Using Custom Transformation Axes for Mid-Air Manipulation of 3D Virtual Objects

Daniel Mendes
Maurício Sousa
danielmendes@ist.utl.pt
antonio.sousa@ist.utl.pt
INESC-ID Lisboa, Instituto Superior Técnico, Universidade de Lisboa

Rodrigo Lorena
rui.rodrigo.lorena@ist.utl.pt
Instituto Superior Técnico, Universidade de Lisboa

Alfredo Ferreira
Joaquim Jorge
alfredo.ferreira@ist.utl.pt
jorgej@acm.org
INESC-ID Lisboa, Instituto Superior Técnico, Universidade de Lisboa

ABSTRACT

Virtual Reality environments are able to offer natural interaction metaphors. However, it is difficult to accurately place virtual objects in the desired position and orientation using gestures in mid-air. Previous research concluded that the separation of degrees-of-freedom (DOF) can lead to better results, but these benefits come with an increase in time when performing complex tasks, due to the additional number of transformations required. In this work, we assess whether custom transformation axes can be used to achieve the accuracy of DOF separation without sacrificing completion time. For this, we developed a new manipulation technique, MAiOR, which offers translation and rotation separation, supporting both 3-DOF and 1-DOF manipulations, using personalized axes for the latter. Additionally, it also has direct 6-DOF manipulation for coarse transformations, and scaled object translation for increased placement. We compared MAiOR against an exclusively 6-DOF approach and a widget-based approach with explicit DOF separation. Results show that, contrary to previous research suggestions, single DOF manipulations are not appealing to users. Instead, users favored 3-DOF manipulations above all, while keeping translation and rotation independent.

CCS CONCEPTS

•**Human-centered computing** →**Interaction techniques;**

KEYWORDS

3D user interfaces, Mid-air object manipulation, Custom manipulation axis, DOF separation, Virtual reality

ACM Reference format:
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 VRST '17, Gothenburg, Sweden, November 8–10, 2017,* 8 pages.
DOI: 10.1145/3139131.3139157

VRST '17, Gothenburg, Sweden

978-1-4503-5548-3/17/11…$15.00
DOI: 10.1145/3139131.3139157

1 INTRODUCTION

Despite the increasing interest in Virtual Reality (VR) and its popularization in recent years, there are still several open challenges regarding interaction in such environments. As in any kind of interactive 3D environments, object manipulation is a key element. The spatial input typically associated with VR setups can offer natural metaphors, allowing users to grab, move and rotate objects in a similar way to how it is done in the physical world. However, mid-air gestures compromise object placement accuracy, whether due to limitations in tracking solutions or human dexterity itself.

Traditional object manipulation interfaces resort to 2D screens, mouse, and keyboard shortcuts to switch between views and apply constraints. On the other hand, multi-touch surfaces, which have already become part of everyday life, allow direct ways of interacting with virtual objects. When designing interactions for these kind of interfaces, and contrary to what happens in VR, it is necessary to develop appropriate mappings between the 2D input and the desired 3D transformations. For touch-based interfaces, approaches that separate degrees-of-freedom (DOF) have become the most common, easing this mapping.

Although a direct mapping can be applied for mid-air gestures in VR, it has already been shown that the constraint of transformations through DOF separation can significantly improve object placement accuracy. However, to perform several transformations on two or more axes, it is necessary to perform multiple consecutive operations. Thus, such approaches, besides compromising the naturalness of the interaction, have a negative impact on the execution time of placement tasks.

In this work we explore custom transformation axes, as an alternative to the traditional fixed frames. We assess if such approach can benefit from the increase of precision associated to DOF separation, but without impacting the time required to perform more complex tasks. For this, we propose a new manipulation technique, MAiOR (Mid-Air Objects on Rails). It supports both direct manipulation and transformation separation, as well as single DOF manipulation.

In the remainder of the document, we start by presenting the state-of-the-art in virtual object manipulation, showing how DOF separation became popular in mouse and touch based interfaces and which are the current proposals for mid-air manipulation. We then describe MAiOR, our technique that implements custom transformation axis in mid-air. We follow with a user evaluation, were we compare MAiOR against a direct 6-DOF approach and single DOF widgets. Finally, we conclude the paper and point out future research directions.

2 RELATED WORK

Traditional interfaces for 2D interaction usually resort to mouse and keyboard for indirect manipulation. In most approaches, users have to use keyboard shortcuts and widgets to select both a transformation and an axis to perform simple manipulations on an object. For instance, the Virtual Handles technique [Conner et al. 1992] uses a set of widgets associated with axes and transformations. This approach is still very popular in 3D editing applications such as Autodesk Maya or Unity 3D.

Multi-touch surfaces allow for more direct approaches. Users can interact directly with objects displayed on the surface. The RST approach, or Two-Point Rotation & Translation [Hancock et al. 2006], is a de facto standard for 2D manipulations on multi-touch devices. It is direct and based on the behavior of objects in the physical world, which makes it easy to learn and fast to execute. For 3D manipulations, there are techniques that extend RST to the third dimension, namely Sticky Fingers [Hancock et al. 2009] and Screen-Space Formulation [Reisman et al. 2009]. Direct manipulations are fast, but can lead to unwanted transformations and offer little accuracy. To address this challenge, DS3 [Martinet et al. 2010] focuses on DOF separation. Other techniques have also addressed this through virtual widgets or gestures on the surface. GimbalBox [Bollensdorff et al. 2012] and tBox [Cohé et al. 2011] use widgets to enable users to independently control translations, rotations and scales on specific axes. Schmidt et al. [Schmidt et al. 2008] introduced an approach for sketch-based interfaces where, after indicating the object to be transformed, the user can draw a stroke and the system responds by creating translation and rotating widgets based on the candidate axis closest to the stroke. Candidate axes include object and world axes. The Multi Touch Gestures [Au et al. 2012] technique rely only on hand gestures to apply constraints and manipulate objects in relation to candidate axes. Touch Sketch [Wu et al. 2015] has a constraints' menu to separate DOF. Studies have concluded that, in the various cases, DOF separation is better, although being less natural. This separation benefits tasks that have higher requirements of accuracy because there are fewer simultaneously modified DOF.

Mid-air manipulation techniques, which resort to spatial tracking of users' hands, are more appropriate to mimic interactions in the physical world. With 6-DOF position and orientation hand tracking, Simple Virtual Hand [Bowman et al. 2004] allows users to move and rotate objects simultaneously. The Handle-Bar [Song et al. 2012] uses only the position of both hands of a user to simultaneously move, rotate and scale objects, using the metaphor of a bar. Spindle+Wheel [Cho and Wartell 2015] uses the same metaphor, but with a 6-DOF hand tracking, it also supports rotations around the bar axis. It has been shown [Mendes et al. 2014] that direct manipulation techniques such as Handle-Bar [Song et al. 2012] and Simple Virtual Hand [Bowman et al. 2004] are the fastest and most natural.

However, these direct approaches may suffer from precision issues. With this challenge in mind, the PRISM [Frees et al. 2007] technique was proposed. It dynamically reduces users' virtual hand movements to decrease instability caused by real hand and increase accuracy. Similar to the multi-touch approaches described earlier, virtual widgets have also been explored in the air. 7-Handle [Nguyen et al. 2014] explores this, but the complex widgets' design made the learning curve very steep, performing worse than a direct 6-DOF approach. The impact of the explicit separation of DOF in mid-air manipulation tasks has also been evaluated [Mendes et al. 2016]. After a comparison of approaches including direct manipulation with 6-DOF, PRISM [Frees et al. 2007] and an implementation of virtual widgets for DOF separation, it was concluded that DOF separation is able to increase precision and prevent unwanted transformations, but it impairs execution time of more complex tasks. From this study, a set of four guidelines for mid-air manipulation techniques was proposed: (1) direct 6-DOF manipulation is suitable for fast and coarse transformations; (2) separating transformations helps prevent unexpected results; (3) 1-DOF transformations are useful for fine adjustments; (4) scaled movements effectively reduce positioning error in translations. Regarding single DOF manipulations, Veit et al. suggested that interactions that ease tasks' decomposition can lead to significant improvements in performance for orientation tasks [Veit et al. 2009].

In this work, we assess whether custom transformation axis with scaled movements allow users to achieve the same level of precision as single DOF manipulations, while minimizing the impact in the tasks' completion time. For this purpose, we developed and evaluated a novel mid-air manipulation technique following the above set of guidelines, with added support for user specified transformation axis.

3 MAiOR

To explore custom transformation axis in mid-air object manipulations, we developed MAiOR (Mid-Air Objects on Rails). It is a technique that follows the aforementioned guidelines for mid-air manipulation [Mendes et al. 2016]. It offers transformation separation and single-DOF manipulation on custom axes, as well as 6-DOF direct manipulation and scaled movements for fast and accurate transformations, respectively. Figure 1 show how to activate available transformations and constraints in MAiOR.

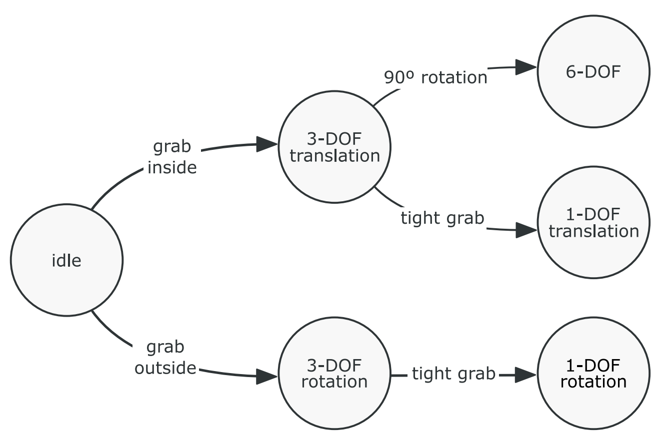
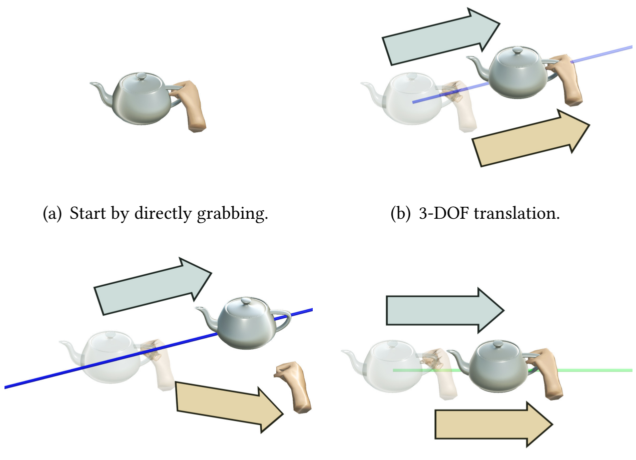


Figure 1: MAiOR's interaction diagram.



(a) Start by directly grabbing. (b) 3-DOF translation. (c) 1-DOF translation (d) Snap to candidate axis.

Figure 2: MAiOR's translation.

3.1 Translation

Translations can be performed in MAiOR by directly grabbing the desired object. Initially, the object will be restricted to 3-DOF translations. A transparent blue axis is drawn from the initial position of the object and passing through its current position (Figure 2). If the axis has a 10 degree or less deviation from any candidate axis, either from world or object reference frames, the closest candidate axis is shown instead, similarly to some approaches for interactive surfaces [Au et al. 2012; Schmidt et al. 2008], and its color is changed to green (world) or yellow (object).

The displayed axis can then be used to restrict translations to 1-DOF exclusively in that same axis. It can be locked at any time after starting a translation, by tightly closing the hand. When the axis is locked, it becomes opaque and infinitely extended for both sides. If a candidate axis has been selected, the object is repositioned in such a way that the translation made since it was initially grabbed is coincident with the axis. Thereafter, variations in user's hand position will be projected on the selected axis, and object translations will follow the metaphor of an object on a rail.

Since it has been shown that scaled movements improve precision in object placement when small adjustments are required [Frees et al. 2007; Mendes et al. 2016], we implemented scaled translations using a fixed scale factor. This can be used both in 3-DOF and 1-DOF translations, by closing the non-dominant hand. Given the mean position error attained in previous evaluations of mid-air manipulations [Mendes et al. 2016], we chose a scaling ratio of $\frac{1}{4}$, aiming for millimetric accuracy. Naturally, this method originates the accumulation of an offset between the object and the hand.

3.2 Direct Manipulation

Since direct manipulation is the most efficient approach for coarse operations that require multiple transformations, MAiOR also offers direct manipulations. While in translation mode, hand rotations are discarded until they achieve a 90 angle or greater in any direction

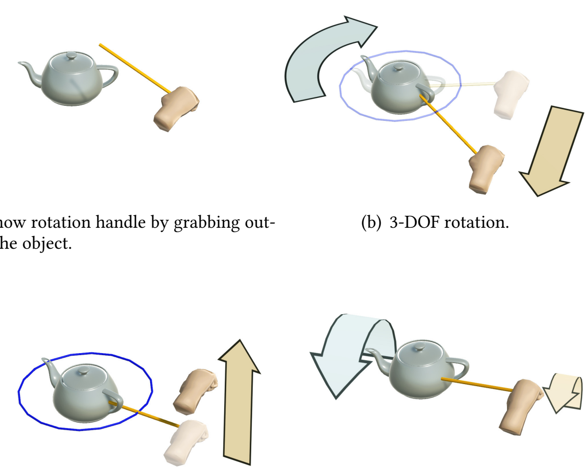
(a) Object grabbed. (b) Hand rotated 45°. (c) Hand rotated 90°.

Figure 3: Unlocking MAiOR's 6-DOF manipulation.

(Figure 3), which triggers 6-DOF transformations. This gesture follows the metaphor of unlocking a door, because the object will no longer be locked to 3-DOF translations. Whenever this mode is activated, the object is immediately rotated by the same amount that the hand has rotated since the grab gesture, in order to ensure that the orientation of the object is consistent with the orientation of the hand.

3.3 Rotation

Following the suggestions from Veit et al. [Veit et al. 2009], we also allow the decomposition of orientation tasks into single DOF manipulations. MAiOR's rotations are based on the same principles of DOF separation applied in translation, as depicted in Figure 4. Users start by rotating the object in 3-DOF and then can select an axis to rotate in 1-DOF. For this, users first have to close a hand outside the objects to enable a virtual bar (Figure 4). This bar will act as a lever for rotating objects. After attaching the bar to the desired object, the object can be rotated in 3-DOF. This approach is based on techniques such as the Handle-Bar [Song et al. 2012]



(a) Show rotation handle by grabbing outside the object. (b) 3-DOF rotation. (c) Motion parallel to locked rotation axis is discarded. (d) Wrist rotation.

Figure 4: MAiOR's rotation.

and the Spindle+Wheel [Cho and Wartell 2015], which have been shown to have good results in previous studies. However, in our approach we use the center of the object instead of a second hand, since the object remains in the same position, and this point is used as center of rotation. The rotation angle to be applied will be calculated according to the variation in the bar's orientation. This is, in turn, determined by the position of the user's hand, so that the intersection point between the bar with the object remains unchanged. In addition, we also implemented rotations around the axis defined by the bar, using wrist rotations.

After entering the 3-DOF rotation mode, a transparent blue circumference is shown around the object illustrating the current object rotation, taking into account hand position's variation (rotations around the wrist are not considered for this purpose). Similar to translation, if the calculated axis of rotation has a deviation of 10 degrees or less from a candidate axis, the circumference will be shown around it, either in green or yellow. Users can then lock the current rotation axis, by tightly closing the hand. Thereafter, to calculate rotation angle, users' hands movements will be projected in the plane defined by the circumference shown.

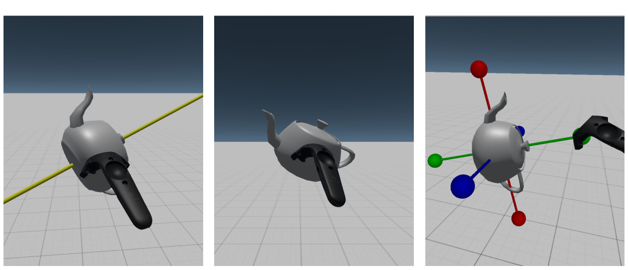
Since scaling rotations has a negative effect on users [Frees et al. 2007; Mendes et al. 2016], we did not implemented an explicit precision mode in rotations. However, it is possible to scale rotations implicitly, due to the concepts behind circular motion. The further away the user's hand is from the object, the more distance will have to be covered for the object to rotate. This can be used to improve precision when fine-tuning object's orientation.

4 USER EVALUATION

To assess whether the custom transformation axis implemented in MAiOR appeal to users and help achieving an accurate and fast object placement, we conducted a user evaluation. We compared MAiOR against two baseline approaches, with a set of object placement tasks with different requirements.

4.1 Baseline Mid-air Manipulation Techniques

As baselines, we chose two techniques that achieved the better results in previous research [Mendes et al. 2016]: a direct 6-DOF approach and an indirect approach based on 3D widgets (Figure 5). The first was identified as the fastest, while the second was the most accurate.



(a) MAiOR. (b) 6-DOF. (c) Widgets.

Figure 5: Evaluated techniques.

4.1.1 6-DOF. The direct 6-DOF approach is often used as a baseline for evaluating other techniques [Frees et al. 2007; Mendes et al. 2016; Nguyen et al. 2014], as it simulates as closely as possible interactions with physical objects. Manipulations start by directly grabbing the desired object. It will then closely follow the user's hand, moving and rotating accordingly. As both translation and rotation are applied simultaneously, there is no transformation separation. The point grabbed in the object will be the center of all transformations, until it is released.

4.1.2 Widgets. This technique is based on the 3D Virtual Handles common in mouse-based interfaces, as initially proposed by Conner et al. [Conner et al. 1992], which was recently adapted to mid-air interactions [Mendes et al. 2016]. It strictly follows explicit DOF separation, allowing only one transformation at a time according to a single axis form the object frame. The widget is composed by three cylinders representing object axes with spherical handles in each end, following a RGB coding for XYZ axes respectively. Translations can be done by grabbing an handle and moving the hand along the corresponding axis. Rotations are performed by also grabbing and handle, but rotating it about the desired axis. The decision to either perform a translation or rotation is made based on the first 10 cm from the handfis path after grabbing the handle. The transformation and axis resulting from that decision will remain locked until the handle is released.

4.2 Prototype

In order to compare MAiOR against the two baselines, we developed a prototype where we implemented the three manipulation techniques.

4.2.1 Setup. We used a setup based on the HTC Vive, which tracks users' head and hands in 6-DOF. While head tracking is made by the headset itself, hand tracking is made through handheld controllers. Each controller has a trigger with 10 levels of pressure and a circular track-pad. We use the trigger to detect if the hand is opened (no pressure), closed (any pressure level from level 1 to 9) or tightly closed (full pressure at level 10). The last pressure level is perceived by a slight click on the trigger, and is used to lock translations and rotations to 1-DOF in MAiOR. This approach to trigger restricted transformations is adaptable to other kinds of force-sensitive devices, and avoids the use of multiple buttons.

4.2.2 Virtual Environment. We developed the prototype using the Unity3D engine. The environments consists of a wide and empty plane area, with shadows but no gravity nor collisions. Users' hands are represented through virtual replicas of the controllers. Objects become transparent whenever users intersect them, and became opaque as soon as they are grabbed in order for improved visual feedback.

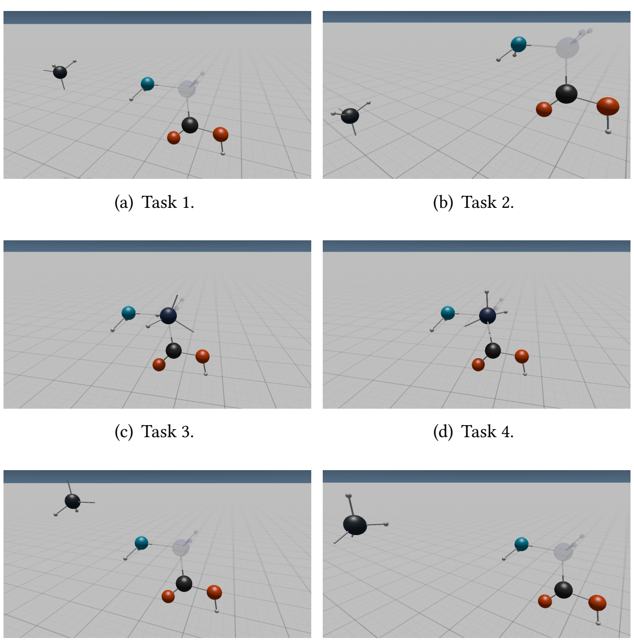
4.3 Methodology

All sessions followed the same structure, and could last a maximum of 70 minutes. After an introduction to the evaluation, participants experimented all techniques. Techniques' order followed a Latin square design, to avoid biased results. For each technique, we started by playing a demo video explaining it, and gave a maximum of 5 minutes for participants to freely explore it and get acquainted with the environment. Then, we asked participants to execute a set of tasks and fulfill a questionnaire.

4.4 Tasks

Participants were asked to completed a set of six docking tasks for each technique. This set was based on previous user evaluation for mid-air manipulations [Mendes et al. 2016]. The objective of all tasks was to place a carbon component on a model of a protein compound (Figure 6). The model was designed so that there was only one correct way to fit it. When the carbon component was placed on the docking model within the error boundaries, its color turned green and the task goal was achieved.To foster an highly accurate object placement, we set error boundaries to less than 1 millimeter for position and 1 degree for orientation. To avoid long user sessions, each task had a maximum time of 2m30s. If the time limit was reached, we considered the attained position and orientation as final and registered it as an unsuccessful attempt. Although some tasks required only translation or rotation transformations, none of those were restricted on any task.

The first two tasks required translations only. For the first task, the carbon needed to be moved only along the X axis. In the second, its position was initially incorrect in all three axes. Third and fourth tasks required rotations only. In the third task participants needed to rotate the carbon about the Z axis, while the fourth they needed to perform rotations about X, Y and Z axes. The last two tasks were the most complex, requiring both translations and rotations. The fifth task required the object to be rotated about the Z axis and moved along X and Y axes. On the final task, participants had to apply a full 6-DOF transformation.



(a) Task 1. (b) Task 2.

(c) Task 3. (d) Task 4.

(e) Task 5. (f) Task 6.

Figure 6: User evaluation tasks.

4.5 Apparatus and Participants

Evaluation sessions were performed in our lab, which has restricted access, thus providing a calm and controlled environment. We had a total of 24 participants, 16 males and 8 female, between the ages of 17 and 33 years old. Most of them hold at least a BSc degree (80%). 83% reported having no previous experience in VR, and 70% had never used any kind of gesture recognition systems, such as Xbox Kinect, Nintendo Wiimote or Playstation Move. Only 16% of the participants use 3D modeling systems at least once a month.

5 RESULTS AND DISCUSSION

During user sessions, we gathered objective data through a logging mechanism and subjective data from the questionnaires. To analyze such data, we used Shapiro-Wilk test to assess data normality. Then, we ran the repeated measures ANOVA test with a Greenhouse-Geisser correction to find significant differences in normal distributed data. Otherwise, we ran Friedman non-parametric test with Wilcoxon Signed-Ranks post-hoc test. In both cases, post-hoc tests used Bonferroni correction (presented sig. values are corrected).

5.1 Task Performance

We measured success rate, completion time and object placement error (position error in millimeters and orientation error in degrees) for each task.

5.1.1 Success Rate. Success rate was defined by the ratio of participants that were able to place the virtual object below the position and orientation tolerated error and within the time limit. Values are shown in Table 1, and statistically significant differences were found (Task 1: $\chi^2(2)$=27.9, p<.0005; Task 3: $\chi^2(2)$=13.875, p=.001; Task 4: $\chi^2(2)$=8, p=.018; Task 5: $\chi^2(2)$=8.4, p=.015). On task one, 6-DOF had a lower success than MAiOR (Z=-3.638, p<.0005) and Widgets (Z=-4.243, p<.0005). On task three, Widgets outperformed both 6-DOF (Z=-3.051, p=.006) and MAiOR (Z=-2.887, p=.012). On task four, Widgets was better than 6-DOF (Z=-2.673, p=.024) and, on task five, better than MAiOR (Z=-2.714, p=.021).

6-DOF attained consistent success rate values. Since it does not possess any kind of transformation's restriction, all tasks are of identical execution. The low success rate relates to the difficulty to accurately place an object in mid-air. Widgets had the highest overall success rate, successfully preventing undesired transformations. However, and as shown in previous research [Mendes et al. 2016], the more complex the task, the higher the time needed to complete it, and more difficult it is to do so within the time limit. MAiOR's success rate also decreased when task complexity

Table 1: Success rate per task for each technique. $*$ and † indicate statistically significant differences.

	T1	T2	T3	T4	T5	T6
MAiOR	79% *	58%	33% *	42%	17% *	13%
6-DOF	17% * †	33%	29% †	25% *	29%	33%
Widgets	92% †	54%	75% * †	67% *	54% *	17%

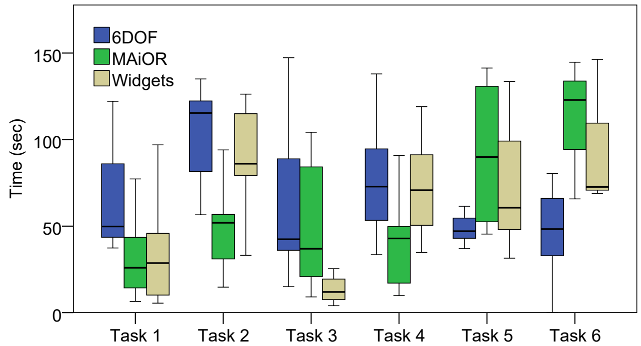


Figure 8: Tasks' completion time, in seconds. The chart presents the median, 1st and 3rd interquartile ranges (boxes) and 95% confidence interval (whiskers).

increased. We believe that this is due to both DOF separation's additional steps, and a not so simple rotation metaphor that made all tasks involving rotations more difficult. Indeed, the success rate increased from task 3 to task 4, suggesting that experience may have affected participants' performance in rotation tasks.

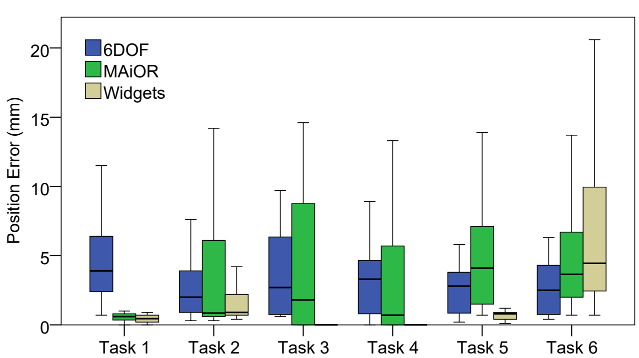
5.1.2 Completion Time. To analyze completion time, we only considered times attained by participants who achieved tasks' goal within the time limit, which are depicted in Figure 8. Although we did not find any statistical significant differences, we verified in tasks 2 and 4 a tendency for MAiOR (Task 2: avg=48s; Task 4: avg=39s) to be faster than Widgets (Task 2: avg=92s; Task 4: avg=72s). Taking into account individual times, 8 out of 9 (task 2) and 6 out of 6 (task 4) participants that completed those tasks with both techniques had lower completion times with MAiOR than with Widgets. However, this needs to be confirmed with further testing.

5.1.3 Placement Error. Regarding placement error, we analyzed both position and rotation error (Figure 7). For participants that did not achieve tasks' goal, we considered the current placement when the time expired. Significant differences existed for position error (Task 1: $\chi^2(2)$=29.84, p<.0005; Task 3: $\chi^2(2)$=20.118, p<.0005; Task 4: $\chi^2(2)$=19.5, p<.0005; Task 5: F(1.363,19.085)=10.075, p=.003; Task 6: $\chi^2(2)$=9.264, p=.01). Post-hoc tests showed that for the first task, 6-DOF (avg=4.65mm) was less accurate than both Widgets (avg=0.43mm, Z=-4.108, p<.0005) and MAiOR (avg=0.55mm, Z=-3.921, p<.0005). In task 3, 6-DOF (avg=3.44mm) performed worst than Widgets (avg=0.43mm, Z=-4.108, p<.0005). In both fourth and fifth tasks, Widgets(Task 4: avg=0.00mm; Task 5: avg=0.69mm) outperformed 6-DOF (Task 4: avg=3.38mm, Z=-3.725, p<.0005; Task 5: avg=2.609mm, p=.001) and MAiOR (Task 4: avg=2.94mm, Z=-2.803, p=.015; Task 5: avg=5.00mm, p=.006). However on the sixth task, 6-DOF (avg=2.61mm) proven to be more precise than Widgets (avg=6.52mm, Z=-3.015, p=.009) and MAiOR (avg=4.41mm, Z=-2.521, p=.036), mostly due to time constraints. In rotation only tasks (task 3 and 4), the existence of a positional error with MAiOR implies that participants performed translations, even though MAiOR had transformation separation. This is further discussed in Section 5.3.

Technique used also influenced orientation error (Task 1: $\chi^2(2)$= 32, p<.0005; Task 2: $\chi^2(2)$=22, p<.0005; Task 3: $\chi^2(2)$=13.857, p=.001; Task 4: $\chi^2(2)$=8.4, p=.015; Task 5: $\chi^2(2)$=10, p=.007). In translation only tasks, 6-DOF (Task 1: avg=1.77°; Task 2:avg=1.22°) was the only that caused this kind of error, significantly worse than Widgets (Task 1: avg=0.00°, Z=-3.92, p<.0005; Task 2: avg=0.00°, Z=-3.408, p=.003) and MAiOR (Task 1: avg=0.00°, Z=-3.516, p<.0005; Task 2: avg=0.00°, Z=-3.517, p<.0005), which shows the benefits of translation and rotation separation in preventing unwanted transformations. In the remaining three tasks Widgets (Task 3: avg=0.54°; Task 4: avg=0.77°; Task 5: avg=0.48°) achieved better results than 6-DOF (Task 3: avg=2.42°, Z=-3.743, p<.0005; Task 4: avg=1.39°, Z=-2.722, p=.018; Task 5: avg=1.47°, Z=-3.637, p<.0005), as expected. In these tasks, Widgets were also more precise than MAiOR (Task 3: avg=1.72°, Z=-2.92, p=.012; Task 4: avg=1.27°, Z=-2.442, p=.045; Task 5: avg=2.20°, Z=-2.442, p=.045), possibly due to the difficulty reported by participants to perform rotations with this technique.

5.2 User Preferences

Through questionnaires, we asked participants about their experience with each technique. This included general ease of use, fun factor, ease of manipulating object position and orientation, recall



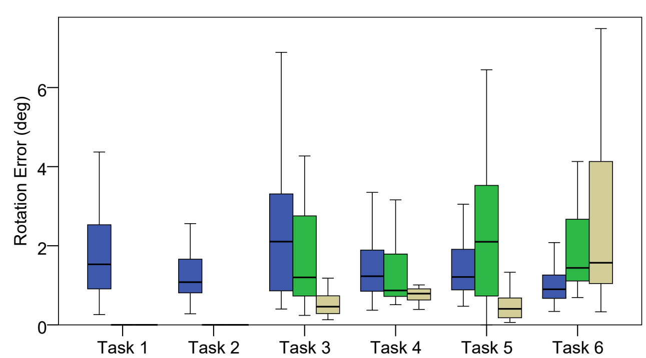


Figure 7: Position error, in millimeters, and rotation error, in degrees. Charts present the median, 1st and 3rd interquartile ranges (boxes) and 95% confidence interval (whiskers).

Table 2: Answers to questionnaires, regarding each criteria (median, interquartile range). ∗ and † indicate statistically significant differences.

	MAiOR	6-DOF	Widgets
Overall easiness	2 (1) *	2 (1)	3 (2) *
Translation easiness	3.5 (2)	4 (2)	4 (2)
Translation recall	3.5 (1) *	5 (1)	5 (1) *
Rotation easiness	2.5 (1) * †	3.5 (2) *	3 (2) †
Rotation recall	3 (1) * †	5 (1) *	4,5 (1) †
Fatigue	3 (2)	3 (2)	3 (1)
Fun	3 (1) *	4 (1)	4 (2) *

regarding translation and rotation transformations and overall fatigue. In all questions it was used a Likert Scale from 1 to 5, being 5 the favorable value. Answer's are reported in Table 2.

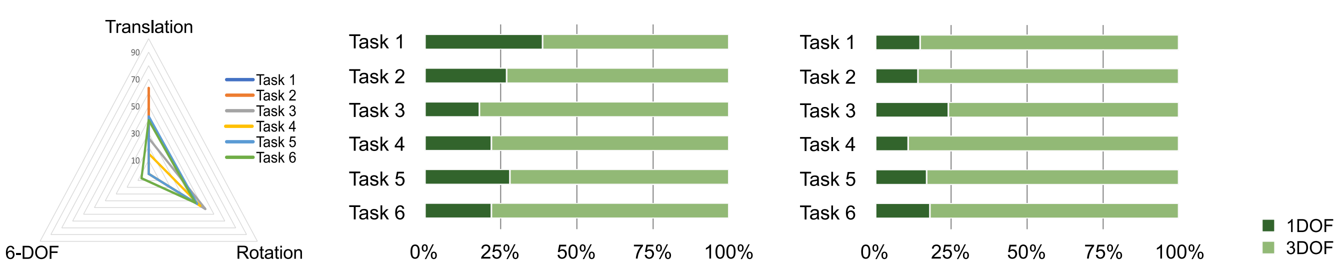
We found statically significant differences in ease of use ($\chi^2(2)$= 6.685, p=.035), translation recall ($\chi^2(2)$=11.541, p=.003), rotation easiness ($\chi^2(2)$=14.427, p=.001) and recall ($\chi^2(2)$=20.738, p<.0005), and fun factor ($\chi^2(2)$=7.69, p=.021). Participants agreed that Widgets were generally easy and more fun to use than MAiOR (easiness: Z=-2.623, p=.027; fun: Z=-2.599, p=.009). While for translations there was no significant difference in easiness between the three techniques, participants agreed that remembering how to translate was easier with Widgets than with MAiOR (Z=-3.246, p=.003). Participants also indicated that was harder to perform and to remember rotations with MAiOR than both 6-DOF (rotation easiness: Z=-3.407, p=.003; rotation recall: Z=-3.578, p<.0005) and Widgets (rotation easiness Z=-2.99, p=.009; rotation recall: Z=-3.691, p<.0005). Differences in recall can be justified by the complexity inherent to the gesture-based grammar, as the available actions are not visible to users and an additional effort must be done to remember how to perform such actions. Contrary to MAiOR, in Widgets and 6-DOF users only needed to remember to grab and drag the object. Moreover, the rotation metaphor adopted in MAiOR was identified as being difficult to get acquainted. According to participants' comments, they needed more time to fully take advantage of MAiOR.

Figure 10: Example of a pose performed by participants.

5.3 Observations

We observed that, when using the direct 6-DOF approach, some participants used specific poses to help reducing hand tremor and to be easier to achieve tasks' goal. They held tight the dominant hand's arm while using the other arm as support, as shown in Figure 10. Other participants mostly resorted to chance. First, they roughly placed the object with position and orientation close to the target, then tried successive grabs and releases hoping one would be within the acceptable threshold.

As previously stated, in tasks 2 and 4 almost all participants had better performance with MAiOR than Widgets. Looking at participants' profiles, we found that in task 2 they had several distinct backgrounds. This suggests that MAiOR's translation approach might be adequate for novice users. On task 4, however, they had backgrounds related to 3D modelling, such as design and architecture. This can mean that MAiOR's rotations require more experience with 3D manipulation concepts, such as rotation axes. More, we noticed that participants had difficulties understanding MAiOR's circumference widget's motion. This might be due to the fact that wrist rotation was not accounted for this. One of the most experienced participants also reported that using a custom pivot for rotations instead of the object's center could be beneficial.



(a) Transformations. (b) Translation. (c) Rotation.

Figure 9: Time distribution, in percentage, between transformations (a), and between 3-DOF and 1-DOF translation (b) and rotation (c), in MAiOR.

Besides task completion time, we logged where participants spent time when manipulating the object. This is illustrated in Figure 9. It is possible to see that in rotation only tasks (task 3 and 4), participants' first instinct was to directly grab the object. This originated some position error that they tried to correct. Also, users favored transformation separation, which is mostly noticed in tasks 5 and 6, where time spent performing independent translations and rotations is evenly distributed.

We also registered, for both independent translations and rotations, time spent in 3-DOF and 1-DOF manipulations in MAiOR, also depicted in Figure 9. We verified that participants preferred to perform transformations in 3-DOF, as their were faster than 1-DOF. These results contradict previous research suggestions [Mendes et al. 2016; Veit et al. 2009]. With MAiOR's scaled movements in translation, participants found the positioning precision attained with 3-DOF adequate. On the other hand, they found it difficult to achieve the required orientation, yet they rarely restricted rotation to a single DOF. Additionally, the tight grab gesture used to constraint transformations originated unintentional activations of 1-DOF, as sometimes participants applied the extra strength without noticing.

6 CONCLUSIONS AND FUTURE WORK

Direct approaches for mid-air object manipulation, albeit natural and fast, lack precision. Previous research showed that DOF separation, common in mouse and touch based interfaces, benefit precision in mid-air manipulations. However, they also increase completion time in complex tasks, due to the additional steps required by performing a single transformation in one axis at a time. In this work we explored personalized transformation axes. Our objective was to assess if they can contribute for faster yet accurate manipulations. For this, we developed a novel technique, MAiOR.

We compared MAiOR against two baseline approaches, one is based on direct 6-DOF manipulations and other on 3D widgets. While MAiOR did not compromised tasks' completion time, Widgets had the best performance overall. Even though, MAiOR showed promising results on isolated transformation tasks. When analyzing individual transformation times in MAiOR, we found that while participants took advantage of transformation separation to prevent unwanted results, they did not sought single DOF manipulations.

The challenge of achieving high levels of precision in mid-air still offers avenues for future research. Our findings regarding 3-DOF vs 1-DOF contradicts suggestions from previous research [Mendes et al. 2016; Veit et al. 2009]. As such, further testing should be carried out in order to clarify this. Additionally, MAiOR's rotation approach was generally reported as hard to grasp for participants not acquainted with 3D software. Investigating into more easy-to-learn metaphors can lead to improved manipulations. Finally, we also intend to explore novel interaction paradigms, resorting to approaches based not only on software, but also on specific hardware, such as haptics. This can be used, for instance, to add additional modalities for feedback to aid in precise object manipulations.

ACKNOWLEDGEMENTS

This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with ref. UID/CEC/50021/2013, through project IT-MEDEX (PTDC/EEISII/6038/2014), and doctoral grant SFRH/BD/91372/2012.

REFERENCES

Oscar Kin-Chung Au, Chiew-Lan Tai, and Hongbo Fu. 2012. Multitouch Gestures for Constrained Transformation of 3D Objects. *Computer Graphics Forum* 31, 2pt3 (2012), 651–660. https://doi.org/10.1111/j.1467-8659.2012.03044.x

Björn Bollensdorff, Uwe Hahne, and Marc Alexa. 2012. The Effect of Perspective Projection in Multi-touch 3D Interaction. In *Proceedings of Graphics Interface 2012 (GI '12)*. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 165–172. http://dl.acm.org/citation.cfm?id=2305276.2305304

Doug Bowman, Ernst Kruijff, Joseph J LaViola Jr, and Ivan P Poupyrev. 2004. *3D User Interfaces: Theory and Practice, CourseSmart eTextbook*. Addison-Wesley.

Isaac Cho and Zachary Wartell. 2015. Evaluation of a bimanual simultaneous 7DOF interaction technique in virtual environments. In *3D User Interfaces (3DUI), 2015 IEEE Symposium on*. IEEE, 133–136.

Aurélie Cohé, Fabrice Dècle, and Martin Hachet. 2011. tBox: A 3D Transformation Widget Designed for Touch-screens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 3005–3008. https://doi.org/10.1145/1978942.1979387

Brookshire D Conner, Scott S Snibbe, Kenneth P Herndon, Daniel C Robbins, Robert C Zeleznik, and Andries Van Dam. 1992. Three-dimensional widgets. In *Proceedings of the 1992 symposium on Interactive 3D graphics*. ACM, 183–188.

Scott Frees, G Drew Kessler, and Edwin Kay. 2007. PRISM interaction for enhancing control in immersive virtual environments. *ACM Transactions on Computer-Human Interaction* 14, 1 (2007), 2.

Mark Hancock, Thomas ten Cate, and Sheelagh Carpendale. 2009. Sticky Tools: Full 6DOF Force-based Interaction for Multi-touch Tables. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS '09)*. ACM, New York, NY, USA, 133–140. https://doi.org/10.1145/1731903.1731930

Mark S Hancock, Frédéric D Vernier, Daniel Wigdor, Sheelagh Carpendale, and Chia Shen. 2006. Rotation and translation mechanisms for tabletop interaction. In *Horizontal Interactive Human-Computer Systems, 2006. TableTop 2006. First IEEE International Workshop on*. IEEE, 8–pp.

Anthony Martinet, Géry Casiez, and Laurent Grisoni. 2010. The Effect of DOF Separation in 3D Manipulation Tasks with Multi-touch Displays. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology (VRST '10)*. ACM, New York, NY, USA, 111–118. https://doi.org/10.1145/1889863.1889888

D. Mendes, F. Fonseca, B. Arajo, A. Ferreira, and J. Jorge. 2014. Mid-air interactions above stereoscopic interactive tables. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*. 3–10. https://doi.org/10.1109/3DUI.2014.6798833

Daniel Mendes, Filipe Relvas, Alfredo Ferreira, and Joaquim Jorge. 2016. The Benefits of DOF Separation in Mid-air 3D Object Manipulation. In *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology (VRST '16)*. ACM, New York, NY, USA, 261–268. https://doi.org/10.1145/2993369.2993396

Thi Thuong Huyen Nguyen, Thierry Duval, and Charles Pontonnier. 2014. A New Direct Manipulation Technique for Immersive 3D Virtual Environments. In *ICAT-EGVE 2014 : the 24th International Conference on Artificial Reality and Telexistence and the 19th Eurographics Symposium on Virtual Environments*. Bremen, Germany, 8. https://hal-imt.archives-ouvertes.fr/hal-01083215

Jason L Reisman, Philip L Davidson, and Jefferson Y Han. 2009. A screen-space formulation for 2D and 3D direct manipulation. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*. ACM, 69–78.

Ryan Schmidt, Karan Singh, and Ravin Balakrishnan. 2008. Sketching and composing widgets for 3d manipulation. In *Computer Graphics Forum*, Vol. 27. Wiley Online Library, 301–310.

Peng Song, Wooi Boon Goh, William Hutama, Chi-Wing Fu, and Xiaopei Liu. 2012. A Handle Bar Metaphor for Virtual Object Manipulation with Mid-air Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 1297–1306. https://doi.org/10.1145/2207676.2208585

Manuel Veit, Antonio Capobianco, and Dominique Bechmann. 2009. Influence of Degrees of Freedom's Manipulation on Performances During Orientation Tasks in Virtual Reality Environments. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology (VRST '09)*. ACM, New York, NY, USA, 51–58. https://doi.org/10.1145/1643928.1643942

Siju Wu, Amine Chellali, Samir Otmane, and Guillaume Moreau. 2015. TouchSketch: A Touch-based Interface for 3D Object Manipulation and Editing. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology (VRST '15)*. ACM, New York, NY, USA, 59–68. https://doi.org/10.1145/2821592.2821606