaging a single button to engage scale, while OTS requires the user alternatively engage and disengage the secondary scale button.
- **H4:** Overall, completion times for the 100% box will be faster than for the 50% and 200% percent cases due to the lack of need to scale.
- **H5:** Overall box sizes, OTS and S+W will outperform OH, because from **H3** they should dominate the scaling sconditions and our protocol has 2 scaling conditions and 1 non-scaling condition.

4. RESULTS

We next present our quantitative and subjective results. All significant results are p=0.05.

4.1 Quantitative Results

We use a 3×3 repeated measures ANOVA and use interaction technique presentation order as the between-subjects factor. Manipulation technique condition has value set {*One-handed (OH), Spindle+Wheel (S+W), One-handed with Two-handed Scale (OTS)*} and target box size has value set {*50%, 100%, 200%*}. These two variables are manipulated within participants.

There is a significant interaction effect on task completion time (F(4,44)=9.486, p<.001, η_p^2=.463). Hence we examine the simple effects for each box size.

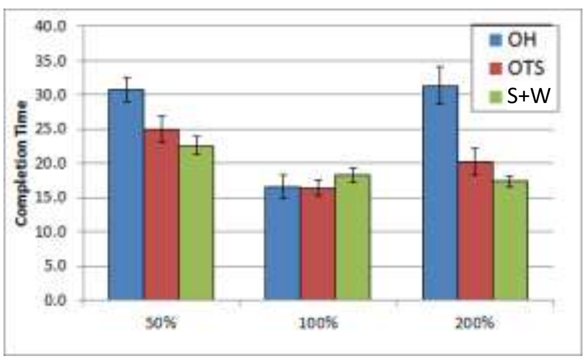

Figure 5: Completion time for different box sizes: 50%, 100% and 200% by different manipulation techniques

There is a significant simple effect for 50% box size (F(2,22)=10.908, p<.005, η_p^2=.498). LSD comparisons show OTS (M=24.99, SD=6.72, p<.005) and S+W (M=22.56, SD=4.52, p<.005) has faster completion time than OH (M=30.76, SD=6.09). This partially confirms **H2** (for sub-case 50%). However the prediction that S+W is faster than OTS for 50% (**H3)** is not supported.

There is also a simple effect on 200% box size (F(2,22)=18.883, p<.001, η_p^2=.632). LSD comparison show OTS (M=20.2, SD=6.71, p<.005) and S+W (M=17.34, SD=2.62, p<.001) has faster completion time than OH (M=31.33, SD=9.36). This completes confirmation of **H2.** Again, however the prediction that S+W is faster than OTS for 200% (**H3)** is not supported.

The simple effect of 100% box is not significant. This is surprising as it fails to support **H1,** that OTS and OH will outperform S+W. Section 5.1 elaborates on possible reasons for this in context of differing results found by others for Spindle+Wheel for 7DOF travel.

For the main effects, there exists significant difference among different box sizes (F(2,22)=30.074, p<.001, η_p^2 =.732). LSD comparisons show that 100% box size (M=17.1, SD=4.27) has faster completion time than 50% (M=26.1, SD=6.38, p<.001) and 200% box size (M=22.9, SD=8.7, p<.001). This confirms **H4**.

In addition, 200% box size has faster completion time than 50% box size (p<.05). This was unexpected, however, it corroborates Cho and Wartell's [4] study of bi-manual, 7DOF navigation. They suggest users found it easier to bring hands together (scale down) than to brings hands apart (scale up) from an initial parallel position. Section 5.2 elaborates on this issue.

Finally, there is a significant main effect on task completion time for interaction techniques (F(2,22)=16.195, p<.001, η_p^2=.596). The LSD comparisons indicate that the completion time of both OTS (M=21.2, SD=7.33, p<.005) and S+W (M=21.85, SD=5.06, p<.001) are faster than OH (M=27.08, SD=8.82). This confirms **H5**.

4.2 Subjective Results

When asked which interaction technique (OH vs. S+W) is better for rotation, six participants answered S+W and the other half answered OH. Eight answered OH is better than S+W for translation, 3 answered both are equivalent, and one answered S+W. When asked which technique is better for scaling, four answered OTS, three rated S+W, three rated OH, one answered OH and OTS are equivalent, and one answered S+W and OTS are equivalent.

Regarding the question about which technique is most intuitive, five answered OTS, four answered S+W, two answered OH and one rated OH and OTS equivalent.

Overall, seven of twelve participants preferred the OTS, three rated S+W and one rated OTS and S+W equivalent. Participants rated arm fatigue after finishing the experiment for each interaction condition (on a 7-point Likert scale, 1 *no fatigue* to 7 *very painful*). There is no significant main effect on arm fatigue rate for interaction condition ($\chi^2(2)$=.054), rates were OH=1.96, OTS=2.0 and S+W=2.04.

5. DISCUSSION

In this section, we compare our results to other related studies and hypothesize reasons for differences.

5.1 Compare to Spindle+Wheel for 7DOF Travel

H1 predicted for 100% box size that both OTS and OH will outperform S+W because S+W always engages scale and in the 100% box size this requires additional physical and mental efforts to maintain a constant scale factor. However, the results do not support **H1** (i.e. OTS, OH and S+W perform the same). In contrast, Cho and Wartell found that OH and a modified Spindle+Wheel (that separated scale with a secondary button) both did perform faster than Spindle+Wheel for trials where scale change was not needed. However, their experiment explores 7DOF travel, adapting Spindle+Wheel using the scene-in-hand metaphor, not 7DOF object manipulation. This may explain the differing outcomes as follows.

Our manipulation task requires a selection step; the manipulation cannot be engaged until the cursor is inside a target box. For Cho and Wartell's 7DOF travel, the travel user interaction is engaged immediately upon button press without requiring the cursor to be inside the target. In the 100% case, their average completion times were 10.6s, 12.1s and 15.5s for OH, SWS, and S+W respectively. For us, in the 100% case, the averages are 16.4s, 16.5s and 17.34s for OTS, OH and S+W. The increase may be explained by the extra time required to move the cursor to the target box for a selection step. Possibly on average this adds an equal increment across all three conditions, leading to a lesser overall percentage difference between interaction techniques and hence lack of significant performance difference between OTS and S+W for 100%. Perhaps repeating our experiment, but as a 7DOF travel task, would find OTS performs better than S+W for the 100% box size.

5.2 200% vs 50% Box Size

The 200% box size has faster completion times than 50% box size for bimanual conditions. This corroborates Cho and Wartell's [4] study of bi-manual 7DOF travel. They suggest that in bi-manual tasks users found it easier to scale down (bring hands together) than to scale up (brings hands apart).

We add a literature review on shoulder range-of-motion that is insightful. McCully et al. [11] study the range-of-motion (ROM) of shoulder rotation in healthy adults. "Shoulder rotation" is the anatomical adjustment that swings the lower arm at the elbow as shown in Figure 6A (next page). McCully et al. examine ROM for internal and external rotation directions when the shoulder is held at 4 different alignments. Figure 6A, B and C sketch three of their tested positions and show the mean ROMs. Internal rotations are positive and external rotations are negative. Figure 6D is our estimate of the typical internal and external rotation limits when the elbow is resting on a chair arm or table in our experiment. This is based on our anecdotal observations in our experiment and interpolating McCully et al.'s results for our typical shoulder alignment. We observe the external ("outward") rotation range is more limited than internal ("inward") rotation range which is consistent with the ROM literature. As the shoulder alignment grows farther away from being aligned with the Y-axis – where the ROM is greatest – the external rotation limit decreases. We find that in bimanual interaction in seated Fish-Tank VR, shoulder alignment ranges between situation A, B and D. In our 3D user interface, scaling up requires bringing the hands apart (external rotation) which is less comfortable than bringing them together (internal rotation) to scale down. We suspect these anatomical issues explain both our and other researchers' observation that bimanually scaling up a small object is slightly slower than scaling down a large object because the range limits cause more clutching in the former.

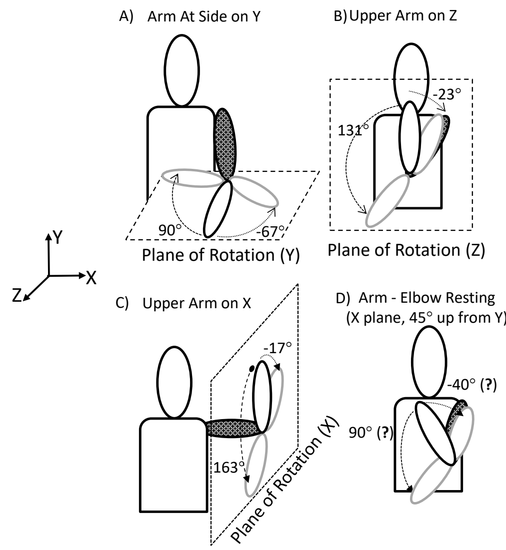

Figure 6: (A), (B) and (C) are shoulder rotation limits (internal >0; external < 0) for healthy adults with upper arms at various positions [11]. (D) is our estimate of the comfortable ROM when the elbow rests on a table as in our Fish-Tank VR environment – this is based on our observations and interpolating results of [11].

5.3 Compare to Prior Free-Hand Techniques

Next, we compare our results to Mendes et al. Both our and their virtual environment neither simulates gravity nor prevents object collision. Mendes et al. do not have a translation offset between the hand and the cursor; they use a horizontal stereo, head-tracked display (*aka* Responsive Workbench) and the user stands. All our conditions include a translation offset to reduce fatigue [16]; the screen is vertical, and the user sits.

Mendes et al.'s 6DOF-Hand, Air TRS and Handle-Bar techniques' relation to Grab-and-Scale and 5DOF+Scale are reviewed in Table 1 and Sections 2.4 and 2.5. A primary difference between Mendes et al.'s and our study's techniques is their use of free-hand tracking versus our use of buttonballs.

Mendes et al.'s Task #1 is a 2DOF task requiring 2D translation (dropping a ball into a solid cube with a bored out cylinder hole). In this case, 3DOF-Hand, 6DOF-Hand, Air TRS and Handle-Bar did not significantly differ; moreover the first 3 techniques only engage one hand for this task. Compare this to our 100% box trials (a 6DOF task) where OH, S+W and OTS did not significantly differ. A commonality across both their results and ours is that two-handed and one-handed techniques performed similarly when the tasks require manipulating less than or equal 6 DOF's.

Next, we compare Mendes et al.'s Task #2 to our 50% and 200% conditions (both tasks require scale adjustment). Mendes et al.'s Task #2 docks a donut into a donut shaped hole bored out of a face of a cube. Translation and size change is required but not rotation. Handle-Bar is faster than the others which performed similarly (6DOF-Hand, Air TRS, 3DOF Hand). In our 50% and 200% conditions scale change is required (7DOF). OTS and S+W perform similarly, but outperform OH. Note, Handle-Bar always simultaneously engages scale (one pinch) while 6DOF-Hand, Air TRS and 3DOF-Hand require a second pinch to engage scale. Again, scale is required for Task #2.

In our task, Spindle+Wheel, which always engages scale ('+,*'), is not faster than OTS, where scaling requires a secondary button ('+,**'). In Mendes et al's Task #2, their technique, which always engage scale ('+,*'), is faster than those that require a secondary scale engagement pinch ('+,**'). A possible hypothesis is that requiring a secondary scale engagement button ('+,**) slows performance compared to class '+,*', but due to ceiling effects it is not measurable for a more complex task, such as our Translate+Rotate+Scale, but is measurable for simpler (faster) tasks, such as Translate+Scale (Mendes et al.'s Task #2).

Next, we compare Mendes et al.'s Task #3 with our 50% and 200% conditions (both are full 7DOF). Mendes et al.'s Task #3 requires docking a half-cylinder into a half-cylinder shaped hole bored out of a cube. This is a 7DOF task but the half-cylinder can fit one of two ways, whereas our 7DOF task requires matching all like color faces – i.e. it can "fit" only one way. They find Handle-Bar and 6DOF-Hand perform similarly, but outperform Air TRS. Handle-Bar always engages scale ('+,*') while 6DOF-Hand and Air TRS require a secondary pinch ('+,**'). It does not appear the presence/absence of a secondary scale pinch alone can explain the time difference. Our 50% and 200% conditions also show no performance difference between the technique that always engages scale, Spindle+Wheel ('+,*' ; ≈Handle-Bar) and the technique that requires a secondary button/pinch to engage scale, OTS ('+,**' ; ≈6DOF-Hand). This suggests that for 7DOF object docking, the presence/absence of a required secondary scale button/pinch does not significantly affect completion time for either free-hand or held devices.

In both our and Mendes et al.'s results the Mapes-Moshell's 5DOF+Scaled derived rotation methods (Handle-Bar, Spindle+Wheel) completion time is not significantly different from the Grab-and-Scale derived methods (6DOF-Hand, OTS). [1]For rotation, Mendes et al. report that users prefer 6DOF-Hand over Handle-Bar. Similarly, our users overall preference are 7/12 preferred OTS compared to 3/12 for S+W. In both cases, more users prefer the bimanual technique where only one hand controls rotation (Table 1, row 4) over the bimanual technique where both hands control rotation (Table 1, row 2 & row 3). Mapes and Moshell's indicate their users preferred bi-manual rotation (5DOF+Scale) over a one-handed rotation [23], but they do not cite specific questionnaire results.

In Task #3, Mendes et al. suggest Handle-Bar outperforms 3DOF-Hand and Air TRS not because of different DOF mappings, but because the selection point (Handle-Bar/Spindle center) is not occluded by the user's hands. But Handle-Bar and 6DOF-Hand do not perform significantly different and 6DOF-Hand does have the occlusion issue. In our case, however, due to the translation offset none of our conditions co-locate the buttonball (and hence hand) with the 3D cursors and we still find Spindle+Wheel (≈Handle-Bar) performs similarly to OTS (≈6DOF-Hand). These results suggest the need to investigate whether adding translation offset versus no offset as an additional factor across all conditions would change the outcomes.

5.4 Free-Hand Spindle+Wheel

As mentioned in Section 3 we ported Spindle+Wheel to a free-hand implementation in our Fish-Tank VR. The user forms a U-shape with her index finger and thumb of the dominant hand to engage the pitch operation. The degree of pitch rotation comes from computing a line between the index finger tip and thumb tip and tracking the rotation of this line around the between-hand axis. Unfortunately, a Kinect + 3Gears' software implementation could not track the index and thumb robustly enough for a formal comparison against buttonball Spindle+Wheel.

We ported this interface to the Leap Motion as well and piloted a comparison of the buttonball versus free-hand Spindle+Wheel [3]. Unfortunately, the Leap Motion while better, but it is still not robust enough for formal evaluation. The U-gesture recognition rate was still significantly less than the "100% recognition rate" of a button press in the buttonball condition. We found roughly 75% recognition rate by the Leap Motion. (Others formally evaluating the Leap Motion found gesture recognition rates insufficient as well and resorted to substituting a second hand to press a keyboard key to engage their interaction techniques [2]). The tracking volume of the Leap Motion is also limited compared to the buttonballs' Polhemus. For buttonball Spindle+Wheel moving a buttonball by 1 cm moves the cursor by 1 cm. Free-hand Spindle+Wheel used the same C/D ratio. While performing the docking task it is fairly common to lose tracking while performing maneuvers equivalent to those with the buttonball because the hands exit the Leap's tracking volume. A user strategy of dividing attention between the docking task on the screen and keeping one's hands constrained in the Leap Motion tracking volume proved impractical. Using a different C/D ratio would change the distance the user maintains between her hands which alters the angular 'leverage' available when performing roll and yaw rotations. These issues prevented a fair comparison free-hand versus buttonball Spindle+Wheel. In the following section we will discuss our future plan.

5.5 Implications for Design and Future Work

Let, $T_1 \equiv T_2$, denote that the task completion time for techniques T_1 and T_2 are not significantly different and $T_1 > T_2$ denote T_1 performs better than T_2 (statistically significantly faster). Then we have the following for isotonic control (row *x* refers to Table 1):

(i) 7DOF object docking: [12]
Handle-Bar ≡ 6DOF-Hand
(free hand, row 2) ≡ (free hand, row 4)

(ii) 7DOF travel docking: [4]
Spindle+Wheel > Spindle
(held device, row 3) > (held device, row 2)

(iii) 7DOF object docking: (*this paper*)
Spindle+Wheel ≡ OTS
(held device, row 3) ≡ (held device, row 4)

(iv) 6DOF object docking: [13]
held device > free hand

Our literature review and result (iv) leaves the possibility of an interaction effect of hand+fingers versus held device implementation of a given bimanual technique and our Section 5.1 results indicate travel and object manipulation tasks that use the same DOF mappings can differ in performance outcomes. This suggests the following 4-way 7DOF experiment should be done:

EX1: {Grab-and-Scale, Spindle+Wheel}×
{hand+fingers, held device}×
{travel, object manipulation}×
{*multiple scale factors*}

From the issues we encounter and discussed in Section 5.4, experiment **EX1** must be done with a robust marker based tracking to allow proper and fair comparison between hand+finger and held device techniques [13].

1. Completion Time:
   a. Grab-and-Scale ≡ Spindle+Wheel (*this paper*)
   b. Grab-and-Scale ≡ Spindle [12]
2. "Average" user subjective preference-wise:
   a. Grab-and-Scale > Spindle+Wheel (*this paper*)
   b. Grab-and-Scale > Spindler

Because (2) is an individual preference, the best general isotonic 7DOF manipulation guideline is:

> **G1:** If satisfying each individual user's preference is of high importance to the interface designer, give the user the option of Spindle+Wheel or Grab-and-Scale derived methods; otherwise use Grab-and-Scale.

Finally Section 5.2, suggests the following expectation for any bimanual isotonic technique that uses the scale technique derived from 5DOF+Scale:

> **G2:** A seated user is likely to be slightly slower at scaling up than scaling down.
>
> Potentially a designer could choose a slightly higher gain factor for scaling up to combat this.

6. CONCLUSION

This paper evaluated 3 input techniques using isotonic 6DOF devices for a 7DOF manipulation task. Overall, we found that when users do not have to scale, all three techniques performed equivalently. If users do have to scale, Spindle+Wheel and One-Handed-with-Two-Hand-Scaling perform better than One-Hand with Scale but similarly to each other. Some users preferred one bimanual technique over the other. Similarities with Mendes et al.'s results lead to a tentative guideline, **G1**, for 7DOF bimanual isotonic interaction. However, our comparison with several other prior suggests several possible interaction effects which indicate need for the further experiment, **EX1**. We plan to perform this experiment and within its context to also analyze the distribution of users' allocation of the different 7DOF's [8].

7. ACKNOWLEDGEMENTS

This work was supported in part by grant W911NF0910241 (PN55836MA) from The U.S. Army Research Office.

REFERENCES

1 Bowman, Doug A., Kruijff, Ernst, LaViola, Joseph J., and Poupyrev, Ivan. *3D User Interfaces: Theory and Practice*. Addison Wesley, 2005.

2 Brown, Michelle A., Stuerzlinger, Wolfgang, and Filho, E. J. Mendon\c. The Performance of Un-instrumented In-air Pointing. In *Proceedings of the 2014 Graphics Interface Conference* (Toronto, Ont., Canada, Canada 2014), Canadian Information Processing Society, 59-66.

3 Cho, Isaac. *Bimanual Stereoscopic Interaction for 3D Visualization*. University of North Carolina at Charlotte, Charlotte, NC, USA, 2013.

4 Cho, Isaac. and Wartell, Zachary. Evaluation of a Bimanual Simultaneous 7DOF Interaction Technique in Virtual Environments. (Arles, France 2015), 3D User Interfaces (3DUI), 2015 IEEE Symposium on.

5 Cutler, Lawrence D., Froehlich, Bernd, and Hanrahan, Pat. Two-handed direct manipulation on the responsive workbench. In *Proceedings of the 1997 symposium on Interactive 3D graphics* (1997), ACM Press, 107--ff.

6 Deering, Michael. High resolution virtual reality. *ACM SIGGRAPH Computer Graphics*, 26, 2 (1992), 195-202.

7 Guiard, Yves. Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model. *Journal of Motor Behavior*, 19 (1987), 486-517.

8 Jacob, Robert J. K., Sibert, Linda E., McFarlane, Daniel C., and Mullen, Jr., M. Preston. Integrality and Separability of Input Devices. *ACM Trans. Comput.-Hum. Interact.*, 1, 1 (March 1994), 3-26.

9 Mapes, Daniel P. and Moshell, J. Michael. A Two Handed Interface for Object Manipulation in Virtual Environments. *Presence*, 4, 4 (1995), 403-416.

10 Masliah, Maurice R and Milgram, Paul. Measuring the allocation of control in a 6 degree-of-freedom docking experiment. (2000), Proceedings of the SIGCHI conference on Human Factors in Computing Systems.

11 McCully, Sean P., Kumar, Naveen, Lazarus, Mark D., and Karduna, Andrew R. Internal and external rotation of the shoulder: Effects of plane, end-range determination, and scapular motion. *Journal of Shoulder and Elbow Surgery* , 14, 6 (2005), 602-610.

12 Mendes, Daniel and Fonseca, Fernando and Araujo, Bruno and Ferreira, Alfredo and Jorge, Joaquim. Mid-air interactions above stereoscopic interactive tables. (2014), 3D User Interfaces (3DUI), 2014 IEEE Symposium on.

13 Moehring, M. and Froehlich, B. Effective manipulation of virtual objects within arm's reach. In *Virtual Reality Conference (VR), 2011 IEEE* (2011), 131-138.

14 Paul Mlyniec, Jason Jerald, Arun Yoganandan F. Jacob Seagull Fernando Toledo Udo Schultheis. Studies in Health Technology and Informatics. IOS Press, 2011.

15 Robinett, Warren and Holloway, Richard. Implementation of flying, scaling and grabbing in virtual worlds. In *Proceedings of the 1992 symposium on Interactive 3D graphics* (1992), ACM Press, 189-192.

16 Shaw, Chris and Green, Mark. Two-handed polygonal surface design. In *Proceedings of the 7th annual ACM symposium on User interface software and technology* (1994), 205-212.

17 Song, Peng, Goh, Wooi Boon, Hutama, William, Fu, Chi-Wing, and Liu, Xiaopei. A handle bar metaphor for virtual object manipulation with mid-air interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2012), 1297-1306.

18 Suma, Evan. osgVirtualEnvironment. http://sourceforge.net/projects/osgve/

19 Taylor II, Russell M., Hudson, Thomas C., Seeger, Adam, Weber, Hans, Juliano, Jeffrey, and Helser, Aron T. VRPN: a device-independent, network-transparent VR peripheral

system. In *Proceedings of the ACM symposium on Virtual reality software and technology* (New York, NY, USA 2001), ACM, 55-61.

20 Ulinski, Amy, Zanbaka, Catherine, Wartell, Zachary, Goolkasian, Paula, and Hodges, Larry F. Two handed selection techniques for volumetric data. In *3D User Interfaces, 2007. 3DUI'07. IEEE Symposium on* (2007).

21 Wang, Robert, Paris, Sylvain, and Popovic, Jovan. 6D hands: markerless hand-tracking for computer aided design. In *Proceedings of the 24th annual ACM symposium on User interface software and technology* (New York, NY, USA 2011), ACM, 549-558.

22 Ware, Colin, Gobrecht, Cyril, and Paton, Mark Andrew. Dynamic Adjustment of Stereo Display Parameters. *IEEE Transactions on Systems, Man and Cybernetics Part A: Systems and Humans*, 28, 1 (January 1998), 56-65.

23 Ware, Colin and Jessome, Danny R. Using the bat: A six-dimensional mouse for object placement. *Computer Graphics and Applications, IEEE*, 8, 6 (1988), 65-70.

24 Wartell, Zachary, Hodges, Larry, and Ribarsky, William. Third-Person Navigation of Whole-Planet Terrain in a Head-Tracked Stereoscopic Environment. In *Proceedings of IEEE Virtual Reality '99 Conference* (March 1999), IEEE Computer Society Press, 141-148.

25 Zhai, Shumin. Investigation of feel for 6DOF inputs: isometric and elastic rate control for manipulation in 3D environments. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (1993), 323-327.