3D mid-air manipulation techniques above stereoscopic tabletops

Fernando Fonseca
INESC-ID/IST/TU Lisbon
fernando.fonseca@ist.utl.pt

Daniel Mendes
INESC-ID/IST/TU Lisbon
danielmendes@ist.utl.pt

Bruno Araújo
INRIA/INESC-ID
bruno.araujo@inria.fr

Alfredo Ferreira
INESC-ID/IST/TU Lisbon
alfredo.ferreira@ist.utl.pt

Joaquim Jorge
INESC-ID/IST/TU Lisbon
jaj@inesc-id.pt

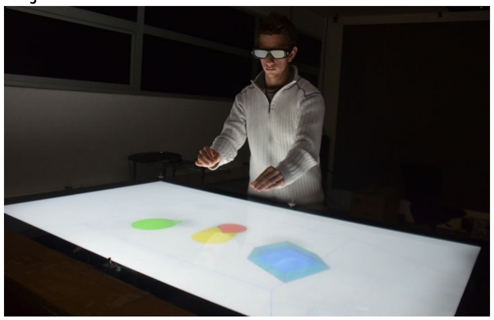

Figure 1: User manipulating objects above our interactive stereoscopic tabletop.

ISIS3D Workshop in conjunction with ITS'13, October 6, 2013, Scotland, UK.

DOI string from ACM form confirmation

Abstract
Stereoscopic tabletops collocate both user interaction and visualization spaces, fostering more natural freehand interaction metaphors to manipulate 3D virtual content. Thanks to a non-intrusive tracking solution, based on depth cameras, users can interact and manipulate virtual objects using both hands. In this position paper, we perform a comparative study of one multi-touch and four mid-air interaction techniques assigning the seven degrees of freedom (translation, rotation and uniform scaling) of 3D manipulation, symmetrically or asymmetrically between both hands.

Author Keywords
3D user interfaces, 3D manipulation, stereoscopic tabletops

ACM Classification Keywords
H.5.2 [User Interfaces]: Interaction styles, Graphical User Interfaces, Input devices and strategies

Introduction
The emergence of affordable 3D projectors, multi-touch technology and non-intrusive tracking solutions based on depth cameras to track user hands and head allows to devise stereoscopic interactive surfaces such as depicted by Figure 1. Such environment enables to design new interaction metaphors where users can interact more naturally with 3D virtual content than using traditional

displays and input devices such as the mouse and the keyboard. The stereoscopic visualization enables to create the illusion that 3D content is lying on top of an interactive surface, fostering direct manipulation methods using user hands. Such scenario is especially appealing for both 3D mock-up visualization and modeling or to support 3D assembly tasks.

In this paper, we explore a novel setup combining affordable non-intrusive tracking technologies with a multi-touch stereoscopic tabletop fostering mid-air hand gestures for 3D manipulations. Using such environment, we implemented different techniques for spatial 3D object manipulations, both direct and indirect, and then performed a comparative study between multi-touch and mid-air interactions in order to assess their adequacy, advantages and limitations to manipulate 3D objects.

Related Work

Several multi-touch solutions have been proposed to manipulate 3D objects as an alternative to traditional WIMP based approaches. To overcome the limitations of interacting with a planar surface to control six degrees of freedom (DOF), several approaches [7, 9, 13] try to mimic physically based manipulations using the number of direct touches in contact with the virtual object being manipulated. While Hancock et al. [7] control the object depth based on the distance between touches, Hilliges et al. [9] take advantages of the space above the surface by estimating the distance from user hands to the screen. Such approaches usually couple translation and rotation together making difficult to control independently degrees of freedom. The Sticky tools proposed by [8] minors such problem by combining direct touches on the virtual content with indirect touches to control the rotation separately. The same approach is followed by Martinet et al. [11] and compared with using multi-views from the 3D scene to separate positional degrees of freedom and to position objects more precisely in 3D space.

A different approach was followed by [4, 12, 15] using touches to interact with 2D graphical controllers. While LTouchIt [12] proposed to use direct touch for translation and widgets for rotations, the tBox [4] proposes a unique box representation around the manipulated object to control independently the object 9 DOF (translation, rotation and scale along each axis). The Triangle Cursor [15] presents a simpler widget controlled by two fingers. However it only enables to control 4 DOF and such as [12] cannot handle with scaling transformations. A similar widget based approach was followed by Daiber et al. [5] to interact with stereoscopic displays. While the distance between two fingers controls the position of the object above the stereoscopic display mimicking a floating balloon, a corkscrew metaphor using one finger allows to move objects below the surface. While this method only controls 3 DOF, the Toucheo [6] system allows users to effectively control 9 DOF and perform complex assembly tasks. However this is done at the cost of additional 2D widgets to control remaining rotations and scaling.

In our stereoscopic multi-touch tabletop environment, we take advantage of the non-intrusive hand tracking technique proposed by Wang et al. [16] which is able to track both hands 6 DOF and its finger, as well as simple hand postures such as pinching or pointing. While Wang et al. illustrate simple manipulations using hand gestures in front of a traditional display, we devise several methods to control the 3D object position, orientation and uniform scaling using a stereoscopic tabletop where both user hands and 3D virtual content are collocated. While hand free manipulations [10] have been proposed using physically based approaches collocating user hands with 3D content thanks to a semi transparent display and a

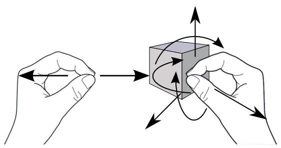

Figure 2: 6-DOF Hand.

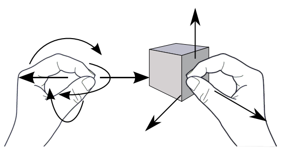

Figure 3: 3-DOF Hand.

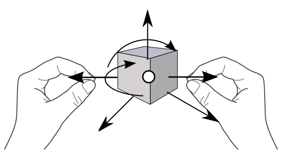

Figure 4: Handle-Bar.

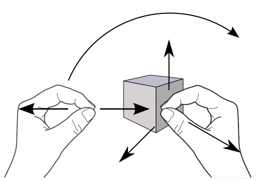

Figure 5: Air TRS.

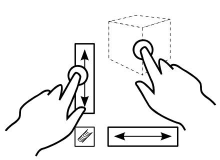

Figure 6: TRS Touch + Widgets.

depth camera tracking solution, we propose bi-manual techniques such as the one used by Mockup Builder [1] however without the need of instrumenting user hands.

Spatial Interactions

Using our setup, we implemented four mid-air manipulation techniques and one based only on touch. All of these techniques provide seven degrees of freedom (DOF): three for translation, three for rotation and a uniform scale, and use hand pinch gestures for selecting objects or starting a specific manipulation.

6-DOF Hand To mimic interactions with physical objects, as close as possible, we use all 6 DOF information provided by the tracker (3 DOF for position and 3 DOF for orientation). While grabbing directly the object with one hand maps the 6 DOF as depicted in Figure 2, an indirect grabbing with the second hand allows to uniformly scale the object based on the distance between both hands.

3-DOF Hand This technique separates translation from both rotation and scale between both hands as depicted by Figure 3. After grabbing the object with one hand, the user can translate it by moving that hand. The rotation is achieved by rotating the wrist corresponding to the other hand by grabbing somewhere in space, while keeping the object grabbed with the first hand. Such as the previous method, uniform scaling is controlled by the distance between both hands.

Handle-Bar Following the previous work of Song et al. [14], we implemented the Handle-bar metaphor in our stereoscopic tabletop setup. This approach mimics a familiar bimanual handle-bar, commonly used, for example, to roast a turkey. In this technique, the object close to the middle point of the user hands is selected as presented by Figure 4. Both translation and rotation can be controlled by moving the handle bar. Changes in the distance between both hands scale the object evenly.

Air TRS This method presented by Figure 5 is similar to the one used by the Mockup Builder [1] system. While grabbing with one hand translates the object, grabbing with both hands will define a two-point Translate-Rotate-Scale (TRS) in 3D.The rotation angle is defined by the variation in the position of one hand relatively to the other. For scaling, the distance between both hands is used.

TRS Touch + Widgets The last technique relies exclusively on touching the surface below the object of interest. We use a set of three widgets presented in Figure 6 to control the seven DOF. The first touch controls the position of the object on a X&Y plane and activates the widget interface. While using his second finger, the user can rotate around X and Z axes, using a rod metaphor [6] widget, and the height of the object, using a balloon like metaphor as the one used by [5]. If the second touch lies outside any widget both Y axes rotation and uniform scaling can be controlled.

Conclusions and Future Work

We carried out a user evaluation with twelve participants, in which we put to test each implemented technique via three tasks within a practical scenario and increasingly difficulty shown in Figure 8 after a training scenario (Figure 7) used to test each technique. Their mean age was 25 years old and all hold a bachelor degree. Most participants own a touch device, such as a smart-phone. However, only five of them had previous experience with stereoscopic visualization and only three had experience with 3D modeling applications. Participants to the user evaluation of the implemented techniques agreed that the 6 DOF Hand approach was more natural to use, since it reproduces the direct interactions with physical objects. Results also showed that regarding the time taken to fulfill task objectives, the Handle-Bar [14] solution was as fast

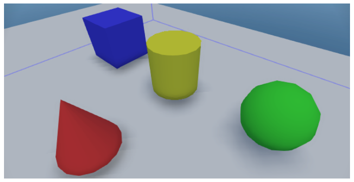

Figure 7: Training scenario for every interaction technique.

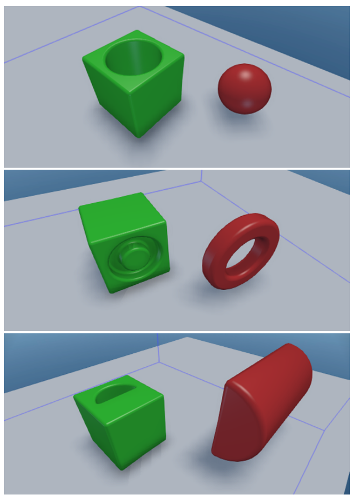

Figure 8: Tasks performed by the test participants

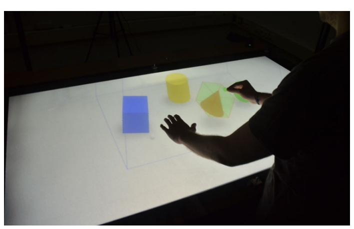

Figure 9: User manipulating objects in our training scenario.

as the 6 DOF Hand. Additionally, we observed that the 6 DOF Hand approach created unwanted occlusions, a consequence of stereoscopic displays already identified by [2, 3] that did not affect the Handle-Bar. We believe that overcoming such challenge will allow us to improve the technique, making it the more appropriate to manipulate three dimensional virtual objects in stereoscopic environments, and reducing the gap between virtual and physical interactions.

As future directions for our work, we intend to improve hand tracking precision, since the resolution offered by the Microsoft Kinect depth cameras is low and created some difficulties. Other devices are appearing, such as the Leap Motion or the newly announced Microsoft Kinect for the XBox One, with Full-HD resolution, which are worthy of attention and future explorations. We also plan to apply the spatial interactions presented here in real-world scenarios, such as fast prototyping for architectural models or assembling engineering parts. We feel that the main conclusion of this work is that stereoscopic visualization reinforces the affordances of direct manipulation and improves both the sense of control and familiarity to make interactions feel more intimate.

Acknowledgements The work described in this paper was partially supported by the Portuguese Foundation for Science and Technology (FCT) through the projects Alberti Digital PTDC/AUR-AQI/108274/2008, CEDAR PTDC/EIA-EIA/116070/2009 and TECTON-3D PTDC/EEI-SII/3154/2012, and under contract Pest-OE/EEI/LA0021/2013.

References

[1] Araùjo, B., Casiez, G., Jorge, J., and Hachet, M. Mockup builder: 3d modeling on and above the surface. *Computers & Graphics 37*, 3 (2013), 165 – 178.

[2] Bruder, G., Steinicke, F., and Stuerzlinger, W. Effects of visual conflicts on 3d selection task performance in stereoscopic display environments. In *Proceedings of 3DUI'13*, IEEE (2013).

[3] Chan, L., Kao, H., Chen, M., Lee, M., Hsu, J., and Hung, Y. Touching the void: direct-touch interaction for intangible displays. In *Proc. of CHI'10*, ACM (2010), 2625–2634.

[4] Cohé, A., Dècle, F., and Hachet, M. tbox: a 3d transformation widget designed for touch-screens. In *Proc. of CHI '11*, ACM (NY,USA, 2011), 3005–3008.

[5] Daiber, F., Falk, E., and Krüger, A. Balloon selection revisited: multi-touch selection techniques for stereoscopic data. In *Proceedings of AVI '12*, ACM (NY, USA, 2012), 441–444.

[6] Hachet, M., Bossavit, B., Cohé, A., and de la Rivière, J.-B. Toucheo: multitouch and stereo combined in a seamless workspace. In *Proc. of UIST'11*, ACM (2011), 587–592.

[7] Hancock, M., Carpendale, S., and Cockburn, A. Shallow-depth 3d interaction: design and evaluation of one-, two- and three-touch techniques. In *Proc. of CHI '07*, ACM (NY, USA, 2007), 1147–1156.

[8] Hancock, M., ten Cate, T., and Carpendale, S. Sticky tools: full 6dof force-based interaction for multi-touch tables. In *Proceedings of ITS '09*, ACM (NY, USA, 2009), 133–140.

[9] Hilliges, O., Izadi, S., Wilson, A. D., Hodges, S., Garcia-Mendoza, A., and Butz, A. Interactions in the air: adding further depth to interactive tabletops. In *Proc. of UIST '09*, ACM (NY, USA, 2009), 139–148.

[10] Hilliges, O., Kim, D., Izadi, S., Weiss, M., and Wilson, A. Holodesk: direct 3d interactions with a situated see-through display. In *Proceedings of CHI'12*, ACM (2012), 2421–2430.

[11] Martinet, A., Casiez, G., and Grisoni, L. The design and evaluation of 3d positioning techniques for multi-touch displays. In *Proc. of 3DUI'10*, IEEE (2010), 115–118.

[12] Mendes, D., Lopes, P., and Ferreira, A. Hands-on interactive tabletop lego application. In *Proceedings of ACE'11*, ACM (NY, USA, 2011), 19:1–19:8.

[13] Reisman, J. L., Davidson, P. L., and Han, J. Y. A screen-space formulation for 2d and 3d direct manipulation. In *Proceedings of UIST '09*, ACM (NY, USA, 2009), 69–78.

[14] Song, P., Goh, W., Hutama, W., Fu, C., and Liu, X. A handle bar metaphor for virtual object manipulation with mid-air interaction. In *Proc. of CHI '12*, ACM (2012), 1297–1306.

[15] Strothoff, S., Valkov, D., and Hinrichs, K. Triangle cursor: interactions with objects above the tabletop. In *Proc of ITS '11*, ACM (NY, USA, 2011), 111–119.

[16] Wang, R., Paris, S., and Popović, J. 6d hands: markerless hand-tracking for computer aided design. In *Proc of UIST'11*, ACM (NY,USA, 2011), 549–558.