

The Design and Evaluation of 3D Positioning Techniques for Multi-touch Displays

Anthony Martinet*

Géry Casiez †
LIFL & INRIA Lille
University of Lille, France

Laurent Grisoni ‡

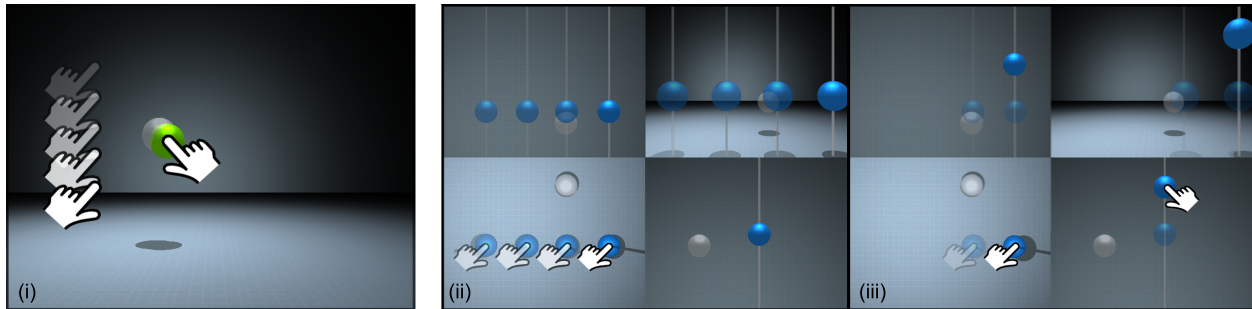


Figure 1: (i) Illustration of the Z-technique. The first finger (right hand in the example) is used for direct positioning in the camera plane while the second finger (left hand) is used for depth positioning in an indirect way. Backward-forward movements move the object farther or closer to the user. (ii and iii) Illustration of the multi-touch viewport technique. The first finger (left hand) is used for 2D positioning in the camera plane corresponding to the viewport (ii) while the second finger (right hand) is used to move the object along the third coordinate (iii). The gray line is used as visual feedback to represent the displacement allowed for the second finger.

ABSTRACT

Multi-touch displays represent a promising technology for the display and manipulation of 3D data. To fully exploit their capabilities, appropriate interaction techniques must be designed. In this paper, we explore the design of free 3D positioning techniques for multi-touch displays to exploit the additional degrees of freedom provided by this technology. Our contribution is two-fold: first we present an interaction technique to extend the standard four viewports technique found in commercial CAD applications, and second we introduce a technique designed to allow free 3D positioning with a single view of the scene. The two techniques were evaluated in a preliminary experiment. The first results incline us to conclude that the two techniques are equivalent in term of performance showing that the Z-technique provides a real alternative to the status quo viewport technique.

Index Terms: H.5.2 [User Interfaces]: Input devices and strategies.

1 INTRODUCTION

Multi-touch input enables users to manipulate and display information by using fingers and hand gestures, with an interaction potential that has never been reached before [12]. Multi-touch displays currently offer a large set of direct interaction techniques using fingers to translate, rotate and scale two dimensional (2D) objects [7].

No interaction technique is currently accepted yet to exploit multi-touch displays for translating objects in 3D environments (although some methods exist, see next section). The development of techniques for multi-touch displays focuses mainly on collab-

orative working environments and 3D object rotation [6]. In this article, we explore the development of 3D positioning techniques for multi-touch displays.

We present related work on 3D manipulation in next section. In section 3 we present two new interaction techniques for 3D positioning and evaluate them in a controlled experiment described in section 4 and section 5.

2 RELATED WORK

A lot of work has been achieved on object manipulation into 3D environments. We categorize previous works into two sets: multi-touch oriented techniques, and generic interaction systems (i.e. using classical mouse pointing systems).

2.1 3D Manipulation with Multi-Touch Inputs

Hancock et al. [6] presented some input interaction techniques to manipulate 3D objects on tabletop displays with limited depth. They discuss interaction using one, two and three fingers. With one touch input, their technique achieves 5 degrees of freedom (DOF) movement by extending the RNT algorithm [8] for 3D rotations coupled with 2D positioning. Two-touch interaction allows to perform pitch and roll rotations with a second finger on the non-dominant hand. With three-touch interaction, users can perform a simultaneous translation and rotation on the surface of the table. Positioning along the depth is proposed as an option by measuring the distance between two fingers. However the technique was mainly designed to control the orientation of a 3D object with 2D positioning. With the depth control position enabled, modifying the orientation of an object changes its depth position without any control. The three-touch technique was shown to be fastest and most accurate as well as preferred by the users. Also on 3D manipulation, Wilson et al. [14] presented a physics-enabled 3D environment with multi-touch input. Their technique is able of modelling both multiple contact points and more sophisticated shapes enabling a large variety of interactions. However adding physics to the interaction allows to rotate objects in a more realistic way but positioning remains limited to two DOF. Finally, Reisman et al. [11] propose

*e-mail: anthony.martinet@lifl.fr

†e-mail: gery.casiez@lifl.fr

‡e-mail: laurent.grisoni@lifl.fr

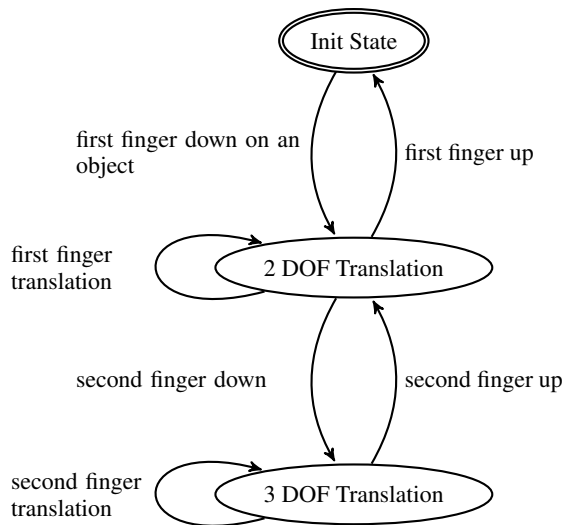


Figure 2: State transition diagram representing the techniques

some energy-minimisation technique for adapting 2D interaction to 3D transformation.

2.2 3D Positioning with Mouse-Related Devices

Mouse-based positioning techniques can be classified into two categories. First, constraint based positioning where the position of the manipulated object is constrained by the attributes of the other objects of the scene [3, 4, 10]. In spite of their widespread use in CAD and sketching software, these techniques introduce visual discontinuities when jumping from visual feature to visual feature making them difficult to use in complex scenes.

Second, free positioning where the manipulated object can be freely positioned in 3D space [9, 2, 13]. One technique (defined as *four-views technique* in the following sections) consists in displaying on screen 3 orthogonal views and one perspective view of the 3D scene. The object position is controlled two DOF at a time using one of the orthogonal views. This technique represents the most widely used one to freely position 3D objects and has been shown to be faster than the handles technique with a single view [10].

To sum up, interaction techniques designed for multi-touch displays allow 3D objects positioning constrained to the screen plane and 3D rotations. 3D positioning techniques using two integral DOF can be separated into constrained and free positioning technique. The four-views techniques represents the most widely used technique for the latter category.

3 PROPOSED INTERACTION METHODS

Our primary goal is to provide users with free 3D positioning techniques using multi-touch displays. We propose two techniques that take advantage of the direct interaction provided by multi-touch displays to control object position in the screen plane. The two techniques proposed explore two different ways of controlling the third degree of freedom: one using direct control and the other using indirect control. The state transition diagram in Figure 2 describes both techniques. It is based on *finger down*, *finger up* and *finger move* events coming from the table SDK.

3.1 Multi-touch viewport

Most CAD software (Autocad, Archicad, Blender ...) propose a display technique splitting the screen (previously described as *four-views technique*). We add multi-touch support enabling 3 DOF manipulation for free 3D positioning. In our case, the screen is divided

into four viewports of equal size and each viewport uses the same field of view for its virtual camera.

The two top states in Figure 2 correspond to the standard viewport technique. When a finger first touches the surface, the corresponding viewport is detected and a ray orthogonal to the view is projected into the scene. The first object in intersection is returned. Subsequent finger movements move the object in the plane parallel to the view passing through the object center.

The multi-touch viewport extension is activated when a second finger touches **another viewport**. On the *finger down* event, the object is first moved to the depth position associated to the second finger position. The user can then adjust the depth position in the direction orthogonal to the view associated to the first finger. We used a visual feedback represented as a line passing through the object center to represent this direction.

3.2 The Z-technique

We designed the second technique to allow 3D positioning using a single view of the scene. We believe this presentation increases the immersion of the user in the 3D environment compared to the multi-touch viewport technique. This technique is inspired from both Venolia and Benko works on 3D interaction [13, 1].

When the first finger is in contact with the surface, a ray coming from the camera center through the finger position is cast in the scene and the closest object in intersection is returned. Positioning is then constrained to the plane parallel to the camera plane passing through the object center.

When a second finger is in contact with the surface we measure its relative motion on the surface and use backward forward movement to control the depth position. Backward - forward movements are measured relative to the user position (in our implementation, we assume that the user position is known). Forward movement moves the object away from the user view and backward movement moves it closer to the user's view. As we measure the relative motion, a succession of finger touches and releases (*clutching*) can be used to position the object at any arbitrary depth position.

As we control the depth position in an indirect way, we use a transfer function to map the finger movement to object displacement. We use a non-linear continuous function that maps the displacement speed to a scale factor [5]. Low finger displacement speed is associated to a low scale factor; high finger displacement speed is associated to a much higher scale factor. In practice, such a mapping provides the user with either accurate positioning or large displacements.

4 PRELIMINARY EXPERIMENT

4.1 Goals

The two techniques are similar for the direct positioning of objects in the plane parallel to the screen but they mainly differ on two aspects : the display of the 3D scene (four views for multi-touch viewport, a single perspective view for the Z-technique) and the control for the depth position (direct interaction for multi-touch viewport, indirect interaction for the Z-technique).

We wanted to investigate the impact of the information presentation and the directness of control for depth positioning on both performance and qualitative feedback. The Z-technique encourages the simultaneous control of the 3 DOF while the multi-touch viewport seems to favour iterative control in each viewport in spite of the multi-touch support we added to the technique. We also expected the Z-technique to increase the feeling of immersion in the environment as the technique presents a single perspective view in full screen compared to the multi-touch viewport where the user focuses on each orthographic view. Finally considering the visual presentation and the directness of control for each technique we wanted to evaluate the influence of the precision required for positioning as a confounding factor.

4.2 Apparatus

The experiment was conducted on an iLight touch table¹ based on the Diffused Illumination technique and consisting of a 100 cm × 70 cm (42 inches) monoscopic display positioned at a distance of 105 cm from the floor. The video-projector under the table was set at 60 Hz with a 1400 × 1050 pixels resolution giving a pixel density of 14 pixels per cm (36 DPI). A camera running at 120 Hz with a 640 × 480 pixels resolution and positioned under the surface recorded the finger movements on the surface. Such a hardware configuration provides a maximal resolution of 6.4 dots per cm (16.25 DPI) for finger tracking.

We used the iLight framework version 1.6 for fingers detection and tracking. Fingers data were sent using TUIO messages to a custom built 3D application based on Ogre3D².

4.3 Task

The task is a three dimensional positioning task based on the docking task introduced by Zhai [15] (Figure 1). Each experimental trial began after the previous target was successfully positioned, and ended with the successful positioning of the current target. Participants were asked to dock a blue sphere in a transparent grey sphere as quick as possible. The blue sphere turned green when well positioned inside the grey sphere. The trial was considered as fully completed when the distance between the two spheres remained below a threshold distance during 0.8s. If the sphere was moved beyond this threshold during these 0.8s, the sphere turned back from green to blue. When the trial was completed, the grey sphere then disappeared and appeared to another position while the green sphere was repositioned at the center of the screen. In addition to perspective and occlusion, we added a ground with shadows projection to improve depth perception. The camera remained static during the whole experiment.

4.4 Participants

Eight males with a mean age of 24 (SD 1.5) participated. 7 were right-handed and 1 was left-handed. All had normal or corrected to normal vision. Participants had a variable experience with virtual reality and 3D applications but this is acceptable as we target experiment to a fundamental physical behaviour. Half of them were frequent users of the touch-table. Others were familiar with tactile technology such as tactile mobile phone but not with multi-touch displays.

4.5 Design

A repeated measures within-subjects design was used. The independent variables were TECHNIQUE (Z-technique and multi-touch viewport), POSITION of the target sphere with 17 positions evenly distributed around a hemi-sphere, and ACCURACY with two levels (low accuracy and high accuracy).

Participants completed five successive BLOCKS of trials. Each BLOCK consisted in 68 trials: 2 repetition of 34 POSITION × ACCURACY combinations. The presentation order of TECHNIQUE was counter-balanced across participants. A break was encouraged after each set of 17 trials.

Before starting the experiment, participants had a 15 minutes training period to get used to the techniques and task. The experiment ended with a qualitative feedback from the participants. The experiment lasted approximately 35 minutes.

5 FIRST RESULTS

The dependent variables for both techniques were *positioning time* and *coordination value*. We also introduced a third variable : the

¹<http://www.immersion.fr>

²<http://www.ogre3d.org>

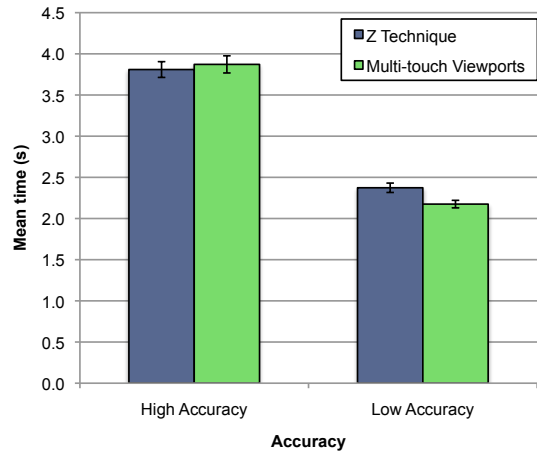


Figure 3: Mean positioning time for each technique and accuracy. Error bars represent 95% confidence interval.

frequency of use of two viewports for the multi-touch viewport technique.

5.1 Positioning Time

Positioning time was defined as the time it takes to successfully position the current sphere from the last successfully positioned target.

Repeated measures analyses of variance found a significant main effect for BLOCK ($F_{3,21} = 5.68$, $p = 0.016$) which shows the presence of a learning effect. Pairwise comparisons showed significant differences ($p < 0.05$) between the first block and the other blocks. As a result we removed the first block for subsequent analysis.

We did not find a significant main effect of TECHNIQUE on the positioning time. The mean positioning time for the Z-technique is 3.10s and 3.03s for the multi-touch viewport technique (Figure 3). A power analysis reveals that our experiment has enough power to detect an effect size equal to 0.75, which represents a difference between the techniques equal to 5%. Further evaluations would be required to prove that no small effect (a difference below 5%) exists.

Repeated measures analyses of variance show a significant main effect of ACCURACY ($F_{1,7} = 330$, $p < 0.0001$) on positioning time. As expected, higher accuracy significantly increases the positioning time (from 2.28s for the low accuracy condition to 3.86s for the high accuracy condition). Interestingly, we found a significant interaction between TECHNIQUE and ACCURACY ($F_{1,7} = 6.38$, $p = 0.039$) with the Z-technique being slower for the low accuracy condition (2.38s) compared to the multi-touch viewport technique (2.18s) ($p = 0.026$). During the experiment we observed that participants could directly position the sphere at the right location by tapping on a second viewport whereas adjustments were required for high accuracy. In contrast, the participants kept a consistent behaviour with the Z-technique no matter the accuracy level.

We found a significant effect for POSITION ($F_{16,112} = 8.11$, $p < 0.0001$) on positioning time but no significant interaction with other factors. This shows that the positioning is faster for the targets requiring no displacement along the depth.

5.2 Coordination

The coordination is computed as the ratio of the length of the shortest path to the length of the actual path measured from the starting point to the destination [15].

Repeated measures analyses of variance did not show any significant main effect or interaction for the different independent variables on the coordination except for POSITION ($F_{16,112} = 24.5$, $p < 0.0001$) showing a higher coordination for the targets requiring

no depth adjustment. The mean coordination value is 0.59 for the Z-technique and 0.62 for the multi-touch viewport technique.

5.3 Use of a Second Viewport

For the multi-touch viewport technique, we analyzed the use of a second viewport to move and adjust depth position. A second viewport is considered as used when two viewports are touched simultaneously at least one time during an individual trial. These trials are marked as 1 and used for computation. On average, a second viewport is used 92% of the time. Repeated measures analyses of variance find a marginal significant main effect of ACCURACY ($F_{1,7} = 5.35$, $p < 0.054$) on the use of a second viewport with the second viewport used more for the low accuracy condition (95.5%) compared to the high accuracy condition (88.1%).

We also measured the frequency of use for each viewport and found that the two viewports closer to the user are used 99.5 % of the time.

5.4 Qualitative Feedback

Among the participants, six preferred the Z-technique, one the multi-touch viewport technique and one claimed to have no preference. The participant who declared to prefer the multi-touch viewport reported a lot of experience with CAD applications using the viewport technique. The participants who declared to prefer the Z-technique found the technique easier to use and appreciated the interaction with the perspective view. The indirect mapping for depth positioning is, according to them, difficult to handle at first, but very efficient once they get used to it.

For the Z-technique, all participants used their dominant hand to control positioning in the camera plane while depth positioning is controlled using non-dominant hand. Two participants showed frequent confusion in the direction to use for depth positioning with backward movements expecting objects to move forward. In a real application, this parameter should be a control setting to be customized for each user.

6 DISCUSSION

We did not find any significant difference for time between the two techniques. Of course this does not mean that the two techniques are strictly equivalent. We have designed an experiment that power analysis revealed to be able to detect a 5% time difference if such exists. Because we failed to find a significant effect, we think it is quite unlikely that one exists. Further experiments would still be required to detect a smaller difference.

We also observed that the multi-touch viewport significantly improves performance with low accuracy thanks to the use of a second viewport that snaps the object at the right depth. In addition the results show that participants emphasize the simultaneous use of two viewports with targets requiring coarse positioning. This suggests that combining two viewports for 3D positioning helps improving performance when high positioning accuracy is not required.

During our experiment, participants used two viewports simultaneously more than 90% of the time, making an effective use of the additional features provided by the multi-touch viewport technique. We also showed that the two top viewports were hardly used with our hardware configuration, suggesting that participants try to reduce their fatigue using the two viewports closer to them. This suggests that the perspective view can be extended to the entire width of the upper half of the screen.

Oh et al. [10] showed that the classic four-view technique is significantly faster compared to other techniques using full screen in a manipulation task. Here we provide the Z-technique that reaches a similar level of performance compared to an improved version of the four-views technique. The interaction with the perspective view in full screen was in addition preferred by most participants. This

makes the Z-technique a real alternative to the multi-touch viewport, especially when the display surface is limited (e.g. mobile phones).

7 CONCLUSION AND FUTURE WORK

In this paper, we have introduced and compared two 3D positioning techniques that allow to freely position an object using a multi-touch display. In a preliminary experiment we showed that these two techniques reach similar level of performance. Furthermore the Z-technique was preferred by most participants. Considering its performance and qualitative feedback, the Z-technique can be used as a base line comparison for other 3D positioning techniques designed for multi-touch displays. As future work, we plan to add the support for objects rotation to control more degrees of freedom.

ACKNOWLEDGMENTS

This work was supported by the ANR project n° ANR-09-CORD-013 "InSTInCT".

REFERENCES

- [1] H. Benko and S. Feiner. Balloon selection: A multi-finger technique for accurate low-fatigue 3d selection. In *3D User Interfaces, 2007. 3DUI '07. IEEE Symposium on*, page 7986, March 2007.
- [2] E. A. Bier. Skitters and jacks: interactive 3D positioning tools. In *Proceedings of the 1986 workshop on Interactive 3D graphics*, pages 183–196, 1987.
- [3] E. A. Bier. Snap-dragging in three dimensions. In *Proceedings of the 1990 symposium on Interactive 3D graphics*, pages 193–204, 1990.
- [4] R. W. Bukowski and C. H. Séquin. Object associations: a simple and practical approach to virtual 3D manipulation. In *Proc. of Symposium on Interactive 3D graphics*, pages 131–138., 1995.
- [5] G. Casiez, D. Vogel, R. Balakrishnan, and A. Cockburn. The impact of control-display gain on user performance in pointing tasks. *Human-Computer Interaction*, 23(3):215 – 250, 2008.
- [6] M. Hancock, S. Carpendale, and A. Cockburn. Shallow-depth 3d interaction: design and evaluation of one-, two- and three-touch techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1147–1156, 2007.
- [7] M. S. Hancock, S. Carpendale, F. D. Vernier, D. Wigdor, and C. Shen. Rotation and translation mechanisms for tabletop interaction. In *Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, pages 79–88, 2006.
- [8] R. Kruger, S. Carpendale, S. D. Scott, and A. Tang. Fluid integration of rotation and translation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 601–610, 2005.
- [9] G. M. Nielson and J. D. R. Olsen. Direct manipulation techniques for 3D objects using 2D locator devices. In *Proc. Workshop on Interactive 3D graphics*, pages 175–182, 1986.
- [10] J. Oh and W. Stuerzlinger. Moving objects with 2D input devices in CAD systems and desktop virtual environments. In *Proceedings of Graphics Interface*, pages 195–202, 2005.
- [11] J. L. Reisman, P. L. Davidson, and J. Y. Han. A screen-space formulation for 2d and 3d direct manipulation. In *UIST '09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pages 69–78, New York, NY, USA, 2009. ACM.
- [12] T. Selker. Touching the future. *Commun. ACM*, 51(12):14–16, 2008.
- [13] D. Venolia. Facile 3D direct manipulation. In *Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems*, pages 31–36, 1993.
- [14] A. D. Wilson, S. Izadi, O. Hilliges, A. Garcia-Mendoza, and D. Kirk. Bringing physics to the surface. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 67–76, 2008.
- [15] S. Zhai and P. Milgram. Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 320–327, 1998.