[Poster] Touch Gestures for Improved 3D Object Manipulation in Mobile Augmented Reality

Philipp Tiefenbacher* Andreas Pflaum† Gerhard Rigoll‡

Institute for Human-Machine Communication
Technische Universität München

Abstract

This work presents three techniques for 3D manipulation on mobile touch devices, taking the specifics of mobile AR scenes into account. We compare the common direct manipulation technique with two indirect techniques, which utilize only the thumbs to perform the transformations.

The evaluation of the manipulation variants is conducted in a mixed reality (MR) environment which takes advantage of the controlled conditions of a full virtual reality (VR) system. A study with 18 participants shows that the two-thumb method tops the other techniques. It performs better with respect to the total manipulation time and total number of gestures.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Interaction styles;

1 Introduction

Experts commonly create AR scenes which thereafter cannot be modified by the user. An intuitive way for the manipulation of existing content allows easy interaction with AR content by average users. Therefore, we present three methods for manipulating 3D items to fill and edit an augmented world.

The envisaged interaction methods start with the direct manipulation method, which is the primary input modality for smart phones. Since mobile AR applications require at least one hand to hold the device, only one hand is available to perform the direct manipulation gestures. That is why our second variant allows indirect manipulation of 3D items. This way both hands can grasp the device. If the device is grasped with both hands, the thumbs are always freely available. Consequently, the third method uses bi-manual thumb gestures. We include three rigid body transformations of objects in 3D space: translation (3 DoF), rotation (3 DoF) and uniform scaling (1 DoF).

The main contributions of this work are the evaluation of three different input variants for 3D transformations on a tracked mobile device. Furthermore, we show that our new two-thumb gestures are the favorite method for 3D touch interaction in AR.

Based on the body-centric design space of [5], our scenario can be described as a combination of mid-air gestures fixed in the world and on-device touch gestures relative to the body. The mid-air and on-device gestures control the device's position and the 3D manipulation, respectively.

This is different from approaches with solid touch walls or tabletops. The camera position on stationary devices is fixed and can be manipulated only virtually [2]. Works of indirect interaction concepts on large stationary devices like tabletops differ from our mobile scenario as it is required to hold an additional device. These

Figure 1: One thumb rotation of the virtual trackball approach (left) and the 2D rotation (right).

characteristics motivate our two indirect methods. Ideas to interact with AR scenes already exist, but the manipulations are only performed on 2D objects and are less general [1] or the authors do not focus on different touch-based interaction methods [4].

2 Touch Gestures

Direct Touch Gestures The selected 3D item follows the finger for a translation in the ground plane. The manipulation stops as soon as the finger loses contact. A detection of two fingers in parallel on the touch screen executes a translation in the third dimension.

The fingers are on opposite sides of the object for a rotation in 2D. Then the fingers rotate around the object's center with the object mirroring the rotation in the same direction. We use a virtual trackball for the missing two directions of rotation. The user touches the object with a single finger and slides to the desired direction of rotation. The object rotates in the same way as a ball would rotate in response to this finger movement. The uniform scaling is based on the two finger-based pinch-to-zoom gesture.

Indirect Touch Gestures We divide the indirect gestures into one-thumb (uni-manual) and two-thumb (bi-manual) gestures. Both indirect gestures implement touch gestures which can be realized with just the two thumbs at the sides of the mobile touch device. The direct gestures consist of one-handed gestures only. This is equivalent to the indirect one-thumb gestures, which rely on one-handed gestures. Our third method, the two-thumb gestures, advances the single-handed gestures to two-handed gestures.

One-thumb gestures: Here, a translation in the ground plane is equivalent to the direct method. The translation in the third dimension cannot be accomplished by the two-finger gesture of the direct method since only one thumb is available. Hence, an additional menu item is added. The one-thumb method uses the trackball-based rotation gesture and a new 2D rotation gesture for the missing direction (Figure 1). When the thumb moves across the virtual trackball, the normal trackball rotation is realized. By the time the thumb moves outside of the displayed circle, the rotation mode changes to 2D rotation. The gesture for scaling a 3D item is based on the movement of the thumb towards and away from the object, which corresponds to scaling down and up, respectively.

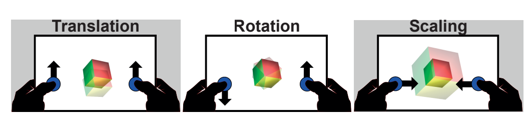

Figure 2: Indirect two-thumb gestures for the 3D manipulation.

Two-thumb gestures: The main difference to the gestures described previously is that every gesture should require both thumbs.

*e-mail:Philipp.Tiefenbacher@tum.de
†e-mail:Andreas.Pflaum@mytum.de
‡e-mail:Rigoll@tum.de

IEEE International Symposium on Mixed and Augmented Reality 2014
Science and Technology Proceedings
10 - 12 September 2014, Munich, Germany
978-1-4799-6184-9/13/$31.00 ©2014 IEEE

For translation, the thumbs move into the same direction while touching the screen as illustrated in Figure 2.

Moving the thumbs horizontally but in opposite directions summarizes the gesture for scaling.
In the same fashion, opposite vertical thumb movements enable the rotation gesture as illustrated in Figure 2. The gesture is equivalent to the direct 2D rotation.
The proposed two-thumb gestures allow transformations on one axis only. As a result, the user picks the desired axis (X, Y or Z) in the menu. This way the two-thumb gestures separate the transformation the most. The menu of the other methods includes items for the gestures.

3 User Study

The proposed gestures are implemented on a 11.6" Windows 8 mobile PC. The user study is performed in a CAVE. Lee et al. [3] show that the completion times of the same task in MR and AR are not significantly different. This finding validates our AR study in a CAVE.

The position of the user's head determines the rendered scene in the CAVE, whereas the mobile device's position describes the scene on the display of the mobile device in parallel. The scene on the tablet PC includes the virtual industrial apparatus of the CAVE, but also AR elements which are rendered solely on the mobile device.

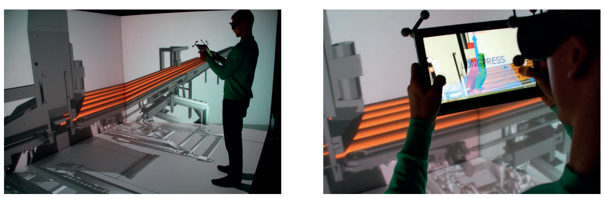

Figure 3: A person standing inside the CAVE. The head and the mobile device are tracked. The mobile device shows a machine part (green). The CAVE does not show this 3D illustration.

Scenario Figure 3 presents the main task of the study. The subject steps into the role of an industrial plant fitter, who inserts three 3D machine parts (AR objects) at certain positions. We encourage the subjects to move inside the CAVE. For that purpose, the final positions of the three machine parts differ spatially. The individual AR objects are selected through a menu and are loaded into the 3D scene dynamically. Based on the proposed touch gestures, the user has to position, rotate and scale the AR object until it matches the final pose. In addition to the AR parts which are added to the scene, the colored 3D shadows highlight the final poses of this AR items.

The AR object is fitted successfully as soon as the metrics of translation, scaling and rotation meet defined criteria which are estimated empirically. As the mobile device is tracked during the whole process, the subject is able to move around the AR objects. In total, three different AR objects need to be matched. The order and initial pose of the inserted 3D machine parts are constant throughout the study.

The evaluation is based on 18 subjects, four women and 14 men, which completed the study. We applied a within-subject design and altered the order of the input methods based on the Latin square method to reduce fatigue and learning effects. Before starting the task with a new input method, we demonstrated the method and the subjects were also able to test it.

4 Results

Manipulation Time We investigated the time the participants needed to perform the manipulations only. This manipulation time (MT) does not take the time into account where the participants stand or move inside the CAVE in order to get the best view of the AR object. Table 1 shows the manipulation times of all three methods. Based on the Shapiro-Wilk normality test, the MT is distributed normally at significance level of .05. An ANOVA indicated significant differences between the methods, $F(2,34) = 16.83$, $p = .00 < .05$. The post-hoc Tukey test showed that the MT of the indirect bi-manual method is significantly shorter than for the other two methods at $p < .05$. The direct method was hardly different from the one-thumb method.

Method	$\overline{MT}$ (s)	SD MT (s)	$\overline{\#_{gest}}$	SD $\#_{gest}$
Direct	123.64	34.61	59.44	17.68
Ind. Uni-man.	145.27	56.33	55.56	24.08
Ind. Bi-man.	87.32	24.60	41.39	9.36

Table 1: Comparison of the mean and standard deviation (SD) of manipulation times (MT) and the number of touch gestures ($\#_{gest}$).

Number of Touch Gestures Besides the manipulation times of the tasks, we recorded the number of touch gestures. Only touch gestures for the transformation of an AR object are counted. The mean number of gestures was highest for the direct method, however, the one-thumb method again had the biggest variance which is depicted in Table 1. Using the two-thumb method, participants needed only 41.39 gestures for the whole task on average. The data is not distributed normally at a .05 significance level. A non-parametric Friedman test exhibited a significant difference between the methods, $\mathscr{X}^2(2,53) = 13.00$, $p = .00 < .05$. The post-hoc Tukey test revealed that the number of manipulations with both indirect methods was significantly less than with the direct method at $p < .05$. The indirect methods did not differ.

5 Conclusion

The two-thumb method combines the advantages of easy grasping and multi-touch input. It was the fastest method. It needed the least number of gestures and also had the least variance for finishing the task. Additionally, participants commented that both indirect methods were physically less straining than the direct method. In summary, the two-thumb gestures outperformed the common direct manipulation technique and the one-thumb gestures.

The presented study has only examined the performance of the bi-manual gestures on a big-sized mobile device. However, we are confident that the use on small devices such as smart phones shares the same benefits.

Acknowledgements

The research leading to these inventions has received funding from the European Union Seventh Framework Programm (FP7/2007-2013) under grant agreement n° 284573.

References

[1] S. Boring, D. Baur, A. Butz, S. Gustafson, and P. Baudisch. Touch projector: mobile interaction through video. In *Proc. CHI*, pages 2287–2296. ACM, 2010.

[2] M. Hancock, T. ten Cate, and S. Carpendale. Sticky tools: full 6DOF force-based interaction for multi-touch tables. In *Proc. ITS*, pages 133–140. ACM, 2009.

[3] C. Lee, G. A. Rincon, G. Meyer, T. Höllerer, and D. A. Bowman. The Effects of Visual Realism on Search Tasks in Mixed Reality Simulation. *Visualization and Computer Graphics*, 19(4):547–556, 2013.

[4] A. Mossel, B. Venditti, and H. Kaufmann. 3DTouch and HOMER-S: Intuitive Manipulation Techniques for One-Handed Handheld Augmented Reality. In *Proc. VRIC*, pages 12:1–12:10. ACM, 2013.

[5] J. Wagner, M. Nancel, S. G. Gustafson, S. Huot, and W. E. Mackay. Body-centric Design Space for Multi-surface Interaction. In *Proc. CHI*, pages 1299–1308. ACM, 2013.