Combining Bimanual Interaction and Teleportation for 3D Manipulation on Multi-Touch Wall-sized Displays

Jean-Baptiste Louvet, Cédric Fleury

Combining Bimanual Interaction and Teleportation for 3D Manipulation on Multi-Touch Wall-sized Displays

Jean-Baptiste Louvet [1,2,3] *

Cédric Fleury [2,1] †

[1] Inria, Université Paris-Saclay, 91405, Orsay, France
[2] LRI, Univ. Paris-Sud, CNRS, Université Paris-Saclay, 91405, Orsay, France
[3] Normandie Univ, INSA Rouen, UNIHAVRE, UNIROUEN, LITIS, 76000 Rouen, France
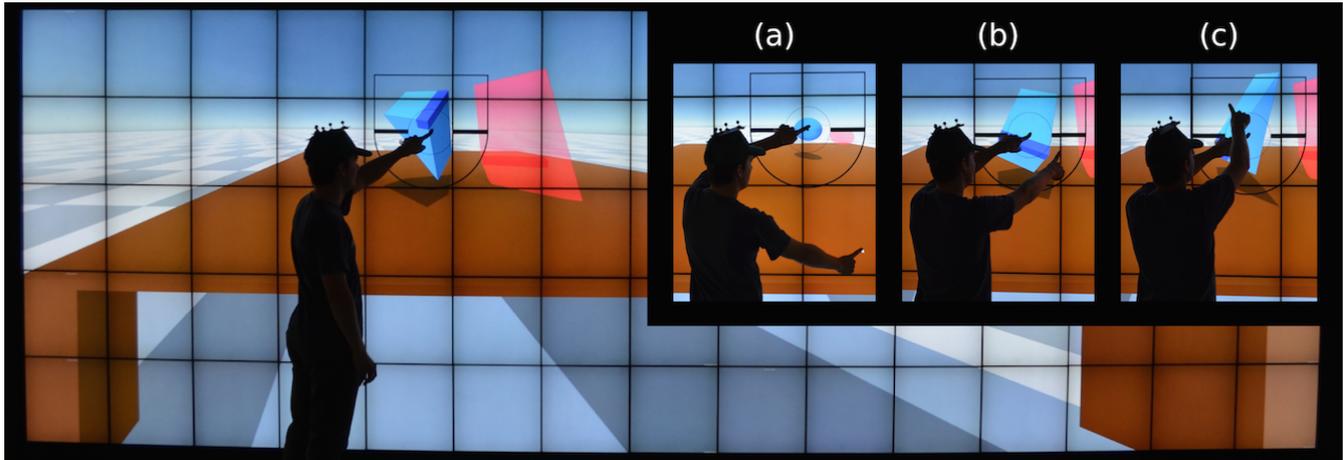
Figure 1: *6 degrees of freedom manipulation of a 3D object on a multi-touch wall-sized display combining bimanual interaction and teleportation. The user is performing a xy translation (main picture), z translation (a), roll rotation (b), and pitch & yaw rotation (c).*

Abstract

While multi-touch devices are well established in our everyday life, they are currently becoming larger and larger. Large screens such as wall-sized displays are now equipped with multi-touch capabilities. Multi-touch wall-sized displays will become widespread in a near future in various places such as public places or meeting rooms. These new devices are an interesting opportunity to interact with 3D virtual environments: the large display surface offers a good immersion, while the multi-touch capabilities could make interaction with 3D content accessible to the general public.

In this paper, we aim to explore touch-based 3D interaction in the situation where users are immersed in a 3D virtual environment and move in front of a vertical wall-sized display. We design In(SITE), a bimanual touch-based technique combined with object teleportation features which enables users to interact on a large wall-sized display. This technique is compared with a standard 3D interaction technique for performing 6 degrees of freedom manipulation tasks on a wall-sized display. The results of two controlled experiments show that participants can reach the same level of performance for completion time and a better precision for fine adjustments of object position with the In(SITE) technique. They also suggest that combining object teleportation with both techniques improves translations in terms of ease of use, fatigue, and user preference.

Keywords: multi-touch interaction, 3D manipulation, wall-sized display, virtual reality

Concepts: •**Human-centered computing** → **Virtual reality; Touch screens;** *Usability testing;*

1 Introduction

Even if touch has been reserved for small devices such as smartphones and tablets in its early years, more and more large screens, such as computer screens, televisions, whiteboards, and wall-sized displays, are now equipped with multi-touch capabilities. Large multi-touch screens will become widespread in a near future in various locations such as public places, meeting rooms, classrooms. For example, we can already see such large multi-touch displays in shopping malls or airports to enable visitors to browse information. At the same time, Microsoft is currently launching its new Surface Hub[1] which is designed for meeting rooms.

Large screen displays are powerful tools to visualize the increasing amount of data from science, industry, business, and society. They are also appropriate to support collaboration among small groups of users. However, these benefits should not be limited to 2D content.

---

*e-mail: jeanbaptiste.louvet@insa-rouen.fr
†e-mail: cedric.fleury@lri.fr

---

[1] https://www.microsoft.com/microsoft-surface-hub/

Large multi-touch devices are a relevant solution to interact with 3D virtual environments. While the large display surface increases immersion in virtual environments, multi-touch input is a simple and efficient way to perform 3D interaction. Touch-based interaction can be easy to use and to learn for none-expert users since most of the people are familiar with touch devices. It also does not require additional external equipment since everything is embedded in the display device, which makes it particularly suitable for public areas, classrooms, and meeting rooms.

Most of the current touch-based 3D interaction techniques are designed either for small screens or consider that users stay static in front of the screen. Consequently, these techniques are not well adapted to the situation where users are immersed in a 3D virtual environment and freely move in front of a large vertical screen. In particular, finger combinations could be hard to achieve if users have to perform actions at the top or the bottom of the screen. In this paper, we propose In(SITE), a bimanual interaction technique designed to perform 6-DOF manipulation of 3D virtual objects using multi-touch input on a wall-sized display. This technique takes advantage of the large interaction space available on the display, in addition to the fact that users can easily move along the screen and perform actions with their two hands. On large wall-sized displays, it can be uncomfortable and inefficient to have to drag objects all along the screen when performing large translations. Consequently, we also propose to combine the In(SITE) technique with object teleportation features in order to overcome this drawback.

Two controlled experiments were performed to assess the usability of the In(SITE) technique in comparison with a standard 3D interaction technique which is used as a baseline for interaction in virtual reality. The first experiment evaluates the two techniques for 3D manipulation tasks which require only translations, while the second experiment focuses mainly on rotations. The first experiment also compares both techniques with and without teleportation features to determine the benefits of teleportation for each technique.

After reviewing the related work (section 2), we present the In(SITE) technique (section 3). Section 4 describes the usability testing of the In(SITE) technique through two controlled experiments. Then, we discuss the implications for touch-based 3D interaction on wall-sized displays (section 5). Finally, we conclude and propose a set of future research directions for 3D interaction on multi-touch wall-sized displays (section 6).

2 Related Work

The increasing accessibility to multi-touch technology has drawn the attention of researchers on the capabilities of touch-based interaction for 3D manipulation. On a state of the art about 3D interaction, Jankowski and Hachet [2015] point out that "*3D transformation widgets need to be reinvented to adapt to the tactile paradigm*", as users of touchscreens do not have access to an unobstructed view of the screen, an accurate pointing, and a direct access to buttons. The emergence of large wall-sized displays with multi-touch capabilities also introduces new constraints which should be taken into account when designing 3D touch-based interaction techniques.

2.1 3D Touch-based Interaction

Multiple approaches using touch input for 3D interaction have been explored. Cohé and Hachet [2012] conducted an experiment to understand how non-technical users *a-priori* interact with multi-touch screens to manipulate 3D objects. The experiment was conducted by showing participants a video of a cube transformation (rotation, scaling, translation) and then asking them to draw a gesture that best matches this transformation on a static image displayed on a

touchscreen. The results of the study made them define a taxonomy based on the analysis of the users' gestures. It contains different strategies applied by participants to perform the actions for rotation, scaling, and translation. This study gave us some guidelines on how to design the interaction technique presented in this paper.

Some techniques try to mimic direct manipulation of an object in 3D such as the *screen-space* technique presented by Reisman et al. [2009]. This technique was designed to ensure that the screen-space projection of the object-space point "touched" by users always remains under their fingertip. It uses a constraint solver to minimize a function that measures the error between the points in the object local-space and the points in screen-space and update the object position and rotation according to the result. Even if these manipulation techniques seem more intuitive, the users can have troubles to perform specific manipulation (especially for rotation) because it is not always easy to predict how the manipulated object will react and to determine what is the correct action to perform. In a similar way, *sticky fingers* presented by Hancock et al. [2009] is a force-based interaction technique designed to maintain the feeling of physical interaction with a virtual object. It is a transposition of the 2D multi-touch interaction paradigm (rotate, spin, scale) to 3D interaction. Jackson et al. [2012] proposes to extend touch-based interface techniques by using hand gestures above the surface.

Some other techniques use multiple simultaneous touch inputs from several fingers or/and from both hands to control each one of the 6 DOF. In particular, Hancock et al. [2007] have tested different 3D manipulation techniques in "shallow-depth" (i.e., in limited depth) involving 1, 2, or 3 touch points, making possible to control 5 or 6 DOF. The user study carried out shows that user preference and performance are higher with techniques involving multiple touch points which enable users to separate DOF. To facilitate the control of the depth, the *multi-touch viewport* presented by Martinet et al. [2010a] divides the screen in four viewports, each one corresponding to a different viewpoint on the 3D scene. Interacting in the first viewport allows a 2 DOF translation of the object, while interacting at the same time in a second viewport makes it possible to control the third DOF of translation of the object. Martinet et al. [2010a] also proposed the *Z-technique* in order to only use a single view of the scene for 3D positioning. When a first finger touches the screen, a raycast is done from the camera center and passing by the touch position and the first object intersected can then be translated in the plane parallel to the camera plane passing through the object center. The depth of the manipulated object is controlled in an indirect way using a second finger. Its up and down movements on the screen are mapped to forward and backward movements of the object relatively to the user position. A docking task experiment did not show significant results on performance but the *Z-technique* was preferred by a majority of participants in comparison to the *multi-touch viewport* technique. Dividing the screen into four viewports does not seem a solution suitable for large wall-sized displays. In addition, Martinet et al. [2010b] combined the *Z-technique* to the *screen-space* technique [Reisman et al. 2009] into a technique called *DS3* and compared it to the *sticky fingers* [Hancock et al. 2009] and *screen-space* technique. The results of this study show that separating the control of translation and rotation, as it is done with DS3, makes the interaction significantly faster.

To improve the understanding of 3D manipulation and the separation of the 6 DOF, Cohé et al. [2011] introduced *tBox*, a 3D transformation widget designed for touchscreens. This widget is inspired by standard box-shaped 3D transformation widgets operated from mouse and keyboard in desktop 3D applications. The widget is a wireframe box that appears on top of the manipulated object. The rotation is carried out with one finger, rotating the widget (with the object) around one of its primary axes, depending of the direction of the gesture. The translation is carried out with one finger on an

edge of the widget, allowing to translate the object in the direction of the selected edge. The scaling is carried out on the direction of the primary axes of the widget, with two fingers selecting two opposed edges of a face and and moving them away from each other.

2.2 Touch-based Interaction on Large Displays

Several previous works started to explore touch-based interaction techniques for 3D manipulation in the context of large vertical displays. For example, Yu et al. [2010] proposed *FI3D*, a touch-based interaction technique to navigate in a 3D scene on a large vertical screen. Traditional pan and zoom can be performed with 2-finger interaction with the display, but manipulation of other DOF can be achieved either by starting the drag gesture on the frame of the screen or just by touching the frame of the screen with the second finger. However, this technique is specific to the exploration of a single object in the 3D scene which does not require any object selection. Touching the frame of the screen is also not suitable with wall-sized displays because they are too large. In most recent work, Lopez et al. [2016] presented a variation of the previous technique in which the *FI3D* technique is used on a separate tablet to explore data displayed a large stereoscopic display. While this variation enables the user to control data on a larger display and deals with stereoscopic screen, the user has to frequently switch between the tablet view and the stereoscopic view. It also requires that the user hold an additional device, which does not seem appropriate when large displays are used in public places. In a similar way, Coffey et al. [2012] manipulated 3D medical data displayed on a large stereoscopic screen through 2D widget on a tabletop. Consequently, this technique implies indirect manipulation of the data displayed on the stereoscopic screen and also requires an additional device. Finally, Gilliot et al. [2014] presented a touch-based interaction technique called *WallPad* to perform direct and indirect manipulation on large wall-sized displays. However, this technique mainly focuses on touch-based interaction with a 2D graphical user interface.

2.3 Touch-based Interaction on Stereoscopic Displays

Some other related work focus on 3D touch-based interaction in the context of stereoscopic vision. For example, Benko and Feiner [2007] introduced the *Balloon Selection* which makes the selection of 3D objects possible in augmented reality settings. This multi-touch interaction technique consists in controlling the 2D position in the screen plane with a finger and to adjust the depth with the finger of the second hand using the metaphor of the string of an helium balloon which can be pulled down or released according to the second finger position. This technique only deals with translations and not with rotations. To improve the technique and to perform rotations, Strothoff et al. [2011] proposed a variation of the *Balloon Selection* called the *Triangle Cursor* in which two fingers of the user (same hand) define the base of an isosceles triangle which is perpendicular to the screen. The top of the triangle is the cursor for selection and manipulation. The user can control the height and the orientation of the triangle by moving these two fingers. The user can performed the additional rotation with his second hand. However, only objects which "come out" of the screen can be selected with both of these techniques. Consequently, they cannot be useful in the general context of immersive systems in which the 3D scene stands behind and in front of the display area.

On a more theoretical point of view, Valkov et al. [2011] studied touch-based interaction with a 2D screen to interact with 3D stereoscopic objects. They showed that users tend to touch between the projections for the two eyes with an offset towards the projection for the dominant eye. They also provided guidelines on how to design touch-based interaction for systems with 3D stereoscopic vision.

2.4 Synthesis

To sum up, many different touch-based interaction techniques have been proposed to perform 6 DOF manipulation of 3D objects. Several exhaustive studies have been achieved to better understand 3D touch-based interaction and provide useful guidelines to design a 3D touch-based interaction technique. In particular, Hancock et al. [2007] provide as guidelines to separate DOF and enable users to simultaneously control them. Even if some previous works have explored large and/or stereoscopic displays, most of the proposed techniques consider systems in which users stay static in front of the display and none of them deal with large wall-sized display in front of which users can freely move. In particular, touch input techniques which combine several fingers from the same hand are not convenient when users have to perform actions at the top or bottom of a large wall-sized display. Long drags all along the screen must also be avoided. In addition, techniques suitable with large displays often require control devices such a tablet or a tabletop. In the context of wall-sized displays in public area, we want to find solutions which avoid these additional devices and we would rather prefer solutions which only use direct interaction on the display.

3 In(SITE) Technique

We designed a new interface for interacting with a 3D virtual environment using multi-touch input: In(SITE), Interface for Spatial Interaction in Tactile Environments. We decided to focus on selection, translation and rotation of objects in virtual environments, so scaling is not present in this first version of the interface.

3.1 Manipulation Mode

The In(SITE) technique provides a widget, presented in Figure 2, which separates manipulation of the 6-DOF for 3D interaction. When the user touches the screen to select an object, a raycast is performed starting from his head and passing by his finger tip, which makes it possible to select the object under his finger tip according to his point of view on the 3D scene. If it reaches an object of the virtual environment, the widget appears on the screen under the user's finger.

Figure 2: *In(SITE) manipulation widget.*

This widget allows the user to fully control the 6-DOF of the selected object using multi-touch input. With his primary finger (i.e., the one used for the selection), positioned at the center of the central circle of the widget, the user controls in a direct way the two *x* and *y* translations of the object in a plane parallel to the screen (see Figure 1(*main picture*)). This method allows co-planar dragging and is called *ObjectCorrection* by Möllers et al. [2012] in the literature. It implies that if the user moves his head, the object will also move to remain under his finger. The central circle of widget is

a dead zone to avoid false-positive detections of secondary fingers while dragging, especially when the touch-detection system is not very precise (as it is the case on our wall-sized display).

The z translation is controlled in an indirect way by interacting with the area outside the widget with a secondary finger (see Figure 1(a)). This interaction design has been strongly influenced by the *Z-technique* [Martinet et al. 2010a]. The main difference is that the gesture of the secondary finger does not have to be vertical but towards the primary finger: when the user's secondary finger gets closer to his primary finger mimicking a large pinch gesture, the object gets closer to him, and when his secondary finger gets away from his primary finger mimicking a large unpinch gesture, the object gets away from him. The direction of the z translation is given by the ray starting from the head of the user and passing through his primary finger. The mapping between the finger displacement and the object displacement is linear.

The upper and lower areas of the widget are used to control the rotation of the object. The lower area allows users to control the roll of the object by doing curve gestures which follow the object rotation (see Figure 1(b)). These curve gestures are suggested to the user by the round shape of the area. The upper area allows users to manipulate the yaw and pitch of the object by doing respectively horizontal and vertical movements (see Figure 1(c)). In the upper area, the yaw and pitch rotations can be combined by doing diagonal movements. The square shape of this area aims to highlight these 2D interaction possibilities. For both rotations, the rotation axis are defined by the gravity center of the manipulated object.

Mode	Translation			Rotation		
	Tx	Ty	Tz	Rx	Ry	Rz
1d(center)	○——○					
1d(center) + 1i(out)	○——○——①					
2d(center + up)	○——○			○——○		
2d(center + down)	○——○					○

Table 1: *Description of the In(SITE) technique using the taxonomy described by Martinet et al. [2010b].*

Table 1 presents the different DOF manipulation modes offered by In(SITE) using the taxonomy introduced by Martinet et al. [2010b]. The primary finger is used to control Tx and Ty in a direct way, as the projection on the screen of the object is directly mapped to the position of the finger. The secondary finger can control Tz, Rx and Ry, or Rz. Tz is controlled in an indirect way using the external area of the widget, Rx and Ry are controlled in a direct way using the upper area of the widget, and Rz is controlled in a direct way using the lower area of the widget.

3.2 Teleportation

As In(SITE) aims to be used in large displays, we combine teleportation with the In(SITE) technique to improve its effectiveness for translations. Indeed, teleportation avoids that the user drags objects all along the screen when he needs to perform large translations.

To perform teleportation, the first step consists in selecting the object that the user wants to teleport. If he does a short touch (less than 1s) on the object, it is selected. As a feedback of the selection, the object color is changed. If he does a long touch (more than 1s), the interface enters directly in the manipulation mode of the In(SITE) technique, without selection. If the user selects an object by mistake, he can deselect it just by touching it again. In this case, the color feedback is removed and the user can select another object.

Once the user has selected an object, he can teleport it anywhere in the environment by doing a short touch on the destination location (either on the floor or on another object). This instantly moves the object above the destination and makes it fall downwards with an animation. The animation stops when the teleported object collides with another object of the environment. This animation is particularly useful for physically simulated virtual environments because the object can be piled up on other objects. Once the animation is over, the color feedback is still present for a short period of time. During this period, the user can catch the object again by touching it. If the object is caught, the interface enters in manipulation mode.

It is also possible to teleport the object without animation by doing a long touch on the destination. In this case the object is directly teleported at the location designated by the intersection between the environment and the ray starting from the user's head and passing by the finger used for the touch. Once the object has been teleported under the user's finger, the interface enters in manipulation mode. The teleportation is especially useful when the user wants to put the object under some other objects.

4 Usability Testing

We conducted two experiments to assess the usability of the In(SITE) technique combined with object teleportation in the context of an immersive virtual environment based on a large wall-sized display. The goal was to determine if the In(SITE) technique could reach the same level of performance than a standard 3D interaction technique for 6-DOF manipulation tasks of 3D objects in terms of completion time of the task and precision. We also considered precision, ease of use, frustration and fatigue of users as important evaluation criteria. In particular, we wanted to assess the four following hypotheses:

H1 Users can reach the same level of performance with In(SITE) in comparison to standard 3D interaction technique for tasks which require only translations.

H2 Users still can reach the same level of performance with In(SITE) in comparison to standard 3D interaction technique if the task requires also rotations.

H3 In(SITE) have additional advantages, in particular in terms of precision, ease of use, frustration, and fatigue.

H4 The teleportation can be beneficial to both techniques.

For the standard 3D interaction technique, we decided to use the virtual ray technique (also known as the ray-casting technique [Jacoby et al. 1994; Mine 1995]) as a baseline for 3D interaction in virtual reality (see Figure 3). Some other techniques have been developed to improve the virtual ray technique such as HOMER [Bowman and Hodges 1997] or Voodoo Dolls [Pierce et al. 1999], but they are still hard to manage for non-expert users and require additional tracking of the users' hand. Consequently, we have chosen to stick to the more standard version of the virtual ray technique since this version is the most commonly used as a mainstream interaction technique in various virtual reality systems such as CAVE. However, a virtual ray is not very efficient for pitch and yaw rotations because the manipulated object is attached to the ray and cannot easily rotate in that case. After pilot experiments, we decided to add a feature to the virtual ray technique for rotation along a vertical axis in order to be as fair as possible in our comparison between the two techniques. Users can thus easily perform yaw rotations and even pitch rotations by turning objects of 90° on the roll axis before using this additional feature (see section 4.1.1 for implementation details).

The two experiments followed the same experimental method, so the common parts are described in the following subsection. The first experiment focused on tasks which require only translations, while the second one concentrated on tasks which require rotations.

4.1 Experimental Method

4.1.1 Apparatus

The experiments were performed on a $5.90m \times 1.96m$ wall-sized display composed of 75 thin-bezel screens for a total resolution of 14400×4800 pixels. Applications ran on a server that distributes the environment to the 10 machines running the wall-sized display. Each machine has one Intel Xeon CPU at 3.7 GHz and a Nvidia Quadro K5000 graphic card. The virtual environment was designed using the game engine *Unity*[2].

While this wall-sized display offers a very large surface with multi-touch capabilities, it does not support stereoscopy. Even if the ideal case would have been to have stereoscopic vision to support immersion in the virtual environment, the device still provide a interesting level of immersion because it supports other cues of immersion: a wide field of view, a complete visual immersion when users are close to the screens and a good resolution. The user's head was also tracked in front of the screen by a *Vicon*[3] system. Displayed images were deformed according to the user head position to match the viewing frustum defined by the user's head and the four corners of the wall-sized displays. Consequently, motion parallax was respected and participants had a sense of depth when they moved.

For the In(SITE) technique, the touch was detected by a *PQLabs*[4] infrared frame surrounding the wall-sized display. This frame is composed of infrared emitters and sensors which shoot and capture infrared light along the screen on the two directions (up-down and right-left). It can thus determine the 2D position of the fingers when they intersect with the infrared light. The main drawback of this system is that fingers can be detected slightly before they touch the screen (around $0.5cm$ from the screen). It can induce false-positive touch detections when fingers or parts of the hand are close to the screen. This system has an accuracy of $1cm$ all over the screen surface and can detect up to 32 contact points.

For the virtual ray technique, the participants held a wireless mouse attached to a Vicon target which was the starting point of the virtual ray (see Figure 3). The Vicon tracking system has an accuracy of $1mm$, but can suffer from jitter, especially for rotations. To reduce the jitter of the virtual ray, we used a *1€ filter* [Casiez et al. 2012] with the following parameters: for translations, min cut-off = 5 and beta = 0.8; for rotations, min cut-off = 0.5 and beta = 0.5. The virtual ray technique was designed as similar as possible to the In(SITE) technique. The first difference was that participants did not select objects by touching them, but by left-clicking while the virtual ray was pointing at them. The second difference was that participants did not use the In(SITE) widget, but had objects directly attached to the ray when they manipulated them. To perform rotations along the vertical axis, participants used the scroll wheel of the mouse. The right button of the mouse was used to fix the rotation of the selected object, and enabled participants to move the virtual ray without modifying the object rotation.

4.1.2 General Task

We asked participants to perform a docking task using the two interaction techniques. We set up a simple environment that enabled participants to perform the task without any navigation actions: a floor and a 1-meter high, 6-meter wide and 3-meter deep table, positioned just behind the screen. Participants had to put an object present at the left of the table in a target placed on the right of the table as fast as possible. The target, which was slightly bigger
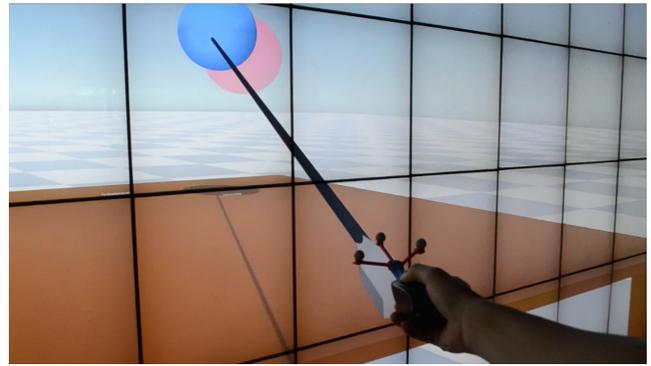
---
[2]http://unity3d.com/
[3]http://www.vicon.com/
[4]http://www.multitouch.com/

Figure 3: *For the virtual ray technique, the ray is co-located with the wireless mouse held by the user.*

than the moved object (set after pilot experiments to 10% for experiment 1 and 20% for experiment 2), was red and turned green when the object was correctly placed inside. In order to validate the completion of the task, the object had to stay in the target for one second, whether it was released or not by the participant. A directional light was placed in the scene, casting shadows vertically on the table and the floor to help participants to perceive the depth in the scene. Physics, including gravity, were disabled during the experiment to avoid bias due to the physics engine.

4.1.3 Participants

16 adults (5 females and 11 males) with a mean age of 24.4 (SD 5.0) participated in both experiments. There were all right-handed. 13 of them use touch devices every day. They had variable experience with 3D visualization systems. They all performed the second experiment after the first one. Participants were not remunerated for their participation.

4.1.4 Data Collection

For each trial, we collected the task completion time. The completion time measurement started when the participant selects the manipulated object for the first time and stopped when the task was accomplished (i.e., when the object stayed in the target for one second). We also collected the number of overshoots. An overshoot was defined by the fact that the manipulated object reached the targeted position (which means that the target turns green), but could not stay within the target during the mandatory 1 second. This usually happened when participants moved the object too fast and could not adjust precisely enough its final position. After the experiments, participants filled out a subjective questionnaire.

4.2 Experiment 1: Translations

The first experiment compared the two interaction techniques and assessed the benefits of the teleportation for 3D manipulation tasks which required only translations. Consequently, we set the TECHNIQUE and the TELEPORTATION as two primary factors. Table 2 describes the TECHNIQUE×TELEPORTATION combinations.

	3D Virtual Ray	In(SITE)
No teleportation	*VRay, NoTele*	*Touch, NoTele*
Teleportation	*VRay+Tele*	*Touch+Tele*

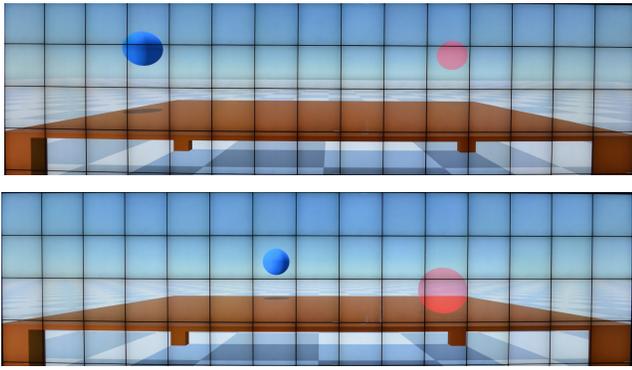**Table 2:** TECHNIQUE×TELEPORTATION *combinations in exp1.*

Figure 4: *Relative positions of the manipulated sphere and its associated target in Floor condition (bottom) and Midair condition (top) for experiment 1.*

For teleportation, both techniques used the process described in section 3.2. The touch and long touch for In(SITE) were replaced by a click and a long click on the mouse in the case of the virtual ray technique. When the participants used teleportation, they could make the object fall down from the top in the direction of the indicated destination. The object stopped when it reached the table in the particular case of this experiment. The vertical position of the target could thus have an impact on the performance. Indeed, the object stopped at a position close to the target when the target was on the table, while it still stopped on the table even if the target was in the air. Consequently, we decided to set the VERTICAL POSITION of the target as a secondary factor to evaluate the influence of the height of the target on the performance.

4.2.1 Task

The manipulation task followed the general task described in section 4.1.2. The manipulated objects were spheres to insure that participants did not need to rotate the objects for putting them in the targets (see Figure 4). We picked 4 random positions for each object, on the left of the table for the manipulated object and on the right for the target, making sure that the objects were at least 50 cm away from each other.

4.2.2 Experimental Design & Procedure

The experiment had a [2 × 2 × 2] within-subject design with the following factors:

- TECHNIQUE, with the two treatments: *VRay* for the virtual ray technique and *Touch* for the In(SITE) technique.

- TELEPORTATION, with the two treatments: *NoTele* when the technique was not combined with teleportation and *+Tele* when the technique was combined with teleportation.

- VERTICAL POSITION, the two levels: *Floor* condition for which the position of the target was on the table and *Midair* condition for which the position of the target was above the table at a random altitude of at least 50 cm (see Figure 4).

Trials were grouped by TECHNIQUE and TELEPORTATION. The order of TECHNIQUE×TELEPORTATION was counterbalanced across participants using a balanced latin square. For each combination of TECHNIQUE×TELEPORTATION, the session began by 4 training trials, during which we explained to the participant how to use the TECHNIQUE with or without the TELEPORTATION. After the training, the participant performed 8 trials: 4 trials for the *Midair* condition using the 4 different object/target positions and 4 trials for the *Floor* condition using the same 4 object/target positions, but

with a zero altitude for the target (target laying on the table). The order of the 8 trials was randomly chosen. Consequently, each one of the 16 participant performed 2 TECHNIQUEs × 2 TELEPORTATIONs × 2 VERTICAL POSITIONs × 4 trials = 32 trials (bringing the total to 512 trials for the whole experiment). The participants were authorized to take a break whenever they wanted between trials, but we encouraged them to take a break between each block of the experiment corresponding to a particular combination TECHNIQUE×TELEPORTATION. Sessions lasted from 25 to 40 minutes depending on the participants.

4.2.3 Results

Task Completion Time An analysis of outliers detected two extreme outliers. The global mean for *completion time* was 14.34*s* (SD: 11.3) while the two trials corresponding to the two outliers lasted more than 130*s*. Since we had recorded during the experiment that two participants faced critical problems during these two trials, we chose to remove these two extreme outliers. One participant faced technique issues, while the other did not understand that the depth was wrong and kept trying to adjust the position during a very long time. Figure 5 illustrates the results for each TECHNIQUE×TELEPORTATION combination and VERTICAL POSITION with regard to task *completion time*.
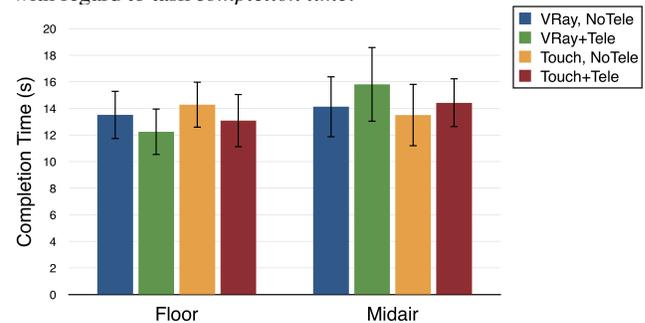

Figure 5: *Mean completion time by* TECHNIQUE×TELEPORTATION *and* VERTICAL POSITION *in experiment 1. Error bars represent 95% confidence intervals.*

A repeated measures ANOVA[5] on *completion time* with the model TECHNIQUE×TELEPORTATION×VERTICAL POSITION revealed no significant effect of TECHNIQUE, TELEPORTATION, and VERTICAL POSITION, but it showed a significant interaction effect of TELEPORTATION×VERTICAL POSITION ($F(1, 14.3) = 8.3$, $p = 0.01$). Pairwise comparisons revealed that techniques combined with teleportation (*+Tele*) were significantly faster in *Floor* condition (avg. 12.66*s*) than in *Midair* condition (avg. 15.12*s* , $p < 0.001$). This showed that teleportation was more efficient when the target was laying on another horizontal object (*Floor* condition). In addition, the overall mean values of *completion time* for the two TECHNIQUEs were very close to each other (avg. *VRay*: 13.92*s* and *Touch*: 13.83*s*), which suggests that there is no practical difference between *VRay* and *Touch* with respect to the *completion time*.

Overshoots For the *overshoots* analysis, we also excluded the two extreme outliers detected during the *completion time* analysis in order to be consistent. Figure 6 shows the mean number of *overshoots* for each TECHNIQUE×TELEPORTATION combination and VERTICAL POSITION. A repeated measures ANOVA on *overshoots* with the model TECHNIQUE×TELEPORTATION×VERTICAL POSITION revealed a significant effect of TECH-

---

[5]All analyses except the ART procedure were performed with the SAS JMP statistical platform. The ART procedure [Wobbrock et al. 2011] was performed with R.
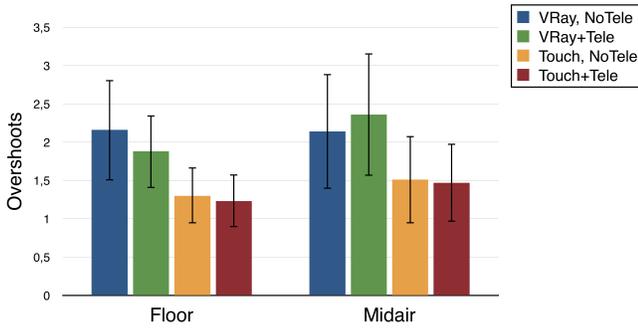
Figure 6: *Mean numbers of overshoots by* TECHNIQUE×TELE-PORTATION *and* VERTICAL POSITION *in experiment 1. Error bars represent 95% confidence intervals.*

NIQUE ($F(1, 14.96) = 3.63$, $p = 0.002$), but no significant effect of TELEPORTATION and VERTICAL POSITION, as well as no significant interaction effects. The *Touch* technique (avg. 1.38) led to significantly less overshoots than the *VRay* technique (avg. 2.13).

Subjective Questionnaire After the experiment, participants had to rate on a 5-point Likert scale if each TECHNIQUE×TELE-PORTATION combination was *mentally demanding*, *physically demanding*, *frustrating*, *easy* to use, *precise* (1:not at all/5:very) and they had also to give an *overall evaluation* (1:bad/5:good). Figure 7 illustrates the results of the subjective questionnaire. To analyze the results of the subjective questionnaire, we used the aligned rank transform (ART) procedure proposed by Wobbrock et al. [2011] with the model TECHNIQUE×TELEPORTATION. It showed a significant effect of TECHNIQUE on the *easy* ($F(1, 45) = 19.21$, $p < 0.001$) and *precise* ($F(1, 45) = 30.49$, $p < 0.001$) criteria. *Touch* was perceived easier to use (avg. 4.31 vs. 3.61) and more precise (avg. 4.47 vs. 3.5) than *VRay*. The analysis also showed a significant effect of TELEPORTATION on the *mentally demanding* ($F(1, 45) = 5.36$, $p = 0.03$), *physically demanding* ($F(1, 45) = 7.64$, $p < 0.001$) and *easy* ($F(1, 45) = 18.41$, $p < 0.001$) criteria, as well as on the *overall evaluation* ($F(1, 45) = 14.57$, $p < 0.001$). *+Tele* was perceived less mentally demanding (avg. 1.91 vs. 2.31), less physically demanding (avg. 2.19 vs. 2.75) and easier to use (avg. 4.28 vs. 3.69) than *NoTele*. *+Tele* was also preferred to *NoTele* (avg. 4.21 vs. 3.47) according to the overall evaluation. Finally, the analysis did not reveal any significant effects of TECHNIQUE and TELEPORTATION on the *frustrating* criterion, as well as any significant interaction effects of TECHNIQUE×TELEPORTATION on all the criteria.

4.3 Experiment 2: Rotations

The second experiment compared the In(SITE) technique with the virtual ray technique for 3D manipulation tasks which required rotations. In order to minimize the influence of translation on task performance, the object and the target were placed close to each other with different orientations. Since the translations to perform were really small and the teleportation did not have an effect on rotations, it seemed reasonable to consider that teleportation did not impact the task performance. Consequently, only the *VRay* and *Touch* TECHNIQUEs both without teleportation were evaluated in this experiment to shorten the global duration.

4.3.1 Task

The manipulation task followed the general task described in section 4.1.2. The manipulated objects were wedges (triangular prisms with one right angle) to insure that there was only one correct ro-
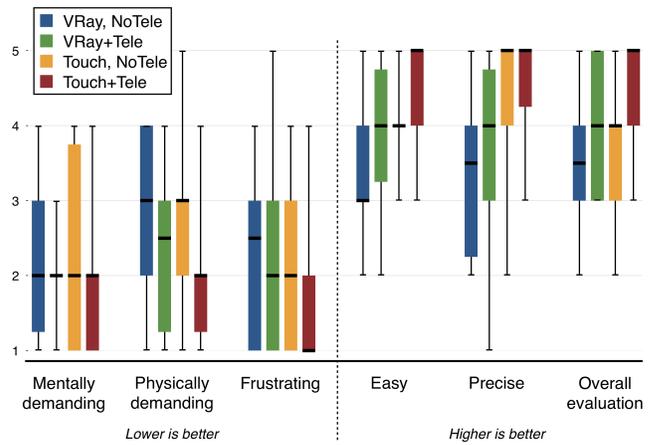

Figure 7: *Boxplots for the answers to the subjective questionnaire of experiment 1 using a 5-point Likert scale. Each boxplot is delimited by the quartile (25% quantile and 75% quantile) of the distribution of the effect over the participants. The median is also represented for each* TECHNIQUE×TELEPORTATION *combination. Error bars represent data minimum and maximum.*
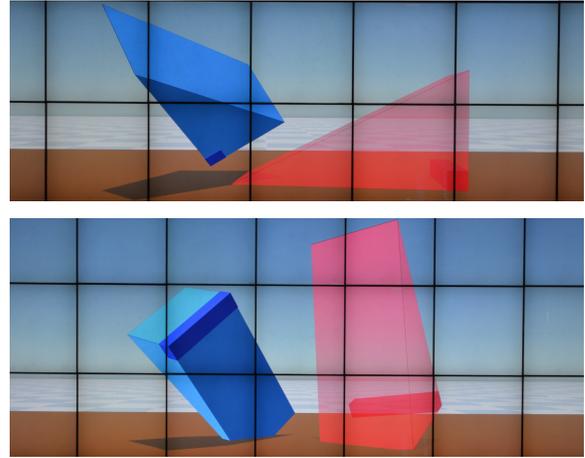

Figure 8: *Relative positions of the manipulated wedge and its associated target in Floor condition (top) and Midair condition (bottom) for experiment 2.*

tation to fit the object in the target. Since it was not always easy to determinate the angle values in the 3D virtual environment, the right angle of the wedge was marked on both the object and the target to help participants identifying the rotation to perform (see Figure 8). The manipulated object and the target were placed 1*m* away from each other and the manipulated object was located on the left of the target. We picked random values for the orientations of both the object and the target. The rotation was considered as correct if the dot product between the manipulated object and the target's up, right, and forward were bigger than or equal to 0.98 (set after pilot experiments).

4.3.2 Experimental Design & Procedure

The experiment had a [2 × 2] within-subject design with the following factors:

- TECHNIQUE, with the two treatments: *VRay* and *Touch* both without teleportation.

- VERTICAL POSITION, the two levels: *Floor* and *Midair* conditions. For the *Midair* positions, the target height was set at 50 cm above the table. For the *Floor* positions, the target position was defined with one side of the target laying on the virtual table (see Figure 8).

Trials were grouped by TECHNIQUE and the order of the TECHNIQUEs were counterbalanced across participants. For each TECHNIQUE, the session began by 4 training trials. After the training, the participant performed 8 trials: 4 trials for the *Floor* condition and 4 trials for the *Midair* condition. The order of the 8 trials was randomly chosen. Consequently, each one of the 16 participant performed 2 TECHNIQUEs × 2 VERTICAL POSITIONs × 4 trials = 16 trials (bringing the total to 256 trials for the whole experiment). The participants were authorized to take a break whenever they wanted between trials, but we encouraged them to take a break between the two blocks of the two TECHNIQUEs. Sessions lasted from 15 to 20 minutes depending on the participants.
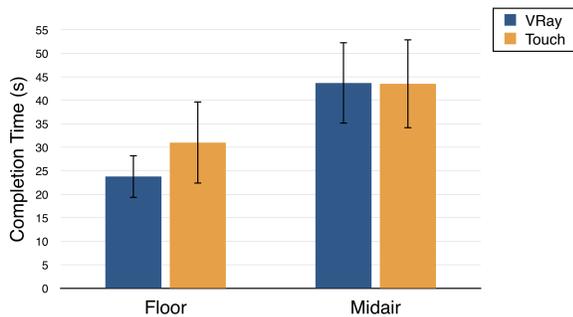

Figure 9: *Mean completion time by* TECHNIQUE *and* VERTICAL POSITION *in experiment 2. Error bars represent 95% confidence intervals.*

4.3.3 Results

Task Completion Time Figure 9 illustrates the results for each TECHNIQUE and VERTICAL POSITION with respect to task *completion time*. A repeated measures ANOVA on *completion time* with the model TECHNIQUE×VERTICAL POSITION showed a significant effect of VERTICAL POSITION ($F(1, 15) = 25.55$, $p < 0.001$). Participants performed the task significantly faster in *Floor* condition (avg. 27.4*s*) than in *Midair* condition (avg. 43.61*s*). However, the analysis did not reveal any significant effects of TECHNIQUE or interaction effect of TECHNIQUE×VERTICAL POSITION.

Overshoots Figure 10 shows the mean number of *overshoots* for each TECHNIQUE and VERTICAL POSITION. A repeated measures ANOVA with the model TECHNIQUE×VERTICAL POSITION showed a significant effect of TECHNIQUE ($F(1, 15) = 20.25$, $p < 0.001$) on *overshoots*. *Touch* (avg. 0.45) led to significantly less overshoots than *VRay* (avg. 0.84). However, the analysis revealed neither a significant effect of VERTICAL POSITION nor a significant interaction effect of TECHNIQUE×VERTICAL POSITION.

Subjective Questionnaire After the experiment, participants filled out a similar questionnaire to the one of experiment 1, but only for the two TECHNIQUE conditions (*Touch* and *VRay*). Pairwise Wilcoxon rank sum tests with Bonferroni corrections among TECHNIQUEs did not reveal any significant differences for all the criteria, and for the *overall evaluation*.
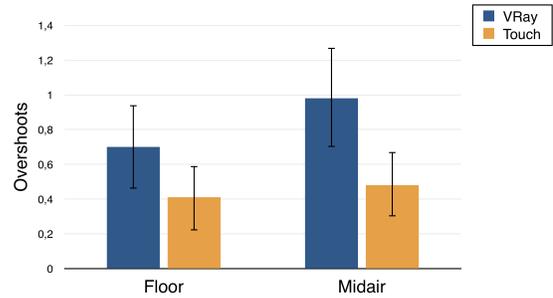

Figure 10: *Mean numbers of overshoots by* TECHNIQUE *and* VERTICAL POSITION *in exp2. Error bars represent 95% confidence intervals.*

5 Discussion

On one hand, the two experiments did not reveal any significant effects of TECHNIQUE on completion time. This result, in addition to the fact that the overall mean values of completion time were very similar for the two techniques, suggests that users can reach the same level of performance with the In(SITE) technique in comparison with a standard 3D interaction technique even if the manipulation tasks require both translations and rotations. It suggests a validation of **H1** and **H2**.

On the other hand, the results of both experiments revealed that the In(SITE) technique led to a significant smaller number of overshoots in comparison to the technique based on a virtual ray. It shows that the In(SITE) technique was more precise for fine adjustments of 3D object position in the immersive context of large vertical displays. This result is consistent with the fact that participants perceived the In(SITE) technique as more precise than the virtual ray technique for manipulation tasks which required only translations. Concerning **H3**, we can note that the In(SITE) technique was more precise for fine adjustments of 3D object position, which was confirmed by the subjective questionnaire only for translations. For translations, the In(SITE) technique was also perceived easier to use by the participants.

Moreover, the results of the first experiments revealed that teleportation significantly improved the performance of the two techniques in terms of completion time for the *Floor* condition. This result seems coherent since users can select a destination for teleportation close to the targeted position in this condition. In addition, the subjective questionnaire of the first experiment showed that teleportation significantly reduced the mental and physical loads. For the mental load, we cannot formulate a strong claim, but it might be explained by the fact that the participants only had to focus on the destination with teleportation and not on the full path to reach the target. For the physical load, it was probably due to the gorilla-arm effect [Hincapié-Ramos et al. 2014] which could affect participants if they hold the 3D virtual ray during a long time or drag an object all along the screen with In(SITE). Teleportation avoids to always raise the arm during interaction and reduces the arm fatigue. This is also confirmed by the fact that participants preferred and found easier to use the two techniques when they were combined with teleportation. These results show that **H4** is confirmed and that the teleportation has a real benefit on large wall-sized displays.

Finally, the difference on *completion time* between the *Floor* and *Midair* conditions in the second experiment can be explained by the fact that the targeted positions in *Floor* condition were defined with one side of the object laying on the virtual table. Consequently, the performed rotations were less complicated and the participants achieved more quickly the manipulation task in *Floor* condition.

6 Conclusion and Future Work

This paper aims to assess the usability of the In(SITE) technique combined with teleportation for 3D manipulation on multi-touch wall-sized displays. The In(SITE) technique is a bimanual touch-based technique for 3D interaction adapted to the situation where users are immersed in a 3D virtual environment and move in front of a large vertical screen. The usability testing shows that the In(SITE) technique can reach the same level of performance than a standard 3D interaction technique with respect to completion time. In addition, the In(SITE) technique improves precision for fine adjustments of 3D object position. The results also show that techniques combined with teleportation were found easier to use, less tiring and were globally preferred by participants for translation tasks.

This paper is a first study of combining a bimanual touch-based interaction with teleportation on large wall-sized display. Since we demonstrated that the In(SITE) technique can be beneficial in such context, we want now to improve it by adding scaling functionalities and to explore different variations of the proposed widget to question the design choices we have done. For example, it will be interesting to study if the choice of the rotation axis of the manipulated object (object gravity center or selection point of the user) or if the choice of the transfer function (which is currently linear) used to control the translation along the *z*-axis as studied by Casiez et al. [2008] have some impacts on performance. In addition, we want to study more in depth the different teleportation strategies such as the animated teleportation (objects falling down when a physical simulation is used) or the direct teleportation (under the user's finger).

The wall-sized display used during the experiments does not support stereoscopy, but we think that the In(SITE) technique could be extended to stereoscopic vision using the guidelines defined by Valkov et al. [2011]. Consequently, it would be interesting to study the In(SITE) technique with another system which support stereoscopy and to compare the results. In addition, we want to adapt and study the In(SITE) technique for systems which do not use head-tracking and viewing frustum deformation. These systems are especially relevant to enable multiple users to collaborate together in front of a same wall-sized display. We think that In(SITE) could be easily adapted to this context by using the center of the virtual camera instead of the user's head as the starting of the raycast when performing object selection and manipulation.

Finally, touch-based 3D interaction techniques could be an issue for remote collaboration in a shared virtual environment. Indeed, when a user interacts with 3D objects through his own multi-touch device, it is impossible for the remote users to understand which objects he is manipulating. It is an issue for the communication between the remote users. We need to provide a feedback which links the user's 2D touch inputs to the manipulated 3D objects. A similar solution to the 2D pointer / 3D ray proposed by Duval and Fleury [2009] could be an interesting way to implement this feedback.

Acknowledgements

References

BENKO, H., AND FEINER, S. 2007. Balloon Selection: A Multi-Finger Technique for Accurate Low-Fatigue 3D Selection. In *2007 IEEE Symposium on 3D User Interfaces*, IEEE Computer Society, 3DUI '07.

BOWMAN, D. A., AND HODGES, L. F. 1997. An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. In *Proc. 1997 Symposium on Interactive 3D Graphics*, ACM, I3D '97, 35–ff.

CASIEZ, G., VOGEL, D., BALAKRISHNAN, R., AND COCKBURN, A. 2008. The Impact of Control-Display Gain on User Performance in Pointing Tasks. *Human-Computer Interaction 23*, 3, 215–250.

CASIEZ, G., ROUSSEL, N., AND VOGEL, D. 2012. 1€ filter: A simple speed-based low-pass filter for noisy input in interactive systems. In *Proc. SIGCHI Conference on Human Factors in Computing Systems*, ACM, CHI '12, 2527–2530.

COFFEY, D., MALBRAATEN, N., LE, T. B., BORAZJANI, I., SOTIROPOULOS, F., ERDMAN, A. G., AND KEEFE, D. F. 2012. Interactive Slice WIM: Navigating and Interrogating Volume Data Sets Using a Multisurface, Multitouch VR Interface. *IEEE Transactions on Visualization and Computer Graphics 18*, 10, 1614–1626.

COHÉ, A., AND HACHET, M. 2012. Understanding User Gestures for Manipulating 3D Objects from Touchscreen Inputs. In *Proc. Graphics Interface 2012*, Canadian Information Processing Society, GI '12, 157–164.

COHÉ, A., DÈCLE, F., AND HACHET, M. 2011. tBox: A 3D Transformation Widget Designed for Touch-screens. In *Proc. SIGCHI Conference on Human Factors in Computing Systems*, ACM, CHI '11, 3005–3008.

DUVAL, T., AND FLEURY, C. 2009. An Asymmetric 2D Pointer/3D Ray for 3D Interaction Within Collaborative Virtual Environments. In *Proc. 14th International Conference on 3D Web Technology*, ACM, Web3D '09, 33–41.

GILLIOT, J., CASIEZ, G., AND ROUSSEL, N. 2014. Direct and Indirect Multi-touch Interaction on a Wall Display. In *Proc. 26th Conference on L'Interaction Homme-Machine*, ACM, IHM '14, 147–152.

HANCOCK, M., CARPENDALE, S., AND COCKBURN, A. 2007. Shallow-depth 3D Interaction: Design and Evaluation of One-, Two- and Three-touch Techniques. In *Proc. SIGCHI Conference on Human Factors in Computing Systems*, ACM, CHI '07, 1147–1156.

HANCOCK, M., TEN CATE, T., AND CARPENDALE, S. 2009. Sticky Tools: Full 6DOF Force-based Interaction for Multi-touch Tables. In *Proc. ACM International Conference on Interactive Tabletops and Surfaces*, ACM, ITS '09, 133–140.

HINCAPIÉ-RAMOS, J. D., GUO, X., MOGHADASIAN, P., AND IRANI, P. 2014. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-air Interactions. In *Proc. SIGCHI Conference on Human Factors in Computing Systems*, ACM, CHI '14, 1063–1072.

JACKSON, B., SCHROEDER, D., AND KEEFE, D. F. 2012. Nailing Down Multi-Touch: Anchored Above the Surface Interaction for 3D Modeling and Navigation. In *Proc. Graphics Interface 2012*, Canadian Information Processing Society, GI '12, 181–184.

JACOBY, R. H., FERNEAU, M., AND HUMPHRIES, J. 1994. Gestural interaction in a virtual environment. In *Proc. of Stereoscopic Displays and Virtual Reality Systems (SPIE)*, 355–364.

JANKOWSKI, J., AND HACHET, M. 2015. Advances in Interaction with 3D Environments. *Computer Graphics Forum 34*, 1, 152–190.

LOPEZ, D., OEHLBERG, L., DOGER, C., AND ISENBERG, T. 2016. Towards An Understanding of Mobile Touch Navigation in a Stereoscopic Viewing Environment for 3D Data Exploration. *IEEE Transactions on Visualization and Computer Graphics 22*, 5, 1616–1629.

MARTINET, A., CASIEZ, G., AND GRISONI, L. 2010. The Design and Evaluation of 3D Positioning Techniques for Multi-touch Displays. In *Proc. 2010 IEEE Symposium on 3D User Interfaces*, IEEE Computer Society, 3DUI '10, 115–118.

MARTINET, A., CASIEZ, G., AND GRISONI, L. 2010. The Effect of DOF Separation in 3D Manipulation Tasks with Multi-touch Displays. In *Proc. 17th ACM Symposium on Virtual Reality Software and Technology*, ACM, VRST '10, 111–118.

MINE, M. R. 1995. Virtual Environment Interaction Techniques. Tech. rep., University of North Carolina, Department of CS.

MÖLLERS, M., ZIMMER, P., AND BORCHERS, J. 2012. Direct Manipulation and the Third Dimension: Co-planar Dragging on 3D Displays. In *Proc. 2012 ACM International Conference on Interactive Tabletops and Surfaces*, ACM, ITS '12, 11–20.

PIERCE, J. S., STEARNS, B. C., AND PAUSCH, R. 1999. Voodoo Dolls: Seamless Interaction at Multiple Scales in Virtual Environments. In *Proc. 1999 Symposium on Interactive 3D Graphics*, ACM, I3D '99, 141–145.

REISMAN, J. L., DAVIDSON, P. L., AND HAN, J. Y. 2009. A Screen-space Formulation for 2D and 3D Direct Manipulation. In *Proc. 22Nd Annual ACM Symposium on User Interface Software and Technology*, ACM, UIST '09, 69–78.

STROTHOFF, S., VALKOV, D., AND HINRICHS, K. 2011. Triangle Cursor: Interactions with Objects Above the Tabletop. In *Proc. ACM International Conference on Interactive Tabletops and Surfaces*, ACM, ITS '11, 111–119.

VALKOV, D., STEINICKE, F., BRUDER, G., AND HINRICHS, K. 2011. 2D Touching of 3D Stereoscopic Objects. In *Proc. SIGCHI Conference on Human Factors in Computing Systems*, ACM, CHI '11, 1353–1362.

WOBBROCK, J. O., FINDLATER, L., GERGLE, D., AND HIGGINS, J. J. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proc. SIGCHI Conference on Human Factors in Computing Systems*, ACM, CHI '11, 143–146.

YU, L., SVETACHOV, P., ISENBERG, P., EVERTS, M. H., AND ISENBERG, T. 2010. FI3D: Direct-Touch Interaction for the Exploration of 3D Scientific Visualization Spaces. *IEEE Transactions on Visualization and Computer Graphics 16*, 6, 1613–1622.