*International Journal of Innovative Computing, Information and Control*
Volume **9**, Number **12**, December **2013**
ICIC International ©2013 ISSN 1349-4198
pp. 4737–4752

# AN INVESTIGATION OF SUITABLE INTERACTIONS FOR 3D MANIPULATION OF DISTANT OBJECTS THROUGH A MOBILE DEVICE

Hai-Ning Liang[1], Cary Williams[2], Myron Semegen[3]
Wolfgang Stuerzlinger[4] and Pourang Irani[2]

[1]Department of Computer Science and Software Engineering
Xi'an Jiaotong-Liverpool University
No. 111, Ren'ai Road, Suzhou Dushu Lake Higher Education Town, Suzhou 215123, P. R. China
haining.liang@xjtlu.edu.cn

[2]Department of Computer Science
University of Manitoba
Winnipeg, Manitoba R3T 2N2, Canada
umwill22@cc.umanitoba.ca; irani@cs.umanitoba.ca

[3]Virtual Reality Centre
Industrial Technology Centre
Winnipeg, MB R3T 6C2, Canada
msemegen@itc.mb.ca

[4]Department of Computer Science and Engineering
York University
Toronto, Ontario M3J 1P3, Canada
wolfgang@cse.yorku.ca

Received August 2012; revised January 2013

Abstract. *In this paper, we present our research linking two types of technologies that are gaining importance: mobile devices and large displays. Mobile devices now come integrated with a variety of sophisticated sensors, including multi-touch displays, gyroscopes and accelerometers, which allow them to capture rich varieties of gestures and make them ideal to be used as input devices to interact with an external display. In this research we explore what types of sensors are easy and intuitive to use to support people's exploration of 3D objects shown in large displays located at a distance. To conduct our research we have developed a device with two multi-touch displays, one on its front and the other in the back, and motion sensors. We then conduct an exploratory study in which we ask participants to provide interactions that they believe are easy and natural to perform operations with 3D objects. We have aggregated the results and found some common patterns about users' preferred input sensors and common interactions. In this paper we report the findings of the study and describe a new interface whose design is informed by the results.*

**Keywords:** Human-computer interaction, Gesture-based interaction, Input devices, Interaction techniques, Multi-display environments, Mobile devices, 3D visualizations

1. **Introduction.** This research aims to bring together two fast-growing types of technologies: (1) mobile devices with multi-touch screens and motion sensors, and (2) large displays. Mobile devices are rapidly becoming the most important, pervasive personal gadget that people have, while large displays are now commonly found in an increasing number of settings, both public and private. Mobile devices with their sophisticated sensing capabilities are becoming a '*ubiquitous input device*' [1] and quite suitable to allow users to interact with large displays that are located at a distance [2-5]. This research

has two main goals. First, given that a device has multiple types of input sensors, what sensors will users find natural and easy to use to interact with 3D objects shown in distant displays? Second, will users have agreement in terms of how to carry out specific interactions with these objects?

There is a growing emphasis to use mobile devices to interact with information shown in a distant display [6-9], with all three major video game companies – i.e., Microsoft, Sony and Nintendo – investigating and starting to develop controllers similar to mobile devices to support enhanced gameplay experiences. There are many potential benefits that can be derived from the union of mobile devices with large displays, including: (1) the additional display on the device can enrich the user experience in a variety of tasks, such as video gameplay, browsing of pictures, web and map navigation, and exploration of 3D virtual environments [4,5,8,10,11]; (2) mobile devices can now support a variety of input modalities, the combination of which is not available in other input devices [1,3,7,12,13]; (3) multiple users can interact with the same distant display with their own devices thus supporting both collaborative public exploration and private browsing [6,8,14]; mobile devices enable interacting with content that is difficult or impossible to reach, for instance in cases where the displays are placed outside of a window or high up in a building [15].

We are at an early stage of leveraging the full capabilities of the mobile devices as input instruments. There are many usability issues that still need to be addressed. There is, for instance, an implicit assumption that the more sensing capabilities a device has, the better it will be received. However, given users' cognitive limitations, ergonomic properties of their hands, and their perceptual constraints, it is not certain whether they can make use of all the sensing features. In addition, it is not clear whether users have any preference as to which input feature they like to use most.

This research is a direct attempt to advance our understanding of how to design mobile devices with a variety of sensing capabilities and interactions that are suitable for manipulating objects located at a distant display through these devices. We are primarily concerned with the manipulation of 3D objects by using finger swipe gestures and motion interactions (see Figure 1). To conduct this research we use a prototype made up of two identical tablet devices placed back-to-back. The devices feature a 10.1 inch multi-touch display, an accelerometer, and a gyroscope, and can therefore detect and capture multiple finger swipes on both sides and movements along all three axes of rotation.

The results of this research are relevant to designers of interactive applications that (1) run on multiple displays of varied sizes, (2) require the manipulation of 3D objects, and (3) need users to be standing at a distance from a display. As shown by gaming

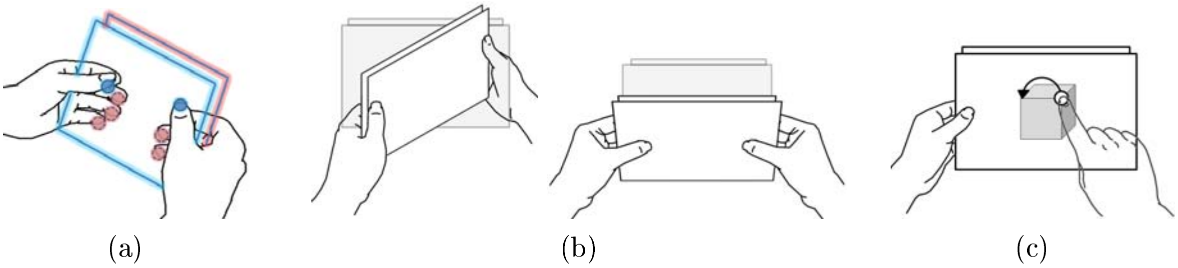

(a) (b) (c)

FIGURE 1. (a) Our prototype made up of two tablet devices placed back-to-back which enables multi-touch input from both sides; (b) two types of motion-based interactions: Yaw (left) and Roll (right); (c) an example of finger swipe interaction.

companies, there is increasing interest in these types of applications, and this current work provides the foundation on top of which further research can take place.

The paper is organized as follows. The next section presents the background and related work. The section after describes our experiment and then discusses the results. The following section outlines a technique we developed based on the findings from the first study and compares this technique with a common technique. Finally, the last summarizes the paper and concludes with the description of some applications of this research.

2. **Background and Related Work.** Display technologies have advanced rapidly and this has led to widespread adoption of high resolution displays that are both large and inexpensive. One area that has benefited from this development is 3D digital images or visualizations because they can now be shown with greater detail and realism. Displaying a static image is not that useful for deep analysis but when made interactive it increases their communicative power and aids in their exploration [16]. Interacting with 3D images on large displays is not trivial, but presents numerous design challenges for interaction designers [11,17-20,40]. One of the main challenges is to find a suitable input device and interaction techniques that allow for natural and expressive interactions [9,17,40]. Researchers have recently begun to investigate the use of mobile devices to interact with 3D environments and objects located on distant displays (e.g., see [4,11]), and this current work aims to make a contribution in this line of research.

Interacting with 3D objects requires an input device with multiple degrees of freedom to allow users to express a variety of actions. Current mobile devices (such as smartphones and tablets) come with a display that accepts multi-touch input along with multiple motion sensors, such as accelerometers and gyroscopes, that allows them to capture and distinguish with high precision movements along all three axis of rotation. Researchers in human-computer interaction define finger swipes performed on the device's touchscreen as *surface gestures*, while motions obtained by rotating the device along any axis are called *motion gestures* [13]. If surface and motion gestures are both provided, it is possible to support a wide variety of interactions.

2.1. **Surface and motion gestures.** *Surface gestures* are based on two-dimensional finger movements, for example sliding the finger on the touchscreen right to unlock the iPhone. This type of gestures has been studied in multi-touch tabletop surfaces (e.g., [21-23]). Results from Wobbrock et al.'s study, for example, suggest that on tabletops, users prefer to use one hand, and often rely on the index finger for interaction [23]. In mobile devices, surface gestures are now well accepted as the main interaction style. Bragdon et al. [24] have found that, in the presence of distractors, gestures offer better performance and also reduced attentional load. Surface techniques, such as Gesture Avatar [25] and Gesture Search [26], show that they support fast, easy target selection and data access.

*Motion gestures* are based on changing the position or orientation of the device, for example by tilting or shaking it. These gestures have been studied for different tasks such as text entry, user identity validation, and content navigation [22,27,28] with tilt being the most common gesture. Recently, Ruiz et al. [13] have investigated in detail what kinds of motion gestures people prefer to perform and found people prefer gestures based on small and unidirectional motions (e.g., flipping left or right) avoiding complex compound movements.

The prior research suggests that people can use both motion and surface gestures. However, little is known about how they would respond to using the two types together and for interacting with large displays. In addition, this prior research is based on one-handed usage of mobile devices as opposed to two-handed or bimanual manipulation, which can be beneficial for many compound interactions and those of higher complexity.

2.2. **Dual-surface interaction.** One main aim of this research is to explore how people make use of all the interaction possibilities given by mobile devices. The touchscreen is located on the front of devices, and is mainly through it that interactions take place. More recently, the back side has been explored as an additional input space [29-33]. This type of interaction is called back-of-device interaction and can be seen in early research systems. RearType [30], for example, allows text-entry using a keypad on the back. HybridTouch lets users issue navigation and browsing commands using surface gestures through a backside touchpad [34]. This work shows that users can perform surfaces gestures through the backside under constrained conditions and for selected activities.

Bimanual interaction is when people use both hands to manipulate the device. For example, Silfverberg et al.'s prototype [31] has two trackpads on the back, and this configuration lets user do zooming actions with one hand and panning actions with the other. Little has been done to assess whether people like to and can use both hands to manipulate mobile devices to interact with distant objects through gestures.

Researchers have experimented using both sides of a device to enable input, and this is called dual-surface input [33]. Yang et al. [33] show that one-handed operations can be enhanced with synchronized interactions using both the back and front of mobile devices, but only for simple target selection and steering tasks. Similarly, Wigdor et al. [32] introduce a device that can require people to use both hands to interact with the back and front sides.

We suspect that bimanual and dual-surface input is suitable for simple tasks but may be cognitively demanding for more complex tasks without extensive training. This is one main reason for this research since we want to assess with our dual-surface bimanual device whether people will find it easy to perform activities that are of higher complexity – in our case 3D manipulations. Given the frequent use of 3D representations in many domains (e.g., video games, biology, architecture), it is timely that we understand how to design tools and devices capable of supporting effective explorative experiences.

2.3. **Manipulation of 3D images on 2D displays.** Manipulating 3D virtual objects on multi-touch 2D surfaces is not trivial [18-21,35,36,40]. Davidson and Han [35] suggest that objects' movement in the $z$-axis can be done using pressure. Hancock et al. [18] propose a technique to perform rotation and translation movements with a single finger, but 3D operations (such as rolling and pitch) require two different touches, one for selecting the object and the other for gesturing. In another technique by Hancock and his colleagues [21], users need to first define a rotational axis using two fingers and then using a third finger to do rotation motions. Studies show that both techniques could be learned but are not '*natural*' and are not easy to perform [37]. The proposed techniques are based mainly on surface gestures performed on or close to surface of the display and not through a mobile device. In all these techniques, researchers are concerned with providing either (1) more ways for users to express their actions (e.g., [21,40]) or (2) interactions that are natural (e.g., [37]). Our work brings together these two concerns because our aim is to explore how users will make use of the multiple mechanisms for expressing their actions through mobile devices as well as to identify which interactions are considered natural to users. Our approach is through the use of a user-elicitation study [13,23].

2.4. **User-elicitation study of natural interactions.** A common approach to identifying new interaction techniques that are considered natural to users is through user-elicitation or guessability studies [13,23], an important component of participatory design in HCI [38]. The main idea of a guessability study is to observe what actions users will follow given the effect of an interaction or gesture. That is, users are provided the outcome of an action and are asked to state how they can reproduce the same effect. From observations across a group of users, we find whether there are patterns and consensus about how a gesture is performed.

We next describe our user-elicitation and discuss the results afterwards.

3. **Research Methodology.** The primary goal of this study was to identify which of the following sensing mechanisms and their combinations of a dual-surface tablet device users would find easy and prefer to use: (1) Front-side multi-touch surface; (2) Back-side multi-touch surface; (3) Gyroscope (for orientation); and/or (4) Accelerometer (for tilt). In addition, we also wanted to assess how users would carry out 3D manipulations with the device and see if any agreement could be identified.

3.1. **Participants and equipment.** We recruited 12 participants (10 male) from a local university between the ages of 22 to 35. All participants had some experience with touch enabled mobile devices.

Our experimental prototype was a dual-surface device created by placing back-to-back two Acer Iconia tablets running Android OS. The prototype had a 10.1″ multi-touch surface on the front and back and were connected through a wireless network. Each tablet supported up to ten touches simultaneously and came with an accelerometer and gyroscope. Users could perform surface gestures by sliding one finger or a set of fingers while motion gestures could be done through rotating (rolling, pitching, or yawing) the device. The device allowed immediate visual feedback of all users' touches on its front surface.

3.2. **Tasks.** Participants were asked to design and perform a gesture (surface, motion, or a hybrid of the two) via the dual-surface device (a cause) that they could potential use to carry out the task (an effect). There were 14 different tasks (see Table 1). We asked participants to do a gesture twice and explain why they chose the gesture. Participants were not told of the difference between surface and motion gestures, but only asked to perform a gesture that they felt comfortable doing.

3.3. **Procedure.** Participants were first handed the device so that they could get a feel for it; they began the experiment when ready. Each participant was asked to define a set of gestures for the above listed 14 different 3D manipulations using the dual-surface device.

The 14 manipulations were graphically demonstrated via 3D animations on a display. After an animation was run once, a researcher would explain the task for clarity. The animation could be replayed as many times as needed. The participant was then asked to create a gesture to reproduce the operation seen in the animation. This could be with any sensory input, or combination of several sensors, they wanted and in whatever manner they wished. While creating their gesture, the participant was asked to think aloud – that is, to verbalize what they were doing and why. Afterward s/he was asked to sketch or write a short description of the gesture on paper. This process was repeated for all the 14 manipulation animations shown in Table 1.

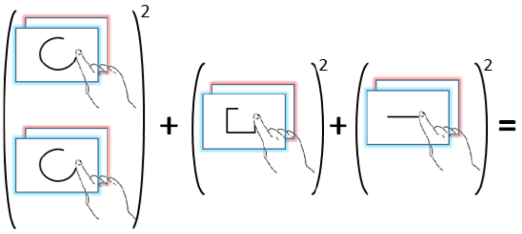
Table 1. The 3D tasks given to participants by category

| **3D Manipulation Tasks** | |
|---|---|
| Manipulation | Animation Descriptions |
| Rotation | |
| About $X$ Axis | Rotate the cube so that the **top face** is facing **forward** |
| About $Y$ Axis | Rotate the cube so that the **left face** is facing **forward** |
| About $Z$ Axis | Rotate the cube so that the **top-right** corner becomes the **top-left** corner |
| Translation | |
| Along $X$ Axis | Move the **red** cube **beside** the **blue** cube (i.e., red cube left side of blue cube) |
| Along $Y$ Axis | Move the **red** cube **on top** of the **blue** cube |
| Along $Z$ Axis | Move or push the **red** cube **back** towards the **blue** cube |
| Stretch | |
| Along $X$ Axis | Stretch the cube **horizontally** to the **right** |
| Along $Y$ Axis | Stretch the cube **vertically up** |
| Along $Z$ Axis | Stretch the cube by **pulling** the cube **forwards** |
| Plane Slicing | |
| $XZ$ plane | Cut the cube into an **upper and lower** portion |
| $YZ$ plane | Cut the cube into a **left and right** portion |
| $XY$ pane | Cut the cube into a **front and back** portion |
| Selection | |
| 2D | Select the cube in the **top-left** corner |
| 3D | Select the cube in the **back bottom-left** corner, hidden behind the front bottom left cube |

3.4. **Results.** From the collected gestures, we were able to create a set of gestures that were natural to users. We grouped identical gestures for each task, and the largest group was chosen as the user-defined gesture for the task. The set composed of the largest group for each task represents the user-defined gesture set. We then calculated an agreement score [13,23] for each task using the group size. The score reflects in one number the degree of consensus among participants. The formula for calculating the agreement scores is:

$$A_t = \sum_{P_i} \left( \frac{P_i}{P_t} \right)^2 \tag{1}$$

where $t$ is a task in the set for all tasks $T$; $P_t$ is the set of proposed gestures for $t$; and $P_i$ is a subset of identical gestures from $P_t$. The range for $A_t$ is between 0 and 1 inclusive. As an example let us assume that for a task, four participants each gave a gesture, but only two are very similar. Then the agreement score would be as follows:



$$A_t = (2/4)^2 + (1/4)^2 + (1/4)^2 = 0.375 \tag{2}$$

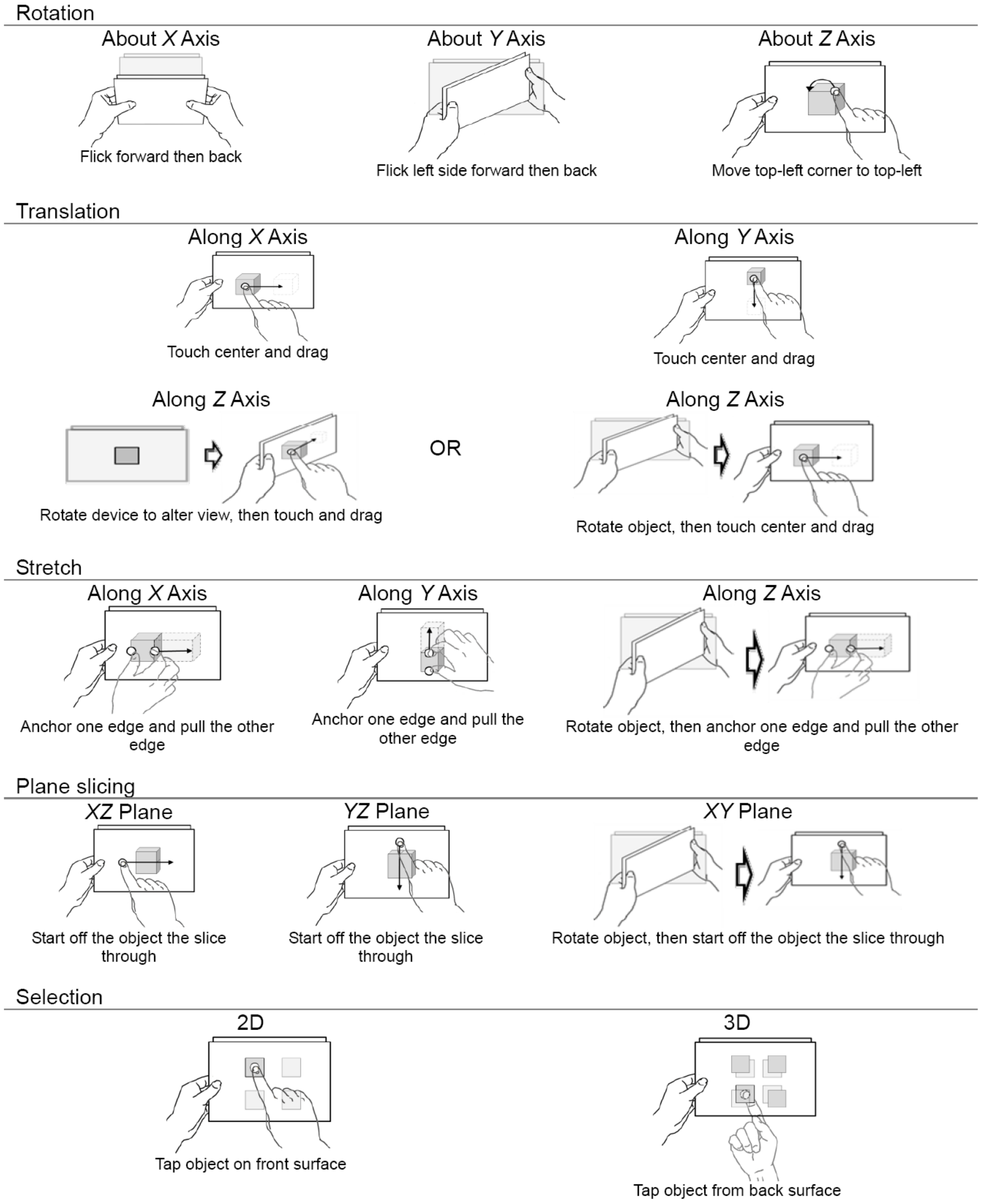


FIGURE 2. Resulting user-defined interaction set

Figure 3 shows the agreement scores for the gesture set, ordered in descending order. The highlighted square shows the gestures with relatively high agreement scores. The scores involving the $Z$ Axis are located at the lower end, indicating a lower consensus. Figure 2 shows the resulting 3D gestures as agreed by the participants.

Figure 4 shows the user-defined gestures grouped by the sensory input. Participants were allowed to use compound gestures. For example, to move an object along the $Z$ Axis, some participants asked if they could rotate the entire scene and then perform a

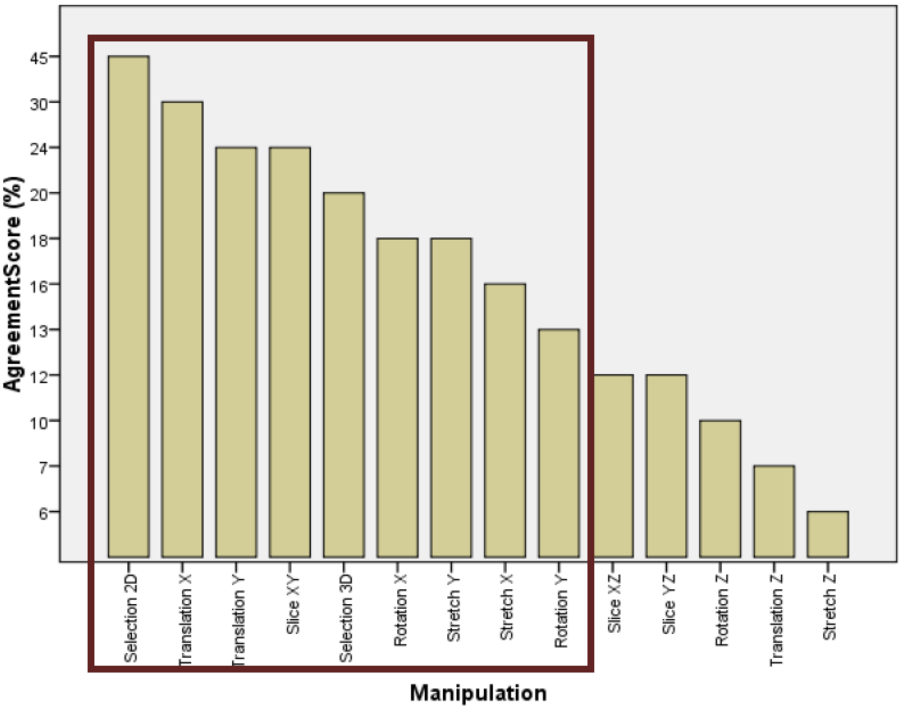
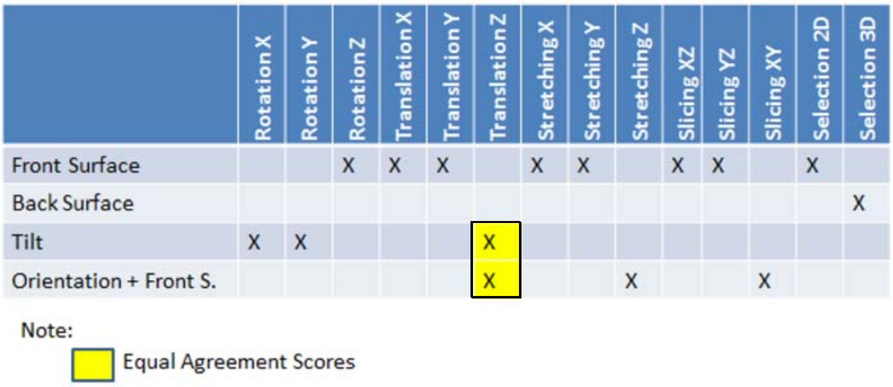


FIGURE 3. Agreement scores for all tasks sorted in descending order



| | RotationX | RotationY | RotationZ | TranslationX | TranslationY | TranslationZ | StretchingX | StretchingY | StretchingZ | Slicing XZ | Slicing YZ | Slicing XY | Selection 2D | Selection 3D |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Front Surface | | | X | X | X | | X | X | | X | X | | X | |
| Back Surface | | | | | | | | | | | | | | X |
| Tilt | X | X | | | | X | | | | | | | | |
| Orientation + Front S. | | | | | | X | | | X | | | X | | |

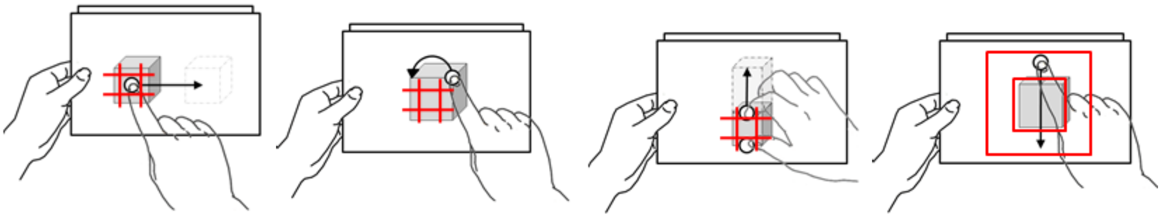
Note:
🟨 Equal Agreement Scores

FIGURE 4. Gestures grouped by sensory input

gesture along the $X$ or $Y$ axes to obtain the same result. The yellow cells correspond to interactions with equal agreement scores for a given input method. The *front* touch-screen surface seems to be most frequently used input method, followed by both *tilt* and *orientation+front surface*, and finally by back-side surface.

3.5. **Discussion.** From Figure 3, we observe that the agreement scores are high for tasks related to the $X$ and $Y$ axes, unlike the scores for tasks in the $Z$ Axis. This shows that gestures along the $Z$ Axis are difficult to perform and not very natural to users. We observed that if a participant could not think of a gesture for manipulating the 3D object along the $Z$ Axis, they would ask if the scene could be rotated in order to perform the manipulation using a gesture along the $X$ and $Y$ axes.

Similarly, from Figure 4, it seems that participants preferred using surface gestures over motion gestures. Figure 3 indicates that participants made use of motion gestures, but mainly for rotation tasks and tasks dealing with the $Z$ Axis. During the study, we



(a) Translation (b) Rotation (c) Scaling (d) Slicing

Figure 5. Specific manipulation regions for four tasks

observed that most participants did not like to make large movements with the dual-surface device to create gestures. This shows that, although participants can make use of motion gestures, there seemed to be some hesitation, perhaps due to their unfamiliarity with motion gestures.

From Figures 2 and 4, we can see that most gestures were carried out on the front-side of the dual-surface device. That is, the front-side was the main input space. Figure 4 shows that the back-side was not used frequently. The few gestures that were performed on the back were unique among participants, and they therefore produced low agreement scores (see Figure 3).

The results also seem to show that when conducting 3D exploratory activities participants find that performing dual-surface interaction using multiple fingers is difficulty. At the very most, participants use two fingers and of the same hand to perform only the three stretch activities (see Figure 2).

Finally, there is one more observation that Figures 2, 3 and 4 show and that is that participants would touch (or begin to make a gesture from certain regions on or around the 3D object to perform interactions. For example, to stretch along the $X$ Axis, many participants would usually begin by touching the midpoint of the object's left and right edges. The same pattern was found for other tasks, especially those with high agreement scores (see Figure 5 for other tasks).

In summary, from this study, we observed that (1) participants preferred to perform actions on the front-side of the dual-surface device; (2) they preferred to perform surface gestures along the $X$ and $Y$ axes; (3) they had difficulty using multiple touches and both surfaces at the same time; and (4) they touched specific regions (or 'hotspots') on virtual objects when performing gestures. These findings led us to modify our experimental device and design a new interface for 3D manipulation of distant objects, SquareGrids.

4. **SquareGrids: An Interface for 3D Manipulation of Distant Objects through Gestures.** SquareGrids used a single-sided multi-touch tablet with an accelerometer and gyroscope (see Figure 6). Based on the gesture input mappings obtained from the first experiment, the touch surface and the accelerometer were used as the primary input mechanisms. In addition, a new graphical interface was developed for the tablet based on the hotspots touched by users when manipulating objects.

The interface was partitioned into 3 major regions, *on-object*, *off-object* and *environment manipulations* (Figure 7). The center of the interface consisted of a 3×3 grid representing the nine regions (or hotspots) that map to the 3D object designated for *on-object* interactions (Region 1 in Figure 7). The middle region (Region 2; area contained within the orange box but outside of the 3×3 grid) was an area designated for *off-object* interactions. A combination of *off-* and *on-object* interactions could be defined. For instance, most participants preferred to start the plane slicing gesture just outside the 3D object's boundaries then slice through the object (see Figure 3 for plane slicing gestures). Outside

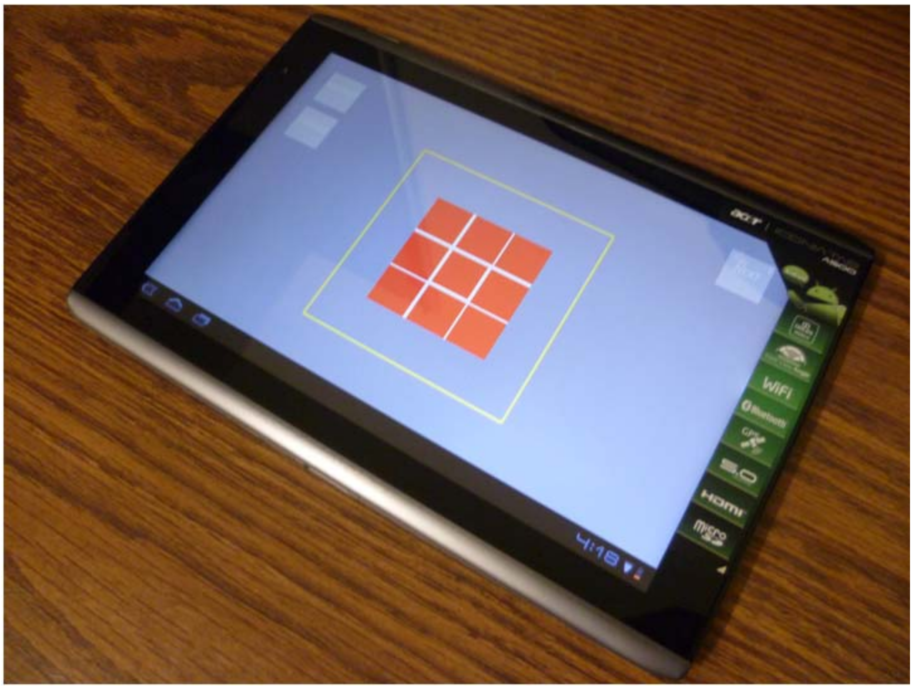
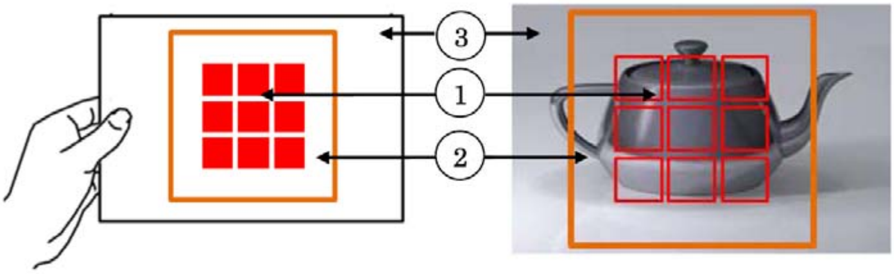


Figure 6. SquareGrids: A potential interface for 3D manipulations of distant objects



(a) View from tablet device (b) View from the large display (without the squares)

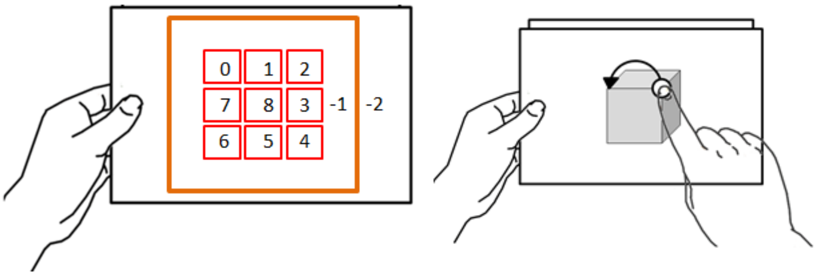
Figure 7. Mapping of the three main regions for (1) *on-object*, (2) *off-object*, (3) *environment manipulations* of the SqureGrids interface (a) to a 3D object displayed on other screen (b).

the orange box was a region for *environment interactions* (Region 3). If gestures were performed in this region, a user could manipulate the entire 3D scene, for example to change the camera's point of view.

Each region and its subdivision were assigned an ID (see Figure 8(a)). As users drag their fingers across the regions of the interface to perform a gesture, a sequence of numbers would be generated. For instance, the gesture in Figure 8(b) would generate the number sequences 2, $-1$, 0. As the gesture is being performed, the gesture recognition engine then checks the number sequence against a set of predefined gesture sequences. Once the engine recognizes the gesture, the correct 3D transformation is invoked. The gesture would continue until the user stopped the gesture motion.

We conducted a preliminary usability study to assess the usability of SqaureGrids against the traditional mouse for 3D manipulations.

4.1. **Participants, equipment, and tasks.** Six male participants between the ages of 23 and 35 were recruited from a local university to participate in this study. All



(a) (b)

FIGURE 8. (a) Assignment of ID numbers to reach region and (b) a user performing a gesture with sequence 2, −1, 0

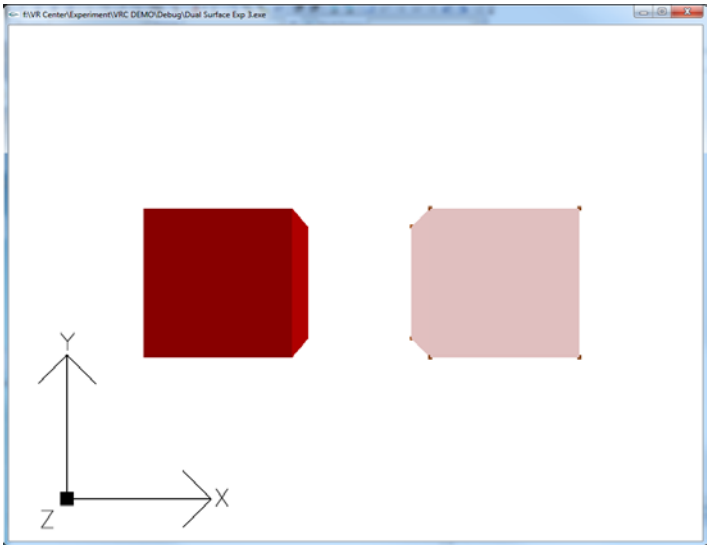


FIGURE 9. The 3D manipulation task: (1) match the left solid to the right solid in terms of size; (2) move the left solid inside the right solid.

participants used computers on a daily basis and are familiar with touch-based mobile devices.

To conduct this experiment we used a desktop computer (with 1.86 GHz Core 2 Duo running Windows XP) with a regular USB mouse and connected to an external 24″ LCD monitor. In addition, we had a laptop (with 2.0 GHz Dual Core and an Intel GMA running Windows XP) connected to another 24″ LCD monitor which was linked to the mobile device prototype via a wireless network.

The task was to manipulate a solid red block by rotating, scaling and/or slicing it so that it would match in size a semi-transparent block and then dock the solid red block inside the semi-transparent block (Figure 9).

4.2. **Conditions and procedure.** This study compared two interfaces: *Mouse* (GUI-based interactions) and *Tablet* (with *SquareGrids*). In the Mouse condition participants interacted with a toolbar to select the manipulation mode and handles on the 3D object to interact with it. In the Tablet condition, participants interact with the 3D object via SquareGrids.

Each trial consisted of these tasks: *Rotation*, *Scaling* or *Plane Slicing* of the 3D object, followed by Translation of the object to dock it inside the semitransparent solid.

We first explained how each of the two interfaces would work and then gave participants practice trials (3 for Translation; 3 for Rotate+Translate; 3 for Scale+Translate; and 3 for Slice+Translate). In the actual experiment, participants repeated the same type of tasks, but these were slightly more complicated. The experiment lasted an hour.

We used a within-subject design. The independent variables were *Interface* (*Mouse* and *Tablet*) and *Task Type*. The order of the presentation of the interface was counterbalanced using a Latin Square design. We collected mainly qualitative feedback in this study because we were more concerned with participants' subjective feedback about the interface.

4.3. **Results.** Results indicated that participants completed the manipulation and docking tasks slightly faster with the traditional Mouse interface. These results could be partly due to the fact that most users were familiar with this type of interface because of frequent use. Although participants had only limited time to learn how to use SquareGrids, they were still able to complete all the given tasks without much difficulty. This shows that SquareGrids was effective. Participants commented that with some more training time they could perform the tasks faster.

As predicted, participants naturally and automatically guided themselves to the hotspots of the Tablet interface to start interacting with the 3D objects. This study therefore confirms the results of our first experiment. One interesting aspect that we observed is that participants had some difficulty making precise motion gestures, but only at the beginning. Gradually, they were able to have better control and used it in conjunction with surface gestures in a highly coordinated manner. Some participants commented that motion gestures made it more interesting and engaging than click-and-drack operations of the mouse. Many participants were actually not in a rush to complete the tasks using the Tablet interface, but wanted to take their time to experiment with the displayed 3D objects using a variety of gestures with the tablet device. This could have explained the somewhat slower performance in the Tablet condition. We believe that this observation indicates that gaming companies can leverage the use of motion sensors in their controllers to add an extra interactive dimension for gamers to enjoy during gameplay. Furthermore, participants found it interesting to have the additional display on the tablet device and faced almost no difficulties to get used to having two displays to look at, switching back and forth with ease. Actually, they commented that the extra information provided in the tablet display helped to increase their accuracy in the tasks. This is useful because it confirms that using two displays of different sizes, with one being static and one mobile, is not disorienting and that they can be complementary. That is, it is possible to design a system for cases in which a user may want to hide some information to be seeing only by her on the mobile device, while being selective as to what information she wants to show in the external display. This scenario is particularly important in public settings or in cases where multiple users are using the same display. In addition, participants commented that they liked the ability to move around with the tablet device, something not possible with the Mouse interface.

Finally, in terms of subjective preference, participants commented that they enjoyed using the Tablet interface more than the Mouse interface and could see themselves using the interface in future applications. They commented that the Tablet interface was "*easy to learn and use*" and that it was also "*very intuitive and interesting to use*". When asked about which interface is better in a public setting or in situations with several users at the same time, they all seemed to agree that they would prefer the Tablet interface. In short, the Tablet interface has the potential to become an effective and efficient input device through which users can interact with 3D information shown in a distant display.

## 5. Conclusions.

### 5.1. Summary.

In this paper we describe the results of an exploratory study about using mobile devices to manipulate 3D objects shown in external displays. This research attempts to answer two questions: (1) what sensory inputs in a tablet mobile device people will find easy and intuitive to use; and (2) what interactions users will find natural to perform with such a device. Our initial prototype presents users with two multi-touch displays, one in the front and the other on the back, to detect surface gestures along with an accelerometer and gyroscope to detect motion gestures. In line with some recent research practices in human-computer interaction, we use a guessability or user-elicitation study in which users are asked to provide what actions come to them naturally and easily when an outcome of an interaction is shown to them. After aggregating the results of the study we find that users in general find it easy and more natural to use the front multi-touch display to perform surface gestures. They are also comfortable with motion gestures but tend to do small motions with the device. If given a choice, users will prefer to use surface gestures than motion gestures. In addition, participants do not find using multiple fingers to easy to perform. In terms of 3D manipulations, we find that users are adept at performing interactions dealing with the $X$ and $Y$ axes but have difficulty with $Y$ axis. Based on the results of the first study, we develop a new interface and compare it against the mouse interface. The results of this study show that our interface has great potential and users indicate that the new interface feels natural, easy to use, and is suitable for performing 3D tasks with objects displayed at a distance.

### 5.2. Implications for user interface design.

A few implications can be derived from this work. *First*, more input modalities may not be better. As our study results suggest, despite the availability of sensors which can detect motion (both tilt and rotation), users have difficulty performing these types of actions. The size of our device could have affected users in making motion gestures, and a smaller device (e.g., a smartphone) would perhaps lend itself better in supporting motions. Therefore, when dealing with tablets of 10.1″ or greater in size, designers should minimize the use of motion gestures. *Second*, although research has shown that the back-side could enrich users' interactive experiences, our results show that users, given the choice of using the front-side, will try to minimize their use of the back-side. Such is the case despite the fact that the back-side would have enabled them to use several fingers simultaneously, potentially facilitating concurrent operations. As such, designers should perhaps maximize the use of the front-side. *Third*, we observed that even using the front-side, users would rarely rely on multiple fingers to perform gestures. This observation indicates that users may have difficulty employing multiple touches at once, and therefore designers should be careful when designing gestures based on multi-finger operations using a 10.1″ handheld tablet.

### 5.3. Applications of this research.

The results of this work have wide applicability in the development of large interactive media displays. There is a trend of large displays being deployed in public settings and more people having mobile devices. Architects for example are experimenting with the displays, to which they refer as media façade, to make public spaces more pleasing and engaging. Similarly, human-computer interaction practitioners are exploring the use of mobile devices for interacting with these media displays. Our results can guide the design of the interactive features of these devices so that interactions with public media displays are natural, simple to use, and effective. This is important because when in public, people do not have time to spend learning how to interact with these displays, and therefore designers should try to make the systems to be as close as possible to walk up and use scenarios.

Another domain which can benefit from this research is video games as there is significant overlap between the context of this research and gameplay: gamers need a state-of-art controller to interact with often 3D environments displayed in a large TV located at a distance. The controller is effectively an input device, while the TV is a large display. There is a strong emphasis for game designers to develop forward-looking controllers that can provide richer gaming experiences. Major gaming companies are now looking into adding a variety of input sensors into the controllers. Their designers have begun experimenting with screens with multi-touch capabilities, touchpads placed on the controllers' backside, and motion sensors attached to the controllers. The results of this research can be used to inform the design of this new generation of controllers and the ways in which gamers can interact with the TV content.

Finally, we also believe that this line of research is important for applications running across several displays, often referred to as multi-display environments [39]. It is not uncommon to find someone who has a smartphone placed next to her tablet device and laptop. In these environments information can be displayed in more than one screen, which is similar to our mobile device-large display scenario. Multi-display environments can be found in offices, meeting rooms, classrooms, libraries, and even in public places. This line of research can shed light for designing the portable devices within these multi-display environments.

**Acknowledgment.** We thank the participants for their time. We would also like to thank the reviewers for their comments and suggestions which have helped to improve the quality of the paper. We acknowledge NSERC and the Virtual Reality Centre for partially funding this project.

## REFERENCES

[1] R. Ballagas, M. Rohs, J. Sheridan and J. Borchers, The smart phone: A ubiquitous input device, *IEEE Pervasive Computing*, vol.5, no.1, pp.70-77, 2006.

[2] T. Pering, R. Ballagas and R. Want, Spontaneous marriages of mobile devices and interactive spaces, *Communications of the ACM*, vol.48, no.9, pp.53-59, 2005.

[3] D. C. McCallum and P. Irani, ARC-Pad: Absolute+relative cursor positioning for large displays with a mobile touchscreen, *Proc. of the ACM Symposium on User Interface Software and Technology*, Victoria, British Columbia, Canada, pp.153-156, 2009.

[4] H.-N. Liang, J. Trenchard, M. Semegen and P. Irani, An exploration of interaction styles in mobile devices for navigating 3D environments, *Proc. of the Asia Pacific Conference on Computer Human Interaction*, Matsue-City, Shimane, Japan, 2012.

[5] S. Boring, M. Jurmu and A. Butz, Scroll, tilt or move it: Using mobile phones to continuously control pointers on large public displays, *Proc. of the Conference of the Australian Computer-Human Interaction Special Interest Group (CHISIG) of the Human Factors and Ergonomics Society of Australia (HFESA)*, Melbourne, Australia, 2009.

[6] S. Boring, S. Gehring, A. Wiethoff, M. Blöckner, J. Schöning and A. Butz, Multi-user interaction on media facades through live video on mobile devices, *Proc. of the International Conference on Human Factors in Computing Systems*, Vancouver, BC, Canada, 2011.

[7] Y. Du, H. Ren, G. Pan and S. Li, Tilt & touch: Mobile phone for 3D interaction, *Proc. of the International Conference on Ubiquitous Computing*, Beijing, China, 2011.

[8] B. Ens, R. Eskicioglu and P. Irani, Visually augmented interfaces for co-located mobile collaboration, *Distributed User Interfaces: Designing Interfaces for the Distributed Ecosystem*, pp.169-176, 2011.

[9] S. Jeon, J. Hwang, G. J. Kim and M. Billinghurst, Interaction with large ubiquitous displays using camera-equipped mobile phones, *Personal and Ubiquitous Computing*, vol.14, no.2, pp.83-94, 2010.

[10] U. Rashid, M. A. Nacenta and A. Quigley, The cost of display switching: A comparison of mobile, large display and hybrid UI configurations, *Proc. of the International Working Conference on Advanced Visual Interfaces*, Capri Island, Italy, 2012.

[11] A. Benzina, A. Dey, M. Tonnis and G. Klinker, Empirical evaluation of mapping functions for navigation in virtual reality using phones with integrated sensors, *Proc. of the Asia Pacific Conference on Computer Human Interaction*, Matsue-City, Shimane, Japan, 2012.

[12] T.-H. Chang and Y. Li, Deep shot: A framework for migrating tasks across devices using mobile phone cameras, *Proc. of the International Conference on Human Factors in Computing Systems*, Vancouver, BC, Canada, 2011.

[13] J. Ruiz, Y. Li and E. Lank, User-defined motion gestures for mobile interaction, *Proc. of the International Conference on Human Factors in Computing Systems*, Vancouver, BC, Canada, pp.197-206, 2011.

[14] E. Eriksson, T. R. Hansen and A. Lykke-Olesen, Reclaiming public space-designing for public interaction with private devices, *Proc. of the International Conference on Tangible and Embedded Interaction*, Baton Rouge, LA, USA, 2007.

[15] S. Boring, D. Baur, A. Butz, S. Gustafson and P. Baudisch, Touch projector: Mobile interaction through video, *Proc. of the International Conference on Human Factors in Computing Systems*, Atlanta, GA, USA, pp.2287-2296, 2010.

[16] H.-N. Liang and K. Sedig, Role of interaction in enhancing the epistemic utility of 3D mathematical visualizations, *International Journal of Computers for Mathematical Learning*, vol.15, no.3, pp.191-224, 2010.

[17] T. Grossman, R. Balakrishnan, G. Kurtenbach, G. Fitzmaurice, A. Khan and B. Buxton, Interaction techniques for 3D modeling on large displays, *Proc. of the Symposium on Interactive 3D Graphics*, New York, NY, USA, pp.17-23, 2001.

[18] M. Hancock, S. Carpendale and A. Cockburn, Shallow-depth 3D interaction: Design and evaluation of one-, two- and three-touch techniques, *Proc. of the Interactional Conference on Human Factors in Computing Systems*, San Jose, CA, USA, pp.1147-1156, 2007.

[19] A. Martinet, G. Casiez and L. Grisoni, The effect of DOF separation in 3D manipulation tasks with multi-touch displays, *Proc. of the Symposium on Virtual Reality Software and Technology*, Hong Kong, China, pp.111-118, 2010.

[20] L. Vlaming, C. Collins, M. Hancock, M. A. Nacenta, T. Isenberg and S. Carpendale, Integrating 2D mouse emulation with 3D manipulation for visualizations on a multi-touch table, *Proc. of the International Conference on Interactive Tabletops and Surfaces*, Saarbrucken, Germany, pp.221-230, 2012.

[21] M. Hancock, T. Cate and S. Carpendale, Sticky tools: Full 6DOF force-based interaction for multi-touch tables, *Proc. of the International Conference on Interactive Tabletops and Surfaces*, Banff, Alberta, Canada, pp.145-152, 2009.

[22] U. Hinrichs and S. Carpendale, Gestures in the wild: Studying multi-touch gesture sequences on interactive tabletop exhibits, *Proc. of the International Conference on Human Factors in Computing Systems*, Vancouver, BC, Canada, pp.3023-3032, 2011.

[23] J. O. Wobbrock, M. R. Morris and A. D. Wilson, User-defined gestures for surface computing, *Proc. of the International Conference on Human Factors in Computing Systems*, Boston, MA, USA, pp.1083-1092, 2009.

[24] A. Bragdon, E. Nelson, Y. Li and K. Hinckley, Experimental analysis of touch-screen gesture designs in mobile environments, *Proc. of the International Conference on Human Factors in Computing Systems*, Vancouver, BC, Canada, pp.403-412, 2011.

[25] H. Lü and Y. Li, Gesture avatar: A technique for operating mobile user interfaces using gestures, *Proc. of the International Conference on Human Factors in Computing Systems*, Vancouver, BC, Canada, 2011.

[26] Y. Li, Gesture search: A tool for fast mobile data access, *Proc. of the Symposium on User Interface Software and Technology*, New York, NY, USA, pp.207-216, 2010.

[27] D. Wigdor and R. Balakrishnan, TiltText: Using tilt for text input to mobile phones, *Proc. of the ACM Symposium on User Interface Software an Technology*, pp.81-90, 2003.

[28] M. Rahman, S. Gustafson, P. Irani and S. Subramanian, Tilt techniques: Investigating the dexterity of wrist-based input, *Proc. of the ACM Conference on Human Factors in Computing Systems*, Boston, MA, USA, pp.1943-1952, 2009.

[29] P. Baudisch and G. Chu, Back-of-device interaction allows creating very small touch devices, *Proc. of the ACM Conference on Human Factors in Computing Systems*, Boston, MA, USA, pp.1923-1932, 2009.

[30] J. Scott, S. Izadi, D. Ruszkowski, X. Bi and R. Balakrishnan, RearType: Text entry using keys on the back of a device, *Proc. of the International Conference on Human Computer Interaction with Mobile Devices and Services*, Lisbon, Portugal, pp.171-180, 2010.

[31] M. Silfverberg, P. Korhonen and I. S. MacKenzie, Zoomig and panning content on a display screen, *United States Patent 7075513*, 2006.

[32] D. Wigdor, C. Forlines, P. Baudisch, J. Barnwell and C. Shen, LucidTouch: A see-through mobile device, *Proc. of the ACM's 20th Symposium on User Interface Software and Technology*, pp.269-278, 2007.

[33] X.-D. Yang, E. Mak, P. Irani and W. F. Bischof, Dual-surface input: Augmenting one-handed interaction with coordinated front and behind-the-screen input, *Proc. of the International Conference on Human Computer Interaction with Mobile Devices and Services*, Bonn, Germany, pp.1-10, 2009.

[34] M. Sugimoto and K. Hiroki, HybridTouch: An intuitive manipulation technique for PDAs using their front and rear surfaces, *Proc. of the International Conference on Human Computer Interaction with Mobile Devices and Services – Extended Abstracts*, pp.137-140, 2006.

[35] P. L. Davidson and J. Y. Han, Extending 2D object arrangement with pressure-sensitive layering cues, *Proc. of the ACM Symposium on User Interface Software an Technology*, pp.87-90, 2008.

[36] A. Wilson, S. Izadi, O. Hilliges, A. Garcia-Mendoza and D. Kirk, Bringing physics to the surface, *Proc. of the ACM Symposium on User Interface Software an Technology*, Monterey, CA, USA, pp.67-76, 2008.

[37] O. Hilliges, S. Izadi, A. D. Wilson, S. Hodges, A. Garcia-Mendoza and A. But, Interactions in the air: Adding further depth to interactive tabletops, *Proc. of the ACM Symposium on User Interface Software an Technology*, Victoria, BC, Canada, pp.139-147, 2009.

[38] D. Schuler, *Participatory Design: Principles and Practices*, L. Erlbaum Associates, Hillsdale, NJ, USA, 1993.

[39] M. A. Nacenta, C. Gutwin, D. Aliakseyeu and S. Subramanian, There and back again: Cross-display object movement in multi-display environments, *Human-Computer Interaction*, vol.24, no.1, pp.170-229, 2009.

[40] C. Williams, X.-D. Yang, G. Partridge, J. Usiskin-Miller, A. Major and P. Irani, TZee: Exploiting the lighting properties of multi-touch tabletops for tangible 3D interactions, *Proc. of the 29th International Conference on Human Factors in Computing Systems*, Vancouver, BC, pp.1363-1372, 2011.