Available online at www.sciencedirect.com

SciVerse ScienceDirect

Int. J. Human-Computer Studies 70 (2012) 690–702

International Journal of Human-Computer Studies

www.elsevier.com/locate/ijhcs

MultiPoint: Comparing laser and manual pointing as remote input in large display interactions

Amartya Banerjee, Jesse Burstyn, Audrey Girouard*,[1], Roel Vertegaal

Human Media Lab, School of Computing, Queen's University, Kingston, Ontario, Canada K7L 3N6

Received 1 July 2011; received in revised form 1 March 2012; accepted 18 May 2012
Available online 31 May 2012

Abstract

We present MultiPoint, a set of perspective-based remote pointing techniques that allows users to perform bimanual and multi-finger remote manipulation of graphical objects on large displays. We conducted two empirical studies that compared remote pointing techniques performed using fingers and laser pointers, in single and multi-finger pointing interactions. We explored three types of manual selection gestures: squeeze, breach and trigger. The fastest and most preferred technique was the trigger gesture in the single point experiment and the unimanual breach gesture in the multi-finger pointing study. The laser pointer obtained mixed results: it is fast, but inaccurate in single point, and it obtained the lowest ranking and performance in the multipoint experiment. Our results suggest MultiPoint interaction techniques are superior in performance and accuracy to traditional laser pointers for interacting with graphical objects on a large display from a distance.

Keywords: Multi-touch; Remote interaction; Large display; Input device; Interaction technique

1. Introduction

Over the past few years, interactive large displays have gained traction as a vehicle for public and large-scale media—with applications in advertising, information visualization, and public collaboration (Ball and North, 2007; Brignull and Rogers, 2003). For example CityWall, a large multi-touch display installed at a central location in Helsinki, provided people with an engaging and highly interactive interface in an urban environment (Peltonen et al., 2008). The popularity of large interactive displays in these applications can, in large part, be attributed to their significantly increased screen real estate, which provides more pixels for collaboration, higher densities of information, or better visibility at a distance (Bi and Balakrishnan, 2009). Since large displays provide more physical space in front of the display, they also allow for multi-user applications that are not easily accommodated or communicated via standard desktop monitors (Vogel and Balakrishnan, 2005).

We believe this presents an opportunity to explore interaction techniques that capitalize on the inherent strength of large displays—greater screen real estate—when physical input devices are not readily available. While many innovative techniques have been proposed in the literature to deal with the difficulties in pointing at hard-to-reach parts of a large display, the majority focus on within-arms-reach interactions through touch or multi-touch, with the underlying assumption that the user stands sufficiently close to the screen to touch its surface (Brignull and Rogers, 2003; Myers et al., 2002; Peltonen et al., 2008). Alternatively, they require users to navigate a mouse cursor using some form of traditional pointing device (Baudisch et al., 2007).

1.1. Issues with walk-up-and-use

As Ringel et al. (2001) point out, the classic problem with multi-touch large display interactions is that users are

---

*Corresponding author. Tel.: +1 613 520 5644; fax: +1 613 520 6623.
E-mail addresses: banerjee@cs.queensu.ca (A. Banerjee), jesse@cs.queensu.ca (J. Burstyn), audrey_girouard@carleton.ca (A. Girouard), roel@cs.queensu.ca (R. Vertegaal).

[1] Present address: Carleton School of Information Technology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, Canada K1S 5B6

1071-5819/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ijhcs.2012.05.009

required to walk up to the display to touch objects that are within arm's reach. Not only does this limit interaction with objects that are out of reach, walking results in a much lower Fitts' law performance than manual pointing (Oakley et al., 2008). Streitz et al. (1999) proposed the use of physics as a potential solution for this problem. However, when users step back from the display to view the contents of the entire screen, they can no longer interact with the graphics until they step forward to touch the screen. In the realm of seated cooperative work scenarios, we often observed a plenary turn taking mechanism, with only one user presenting in front of the screen. We believe this is, at least in part, due to the time required to get up and walk to the screen.

1.2. Issues with remote pointing

One solution is to use remote input techniques that allow users to point at large displays from a distance. One method explored is through the use of laser pointers (Myers et al., 2002). The laser pointer can be used from just about any position in front of the display. Unlike mice or styli, laser pointers do not require a surface to track cursor position. However, they present some limitations. First, one has to carry a laser pointer at all times. Second, multipoint techniques are mostly unavailable unless one uses a laser pointer in each hand.

An alternative method is direct freehand pointing, in which computer vision or another input method detects the location of fingers at a distance from the display (Vogel and Balakrishnan, 2005). As with laser pointers, one can perform ray casting using the vector of a pointing finger. However, when multipoint gestures are considered, it is no longer evident which fingers are participating in the gesture, or even that the fingers are directed at the display. As a solution for this, (Jota et al., 2010) explored an image-plane or perspective-based pointing technique (Pierce et al., 1997) that takes into account the line of sight of the user: fingers are directed at the display when they are within the boundary box perceived from the user's perspective. While their system allowed for bimanual input, it did not allow for multipoint gesturing between the hands, or within fingers of one hand.

1.3. Multipoint: multi-touch inspired gestures at a distance

MultiPoint enables users to remotely manipulate content on a large display. By performing multi-touch inspired in-air gestures, a user can perform manipulations similar to those afforded by a touch enabled interactive surface. MultiPoint employs image-plane or perspective-based pointing (Fig. 1) that follows a user's line of sight. Users can perform manipulations either bimanually, or simply with a single hand.

In this paper, we report on two experiments designed to investigate MultiPoint's potential. We explore the affordances associated with in-air interactions and compare them with laser pointer-based interactions. Our first experiment compares remote perspective-based pointing to laser pointing in a single point manipulation task (Fig. 2a). In addition, this experiment evaluates three selection techniques for remote content that have not been compared previously, including one introduced in the g-speak system (Oblong Industries). The second experiment measures the performance of remote multipoint input by comparing unimanual multipoint, bimanual multipoint, and dual laser pointing (Fig. 2b). We conclude with a discussion of the design space surrounding MultiPoint and provide conclusions regarding the suitability of each technique for systems that benefit from in-air interaction.

2. Related work

A large body of literature investigates solutions for walk-up-and-use and remote pointing. MultiPoint builds upon the following areas of previous research: (1) touch-based interaction; (2) device-based remote interaction techniques; (3) device-less remote interaction techniques.

2.1. Touch-based interaction

Touch-based multi-touch tabletop technologies like SmartSkin (Rekimoto, 2002) and DiamondTouch (Dietz and Leigh, 2001) could be used to interact with large upright wall displays. Barehands (Ringel et al., 2001) and Touchlight (Wilson, 2004) use computer vision to track bare, unmarked hands pressing against a vertical surface.

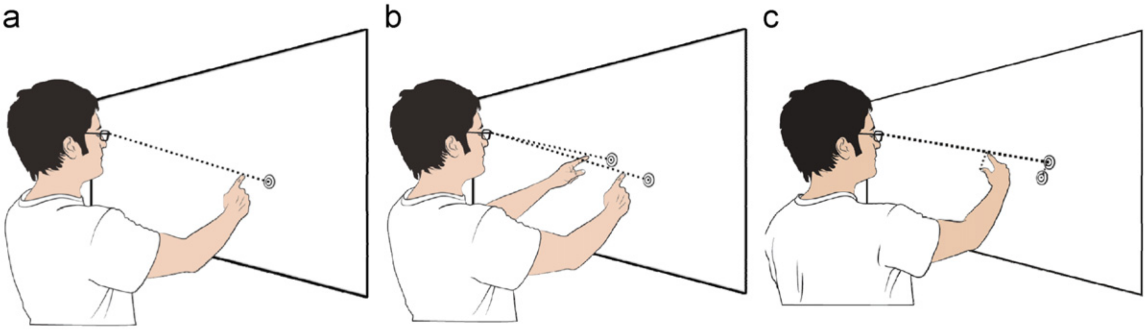

Fig. 1. Remote multipoint techniques. (a) Remote perspective-based single point , (b) Bimanual remote perspective-based multipoint and (c) Unimanual remote perspective-based multipoint

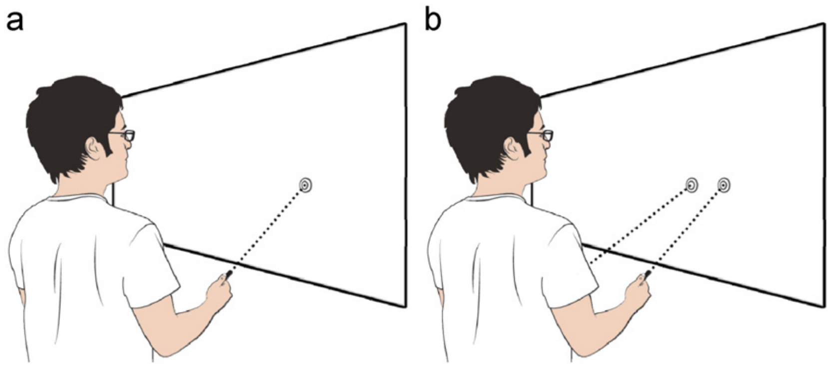

Fig. 2. Laser pointer techniques. (a) Laser pointer based single point and (b) Laser pointer based multipoint

However, these technologies lack the ability to provide remote interaction as both require the hand to be almost in contact with either the tabletop, or a touch-sensitive upright surface to detect the hand image.

Visual Touchpad (Malik and Laszlo, 2004) is a vision-based touch technology emulating touch-based systems by providing an external touchpad mapped 1:1 to the display. With access to an entire 2D hand image, it does not suffer from the finger ambiguity problem of the other systems. It does lack accuracy, as a small position change on the touchpad equates to a large change on the display. To reduce this problem, Touch Projector (Boring et al., 2010) lets users interact with screens at a distance using a freeze frame or zoomed video image on their mobile device. The device tracks itself with respect to the surrounding displays, and a touch on the video image corresponded to a touch event on the target display in view. To design MultiPoint, we drew on this body of prior research to explore the affordances associated with rich sensor data, including but not limited to, touch input for large displays and arm or hand hover information. (Video S1)

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.ijhcs.2012.05.009.

2.2. Remote interaction

Researchers have also designed interaction techniques that allow the user to point and interact with large displays at a distance. We identify related work that use physical devices to perform remote interactions, as well as device-less input.

2.2.1. Device-based interaction

Researchers have applied traditional input devices to large display interactions. In PointRight (Johanson et al., 2002) and I-Room (Tate et al., 2010), the user can use a standard mouse to move the cursor across a display surface composed of different screens. Spotlight (Khan et al., 2005) allows a user to control a large highlighted region across a large display from afar using a mouse, to direct the visual attention of an audience. However, a mouse requires a surface to operate upon.

Extending traditional input devices, Baudisch et al. (2007) developed Soap, an in-air pointing device using an optical mouse encased in a fabric hull. The relative movement between the hull and the sensor was used to define cursor position. Soap provided tactile feedback and interaction techniques for fast cursor navigation across long distances, but it lacked comparison to other remote input devices.

A laser pointer is a common device for remote interactions with large displays (Bolt, 1980; Jota et al., 2010). Myers et al. (2002) assessed the performance of laser pointers in selecting an object on a large screen and compared it to using a mouse; tapping directly on the screen; and a handheld device to capture an area of interest on the screen. The laser pointer recorded the lowest performance. While the laser pointer provides an intuitive way to randomly access any portion of a wall sized display, natural hand jitter makes it difficult to use for accurate target acquisition tasks, particularly for smaller targets. Moreover, ordinary laser pointers have only two degrees of freedom, which limits their use for complicated tasks. Sceptre (Wienss et al., 2006) and Structured Laser Pointer (Qin et al., 2010) presented enhanced laser pointing systems detecting the laser pointer's rotation along its emitting axis.

Pinch Gloves (Bowman et al., 2001) contain electronic sensors embedded in the fingertips of a glove to detect contact between the fingers. Used in virtual reality applications, Pinch Gloves can be employed to assign interactive functions corresponding to touches detected between fingertips. However, these gloves are not designed to facilitate pointing and require a controller unit connected to the gloves with wires.

VisionWand (Cao and Balakrishnan, 2003) uses simple computer vision to track the colored tips of a plastic wand to interact with large wall displays, close-up and from a distance. The inherent presence of a device is the main disadvantage of VisionWand, and of all device-based interaction techniques. The need to carry a specialized device at all times limits casual users, and the number of interactions is restricted by the number of devices available. Finally, in their exploration of pan-and-zoom techniques, Nancel et al. (2011) observed that bimanual input and linear gestures improved performance.

2.2.2. Device-less interaction

Device-less interaction techniques can alleviate the shortcoming of device-based techniques by relying on computer vision to detect hand and finger movements, typically through markers placed on the hands. The major advantage of such vision-based techniques is their ability to track multiple fingers uniquely. However, such remote interaction techniques lack explicit discrete direct inputs such as buttons, making selection techniques and clicks non-trivial.

Wilson (2006) used pinching as a technique for cursor control through robust marker-less computer vision. However, interaction was limited, and required the gesture to be performed over a set background (keyboard), with a close range camera.

Shadow Reaching (Shoemaker et al., 2007) applied a perspective projection to a shadow representation of the user to enable manipulation of distant objects on a large display. The system allows users to interact at a distance while the shadow representation aids in maintaining context in collaborative environments.

The Head Crusher technique casts a ray from the user's eye through the point midway between the user's forefinger and thumb, and onto the scene (Pierce et al., 1997). The object is acquired when it intersects with the ray. Vogel and Balakrishnan (2005) explored single hand pointing and clicking interactions with large displays from a distance. They found ray casting an effective pointing method, and proposed AirTap as a clicking technique for single clicks. Jota et al. (2010) compared four pointing techniques: laser, arrow, image plane and fixed origin. They demonstrated that taking the user's line of sight (i.e. perspective) into account improves performance for tasks requiring more accuracy. Their work was restricted to single, unimanual interactions. Nancel et al. (2011) used bimanual interaction techniques to pan-and-zoom content on a large display.

To our knowledge, the only remote bimanual multipoint systems are the g-speak spatial operating environment (Oblong Industries) and virtual reality applications using Pinch Gloves. In g-speak, the user points at a target by making a trigger gesture (finger pointed towards display, thumb towards the ceiling), and selects by lowering the thumb on top of the index finger (Zigelbaum et al., 2010). However, there are no evaluations of g-speak or of the trigger technique.

3. Multipoint interaction techniques

Most of the present interaction techniques for large displays are limited to up-close interactions using a pen or direct touch. The few systems that do allow interaction from a distance suffer from one or more of issues: an inability to differentiate between the two hands and/or between fingers (Shoemaker et al., 2007), or a trade-off between quick pointing and accurate target acquisition (Vogel and Balakrishnan, 2005). Based on these shortcomings, we have designed a set of interaction techniques called MultiPoint. MultiPoint allows for accurate target acquisition and quick manipulation on large displays from a distance, while eliminating the need for a handheld input device.

MultiPoint uses remote perspective-based pointing gestures, and accommodates both single point and multipoint interactions. By tracking the location of the eyes as well as the location of the index finger and thumb (for unimanual interactions) or the location of both index fingers (for bimanual interactions), the system calculates the position of the cursor(s) on the large display (Fig. 1). This perspective-based technique provides the user, as well as observers, with a more accurate mental model of the mapping between hand location and click location. This is akin to Kendon's work in social anthropology, which classified pointing gestures in the context of what is being pointed at (Kendon, 2004).

3.1. Remote selection techniques

We developed two selection gestures to generate remote click events on a large display, a squeezing gesture and a breach gesture, and we implemented the trigger selection gesture. The user performs these gestures while pointing at the display using his index finger (Jota et al., 2010). Other techniques such as Head Crusher (Pierce et al., 1997) and AirTap (Vogel and Balakrishnan, 2005) were considered. These two techniques were eliminated since both would result in a change in the cursor location during selection. Moreover, the Head Crusher technique uses finger movements similar to a pinch-to-scale gesture that may confuse users accustomed to basic multi-touch gestures.

3.1.1. Squeeze gesture

This gesture is based on the idea of grabbing distant objects. In the squeeze gesture, the user starts with a flat hand, pointed at the display. To click, i.e. generate a click-down event, the user keeps the index pointed at the target,

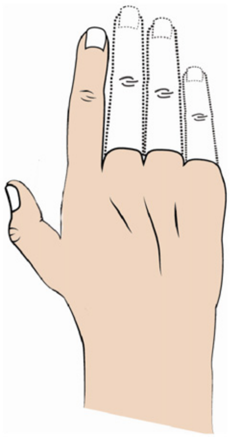

Fig. 3. Remote selection technique—Squeeze gesture. The dotted lines indicate the initial state (flat hand), and the plain lines indicate the selection state (squeezed).

Fig. 4. Remote selection technique—Breach gesture. The dotted lines indicate the initial state (close to the body), and the plain lines indicate the selection state (passed the invisible threshold).

and clenches his middle, ring and little finger (Fig. 3). To generate a click-up event, the user unclenches the last three fingers. The position of the thumb is inconsequential. The configuration of the hand during the click-down event is similar to the Sticky Finger interaction technique for 3 D immersive environments (Pierce et al., 1997). The gesture can result in a minor displacement of the index finger. However, compared to the length of the vector for ray casting with laser pointers, the longer perspective-based pointing vector diminishes most of the potential cursor jitter while clicking.

3.1.2. Breach gesture

This selection mimics the act of touching an invisible touch screen located within arm's reach (Fig. 4). In the breach gesture, the user points at the target using their index finger and pushes their hand towards the screen to select. Subramanian et al. (2006) proposed Pin-Through, a selection gesture for pen-based interaction on tabletops that is similar to the breach gesture. Though Pin-Through recorded low user satisfaction, the breach gesture is simpler. Furthermore, the differences in ergonomic properties between tabletops and vertical displays for analogs' movements motivate further investigation.

A click-down event is generated when the index finger crosses a distance threshold. The click-up event is generated when the index finger is closer than the distance threshold. The index's position and the distance threshold are measured from the user's nose bridge. The threshold is located at two third of an arm's length and is calibrated for each user. This threshold was decided upon based on pilot studies conducted during the design phase. We found that, on an average, most users felt comfortable with click-activation at this distance; full extension of the arms resulted in greater fatigue while shorter distances resulted in the user's hands dominating their field of vision.

3.1.3. Trigger gesture

The gesture uses the metaphor of shooting a gun to select (Fig. 5). The user positions their hand vertically, with the thumb pointing to the ceiling. To select, the user lowers their thumb towards the display, on top of the index finger. This gesture was introduced by Grossman et al. (2004), and reused in the g-speak system (Zigelbaum et al., 2010).

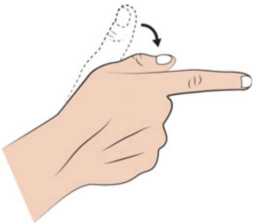

Fig. 5. Remote selection technique—Trigger gesture (right).

3.2. Remote single point

In the remote single point, the cursor is located at the intersection of the display plane and the nose-index vector (Fig. 1a). The nose-index vector is determined through two points in space: the location of the nose bridge, and the location of the index finger (Jota et al., 2010).

In remote single point mode, the user can perform the selection and translation actions. To translate a target, the user selects it, moves their finger to the desired location, and deselects the target.

3.3. Remote multipoint

MultiPoint enables the user to perform in-air bimanual and unimanual multi-touch gestures from a distance. Bimanual remote multipoint gestures use the index of each hand to perform each action, where each index becomes a cursor. Unimanual actions use the index finger and the thumb of the same hand as cursors.

To scale, or zoom, a target, users can choose to perform a single-handed or a bimanual pinch gesture. To rotate, users rotate their arms (or fingers) in a circular path. In unimanual multipoint, the user is required to move both the index finger and the thumb to make the target rotate or scale.

3.3.1. Bimanual multipoint

Bimanual multipoint uses two nose-index vectors to determine the cursor position on the display (Fig. 1b), essentially doubling remote single point. The squeeze, the breach and the trigger interaction techniques are all valid for bimanual multipoint object selection.

3.3.2. Unimanual multipoint

In unimanual multipoint (Fig. 1c), the nose-index vector determines the location of the index cursor. However, we cannot use the same technique to calculate the thumb cursor position: the perspective compounds the distance between the two cursors, making it impossible to select small targets unless the two fingers are touching. Hence,

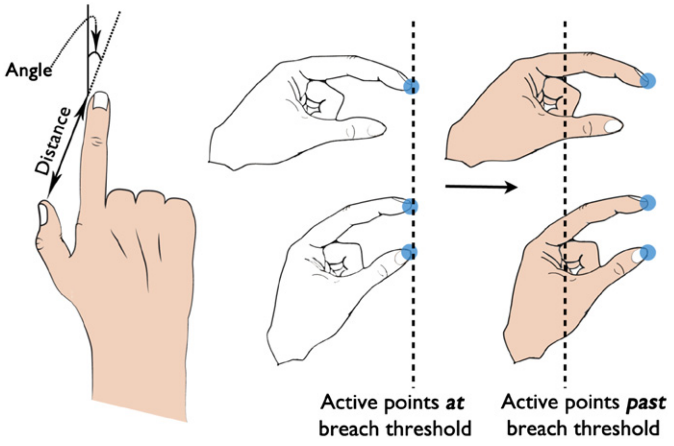

Fig. 6. Unimanual multipoint. Left: the thumb cursor position is determined through the thumb-to-index distance and angle. Right: index breach only (top), thumb and index breach (bottom). Hand configuration while crossing the breach threshold determines the number of active points (in blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

we calculate the thumb cursor position from the index position (Fig. 6, left). This creates more natural and expected cursor movements. The distance between the index cursor and the thumb cursor is proportional to the absolute distance of the fingers, and the angle of the two cursors is identical to that of the two fingers. The distance and angle are taken in 2D space, by projecting the two fingers onto a plane parallel to the display.

Unimanual multipoint restricts which gestures can be used for selection. First, the technique must not use the thumb or the index finger to select. We must eliminate the trigger gesture as it uses the thumb to select, making it impossible to perform multipoint gestures, such as a pinch gesture. Second, the technique must not influence pointing accuracy. Pilot studies showed that the squeeze gesture was hard to correctly apply while performing unimanual interaction. Hence, we restrict unimanual multipoint selection to the breach gesture.

Manipulation mode, i.e. single point or multipoint, is determined based on the configuration of the hand when the breach threshold is crossed. The user can invoke multipoint manipulations by crossing the breach threshold with the index finger and the thumb simultaneously; crossing the breach threshold with the only the index finger, or with the index finger preceding the thumb, results in single point manipulation (Fig. 6, right).

3.4. Click feedback

MultiPoint provides the user with cursors that indicate the specific location of each click event. Since cursor position is calculated by tracking the nose bridge rather than the eyes, there may be a perceived shift in the one-to-one mapping of the cursor position due to ocular dominance. To mitigate this effect, the cursor's horizontal position is calibrated to the user's dominant eye. In addition, using perspective-based cursors can lead to an occlusion of the cursor by the hand (Jota et al., 2010). To address this issue, we placed the cursor a small distance above its calculated position (50 pixels). This offset—with the user standing away from the display—is small enough to not affect the user's perception of directness while alleviating cursor occlusion by the hand. The click-point is resolved to the center of the cursor.

We also incorporated visual feedback in MultiPoint to help participants perceive click events. We selected a progressive indicator instead of a binary one to provide continuous feedback. We display two vertical progress bars, on each side of the display, each representing a cursor (left or right hand, or thumb and index fingers). The progress bar's background color corresponds to each cursor's color. The progress bars turn green at each successful selection.

3.5. Laser pointing

A mouse or a similar pointing device requires a surface to operate on, restricting the user's position. Therefore, we evaluated the MultiPoint interactions techniques against another absolute, surface-less, in-air input device: laser pointing, a commonly used remote pointing technique (Fig. 1d). Single point interactions were performed by holding a wooden dowel emulating a pointer. Bimanual pointing was performed by holding a pointer in each hand. Unimanual interaction cannot be performed through laser pointing: holding two pointers in a single hand is not practical for most users.

4. Multipoint apparatus

Our system uses eight Vicon T40 cameras to track passive infrared retroreflective markers arranged in unique shapes (Fig. 7). We receive data through the Vicon MX Giganet, an image processor that uses a triangulation algorithm to convert the multiple 2D images from each camera to a coordinate in 3D space. Each marker was tracked at 100 Hz, with a precision of 3 mm in a room-sized 3D volume.

Our large display measured 1.65 m × 1.2 m, and was back-projected using a Toshiba X300 short-throw projector running at a resolution of 1024 × 768. MultiPoint was written in C# with WPF4.0.

To track motion with MultiPoint, we affixed marker arrangements on seven objects. For squeeze and breach selection, the user wore gloves: a right glove for single point, and a left and right gloves bimanual multipoint (Fig. 7A and B). We used special left and right gloves for trigger selection that include markers on the thumb (left glove not shown, right glove is Fig. 7C). Unimanual multipoint used the same right glove as the trigger gesture (Fig. 7C). The user wore glasses for all MultiPoint techniques (Fig. 7D). They tracked the orientation of the head and the nose bridge.

Fig. 7. Marker arrangements: left glove (A), right glove, squeezing (B), right glove for unimanual multipoint and trigger (C), glasses (D), left laser pointer (E), right laser pointer (F).

We also created two laser pointers using wooden dowels and markers (Fig. 7 E and F). To simulate clicking a button on the laser pointer, the user occluded a smaller marker located near the thumb. This allowed for click activation while minimizing cursor jitter in comparison with depressing a physical button.

5. Experiment 1: Single point

In our first experiment, our objective was to measure the speed and accuracy of single point interactions. To do so, we compared the performance of remote perspective-based pointing using three selection techniques against a laser pointer in a selection, drag and docking task. This experiment served as baseline for our main goal, to evaluate remote perspective-based multipoint gestures, accomplished in the second experiment. The design of the experimental task was based on the work of Forlines and Balakrishnan (2008).

5.1. Task

Participants were asked to point to a start location, select the target and drag it to the dock location "as quickly and as accurately as possible". The target was equidistant from the start location and the dock, and randomly located within those constraints (Fig. 8).

Four measures were collected: selection time, selection errors, docking time and docking errors. Selection time reports the time from the start location to the time of successful target selection, while docking time reports the time from successful target selection to the time of successful docking. Selection errors count the number of unsuccessful attempts at selecting the target. Docking errors count the number of unsuccessful attempts at placing the target in the dock.

Only the start location and the docking location were displayed at the beginning of each trial. To start the trial, the participant placed the cursor inside the start location at the center of the top edge of the large display, at which point the target appeared. The goal of the participant was to select and dock the target. A docking was successful if at least 62.5% of the target was placed inside the dock. The target snapped into place when docking was successful, changing the target's color from green to blue.

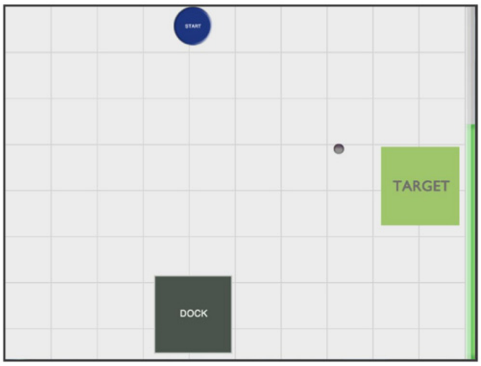

Fig. 8. Sample trial from Experiment 1. The participant begins at the start (blue), acquires the target (green) and drags it to the dock (gray). A progress bar (right) indicates the click state (currently a successful selection). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.2. Design

We used a 4 × 3 × 3 factorial repeated-measures within-subject design. Our variables were interaction technique (remote pointing with squeeze selection, remote pointing with breach selection, remote pointing with trigger selection, and laser pointer), target width (64, 128 and 192 pixels), and target distance (400, 500, and 600 pixels). Each participant performed three trials for each combination of factors, for a total of 108 trials (4 interaction techniques × 3 target widths × 3 target distances × 3 trials). Participants were located two meters from the screen. We randomized the interaction techniques first, then we randomized among target variables (target width, target distance). Each experimental session lasted about 40 min. Participants trained with each interaction technique until they achieved less than 10% improvement between trials.

5.2.1. Preferences

Participants were asked to rate each interaction technique on two criteria: if they were easy to use and if they felt natural to use. The questions were structured using a

5-point Likert scale. Additionally, participants were asked to rank all four single point interaction techniques on their ease of use, then rank which technique they thought allowed for faster task completion.

5.2.2. Participants

12 participants (3 females) between 18 and 30 years old took part in the study. Each subject had some familiarity with multi-touch gestures, e.g., on a smartphone or a laptop. They were paid $10 for their participation.

5.3. Hypothesis

We hypothesized that laser pointing would be preferred over remote perspective-based pointing techniques (H1). This prediction was based on prior work (Jota et al., 2010) that demonstrated that laser pointing results in lower muscular fatigue, as the arm rests against the body instead of being raised in the air.

When comparing each remote selection technique, we expected both the squeeze gesture and the trigger gesture to be faster and more accurate, as well as less mentally demanding, than the breach gesture (H2). We expected this result because the breach gesture requires greater coordination between the selection and pointing actions: the fingers must move along a 2D plane in order to point at a target, and move towards the display to select.

5.4. Results

5.4.1. Performance analysis

We analyzed the four measures collected by performing a repeated measures factorial analysis of variance (ANOVA) using interaction technique (4) × target distance (3) × target width (3) on selection time, docking time, selection errors, and docking errors.

Time Analysis (Fig. 9): For selection time, results show that interaction technique was a significant factor ($F(3,30)=14.206$, $p<0.001$). Pairwise post-hoc tests with Bonferroni corrected comparisons show significance between the breach gesture and every other interaction technique, with the breach gesture being the slowest. We found significant differences for both target distance ($F(2,20)=3.921$, $p<0.05$) and target size ($F(2,20)=25.049$, $p<0.001$).

For docking time, interaction technique was also found to be a significant factor ($F(3,30)=12.726$, $p<0.001$). Pairwise Bonferroni corrected post-hoc comparisons show significance between the breach gesture and the squeeze gesture, as well as the trigger gesture, the breach gesture being significantly slower. Target size ($F(2,20)=17.943$, $p<0.001$) and target distance ($F(2,20)=50.409$, $p<0.001$) were found to be significant factors.

Error Analysis (Fig. 10): We found significant differences between conditions in the target size factor for selection errors ($F(2,20)=13.290$, $p<0.002$). For docking errors, also we found interaction technique to be a significant factor ($F(3,30)=4.490$, $p<0.029$) in addition to target size ($F(2,20)=10.375$, $p<0.002$). However, pairwise Bonferroni corrected post-hoc comparisons did not reveal any differences between specific interaction techniques.

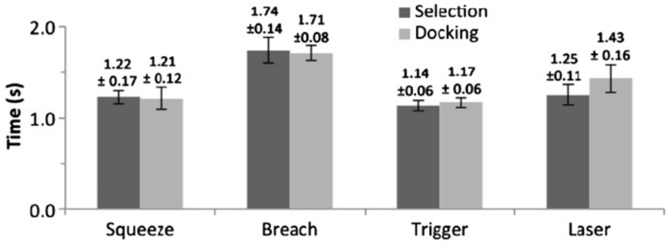

Fig. 9. Mean selection and docking times for the three perspective-based pointing gestures and the laser pointer.

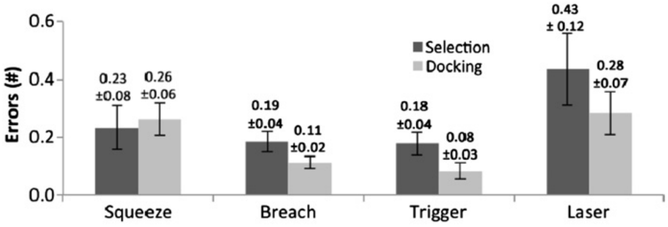

Fig. 10. Mean number of errors for target selection and docking.

5.4.2. Subjective analysis

We found a significant effect of ease of use rankings (Friedman's $\chi^2(3)=9.70$, $p<0.021$), with a preference for remote pointing with trigger selection, followed by the laser pointer, then the squeeze gesture, and with breach having the lowest ranking (Table 1). There was also a significant effect of time completion perception rankings (Friedman's $\chi^2(3)=8.70$, $p<0.034$). Remote pointing with trigger selection was also the highest rated interaction technique on this criterion, with the other three interaction techniques rated in the same order as ease of use.

There was also a significant effect of interaction technique on the ease of use ratings (Friedman's $\chi^2(2)=11.762$, $p<0.003$). Remote pointing with trigger selection had the highest mean rating, above the squeeze, then breach gestures. Similarly, we found a significant effect of interaction technique on ratings of feeling natural (Friedman's $\chi^2(2)=6.950$, $p<0.031$). Again, the remote pointing with trigger selection had the highest mean rating.

5.5. Discussion

The comparison between different interaction techniques for the single point experiment showed significant disparity in temporal performance between the breach gesture and the rest of the techniques. The fastest techniques are, at par, the trigger gesture, the squeeze gesture, and the laser pointer. The breach gesture is the slowest, with significantly higher selection and docking times. Our

Table 1
Cumulative preference ranks for ease of use for each interaction technique for single point.

	Rank			
	First	Second	Third	Fourth
Trigger	5	5	1	1
Laser	5	2	2	3
Squeeze	2	3	6	1
Breach	0	2	3	7

observations indicate that the users were more deliberate, hence slower, with the breach gesture during both target selection and release. This stems from the fact that the breach technique was the only gesture that involved arm movement to select or release the target as opposed to only fingers movements. This confirms our second hypothesis (H2).

Interaction techniques had a significant effect on the number of docking errors. We note that in both type of errors, the trigger gesture had the smallest number of errors and the laser pointer the largest. The high performance of the trigger gesture can be attributed to the minimal movement of the index finger upon activation of the click. The presence of natural hand jitter with the laser pointers interfered with small target acquisition, as pointed out by Myers et al. (2002), resulting in a greater number of selection errors. We observe a similar trend for docking, albeit with fewer cumulative number of errors. We surmise that the effect of hand jitter was reduced due to the margin of error allowed while docking.

It is interesting to note that in a previous comparison among in-air pointing techniques (Jota et al., 2010), the laser pointer was faster than perspective-based pointing for a 1D targeting task. We believe this difference stems from the disparity between the tasks. In a 1D task, hand jitter in the direction perpendicular to the direction of motion is nullified. The 2D nature of our task resulted in the laser pointer performing at par with perspective-based pointing techniques.

From rankings and participant comments, we noted a preference for the trigger gesture, and a dislike for the breach gesture. This is in line with the results reported by Subramanian et al. (2006), where Pin-Through—a technique involving breaching an activation layer to select—recorded lower user preference. While most participants felt that the trigger gesture was the easiest to perform, some mentioned that the squeeze gesture felt more natural. One user remarked that the squeeze gesture was akin to “squeezing the hand as though to grasp an object in real and virtual life”, but another one noted that although “it felt more natural, it was less precise than the trigger”.

When comparing perspective-based pointing against the laser pointer, participants mentioned that using the laser pointers resulted in lower muscular fatigue. We anticipated this, as perspective-based remote pointing requires the index finger to be in the air, between the eyes and the screen. To reach targets in the middle and at the top of the large display, users were required to lift their hand and arm to shoulder levels (or above), which was tiring for users over extended periods of time. Nevertheless, the trigger gesture was preferred by users, and ranked the best both for ease of use and for performance. This result goes against our first hypothesis that stated that the laser pointer would be preferred.

In summary, the competitive temporal performance and lower number of errors for two of the three perspective-based pointing techniques suggest that they can perform at par with laser pointers for single point interactions. These results, combined with user preference for perspective based pointing, prompt us to recommend the trigger gesture for single point interactions.

6. Experiment 2: Multipoint

In our second experiment, we compared the performance of in-air multipoint techniques for both unimanual and bimanual interactions against laser pointers in a standard translate/resize task defined by Forlines and Balakrishnan (2008), adding a 45° rotation of the target to provide a more challenging and realistic abstraction of classic multi-touch photo sorting actions. The goal was to establish whether perspective-based pointing could serve as a viable solution for content manipulation on large displays.

6.1. Task

Before the beginning of each trial, the start and dock locations appeared on the display. The target appeared after the participants placed both cursors inside the start location. Initially, the target was 1.25 times the size of the dock and was rotated 45° counter-clockwise. To dock successfully, each participant was required to rotate, scale and drag (in no particular order) the target inside the dock. The color of the target changed from green to yellow once the rotation and scaling was successful, and to blue once it was correctly docked. Time and error measurements in this experiment were collected identically to those in the first task. Docking was considered successful only if the target was of the correct size and orientation.

6.2. Design

We used a $5x3 \times 3$ factorial repeated-measures within-subject design. Our variables were identical to those in Experiment 1, apart from the interaction techniques. The techniques are as follows:

1. One-handed Point with breach gesture.
2. Two-handed multipoint with squeeze gesture.
3. Two-handed multipoint with breach gesture.
4. Two-handed multipoint with trigger gesture.
5. Two-handed multipoint with laser pointers.

Each participant performed a total of 135 trials (5 interaction techniques × 3 target widths × 3 target distances × 3 trials). Randomization was performed as in Experiment 1. The experimental sessions lasted about 60 min. The participants in this study were the same as the previous experiment. Participants filled out questionnaires similar to the first experiment, this time comparing five interaction techniques instead of four.

6.3. Hypothesis

We hypothesized that all perspective based remote pointing techniques would be faster and more accurate than laser pointers (H3). This prediction was based on the fact that the user needs to compensate for jitter from both laser pointers. In addition, as the user controls two cursors in this condition, we believe perspective based pointing will help the user correlate pointer locations to the corresponding hand. Among the perspective based pointing techniques, we expected unimanual multipoint, using the breach gesture, to be the preferred technique (H4), due both to its similarity to commonly used multi-touch gestures on tabletops and smartphones, and to lower fatigue as the user only has one arm up (Nancel et al., 2011).

6.4. Results

6.4.1. Performance analysis

We performed a repeated measures factorial Analysis of Variance using interaction technique (5) × target distance (3) × target width (3) on selection time, docking time, selection errors, and docking errors.

Time Analysis (Fig. 11): For selection times, results show that interaction technique was a significant factor ($F(4,44)=4.97$, $p<0.013$), in addition to target distance ($F(2,22)=12.61$, $p<0.001$) and target size ($F(2,22)=35.34$, $p<0.001$). Within interaction techniques, pairwise Bonferroni corrected post-hoc analysis showed that bimanual breach was significantly slower than bimanual trigger.

For docking times, results showed interaction technique was a significant factor ($F(4,44)=8.97$, $p<0.001$). Pairwise Bonferroni corrected comparisons identified remote pointing using the trigger gesture as being significantly faster than the laser pointer condition, and the bimanual breach and squeeze gestures, but not unimanual breach. Target size was also found to be a significant factor ($F(2,22)=45.99$, $p<0.001$). We also found an interaction between interaction technique and target size on docking time ($F(8,88)=5.02$, $p<0.013$).

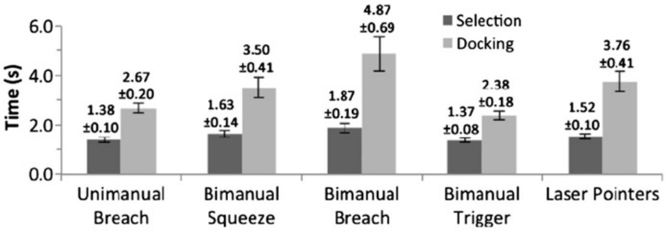

Fig. 11. Mean selection and docking times for the unimanual multipoint remote gesture, the three bimanual multipoint remote gestures and the bimanual laser pointers.

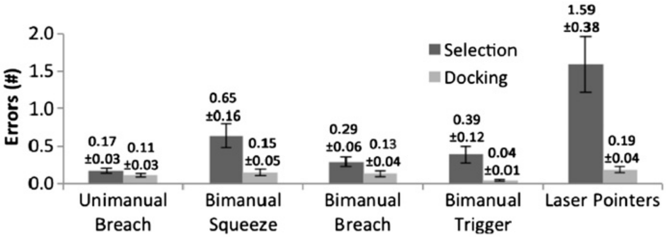

Fig. 12. Mean number of errors for target selection and docking.

Error Analysis (Fig. 12): Results for selection errors showed interaction technique was a significant factor ($F(4,44)=10.08$, $p<0.004$). Pairwise Bonferroni corrected post-hoc comparisons showed significance between the laser pointers and both the squeeze and unimanual gestures, with the laser pointer condition having a larger number of errors. Results showed significance for target distance ($F(2,22)=4.52$, $p<0.029$) and target size ($F(2,22)=18.08$, $p<0.000$). The interaction between interaction technique and target size was also significant ($F(8,88)=6.48$, $p<0.002$).

For docking errors, we only found a significant main effect of target size ($F(2,22)=26.87$, $p<0.001$). However, there was a significant effect of interaction technique by target size ($F(8,88)=3.5$, $p<0.030$).

6.4.2. Subjective analysis

We found a significant effect on rankings of opinions on ease of use (Friedman's $\chi^2(4)=10.80$, $p<0.029$), with unimanual breach and trigger conditions having the highest rankings, followed by the squeeze gesture and laser pointer, with the bimanual breach gesture having significantly lower ranking (Table 2). Likewise, we found a significant effect of participants' rankings of their opinions on which interaction technique allowed faster task completion (Friedman's $\chi^2(4)=10.067$, $p<0.039$). The mean rankings for performance perception are in line with opinions of ease of use.

There was a significant effect of interaction technique on the ease of use ratings (Friedman's $\chi^2(3)=11.972$, $p<0.007$). Remote pointing with unimanual breach gesture had the highest mean rating, above the trigger, squeeze, and the bimanual breach gesture. However, we did not find any significant effect of interaction technique on ratings of feeling natural (Friedman's $\chi^2(3)=7.112$, $p<0.068$).

6.5. Discussion

Our comparison of interaction techniques in the remote multipoint experiment demonstrated significant differences in temporal performance and a discernable disparity in accuracy

Table 2
Cumulative preference ranks for ease of use for each interaction technique for multipoint.

	Rank				
	First	Second	Third	Fourth	Fifth
Unimanual breach	6	0	3	3	0
Bimanual trigger	2	6	3	1	0
Bimanual squeeze	0	3	4	4	1
Laser pointers	3	2	1	1	5
Bimanual breach	1	1	1	3	6

of task completion. Overall, the fastest techniques were the unimanual breach gesture and the bimanual trigger gesture, while the slowest was the bimanual breach gesture.

When comparing selection times between techniques, some results are consistent with the first experiment: the selection times for the bimanual breach gesture were still significantly higher than the bimanual trigger. We observed that all techniques common to Experiment 1, when scaled to their bimanual multipoint equivalents, take at least 20% more time to select the target, with one exception: the bimanual breach gesture. This is in line with a pattern observed in user strategies for this task: most users preferred to place both cursors inside the target at the start of a trial, anticipating an easier transition into multipoint manipulations. The result of this preemptive action was an increase in selection times for bimanual trigger, squeeze and laser pointers. However, this strategy was rarely executed with the bimanual breach gesture, possibly due to the effort and dexterity involved. As a result, selection strategy, and hence selection times, remained constant between experiments for the bimanual breach technique.

Analysis of docking time indicates that the trigger gesture performed significantly better than the bimanual squeeze, the bimanual breach and the twin laser pointers. While this deviated from our third hypothesis, in that we expected all perspective-based multipoint techniques to perform comparably, this underlines the ease of use afforded by the trigger gesture for bimanual multipoint manipulations.

In contrast, the performance of unimanual multipoint was in line with this hypothesis. We observe that the unimanual breach gesture performs well for in-air remote multipoint. Compared to the trigger gesture, we found only a 12% difference in mean docking time, a non-significant difference compared to the next fastest technique (the squeeze gesture, 47% higher). This technique is fast despite it requiring the user to be particularly deliberate while releasing the target due to the breach gesture. It is evident that the unimanual technique allows the user to rotate the target around the wrist while simultaneously performing a scaling gesture using the fingertips and translating with the arm. The trade-off between faster resize and rotate options and slower selection and release operations results in performance that is at par with the trigger gesture (with its faster selection and release but with slower resize and rotate operations due to arm movement). Many users mentioned that the unimanual technique was "easy and efficient" and was preferred among all multipoint techniques.

The bimanual laser pointers accounted for the largest number of selection and docking errors, recording as many selection errors as all other perspective-based techniques combined. The reason for this can again be traced to a user preference of placing both cursors inside the target for concurrent selection to immediately enable multipoint manipulation. In some cases, this resulted in an error for each hand if the target was not acquired.

Overall, results from Experiment 2 confirm our fourth hypothesis: the unimanual condition is preferred. This technique outperformed bimanual laser pointers with temporal performance on a par with the trigger gesture. In addition, the unimanual technique recorded the lowest number of errors overall. Since this is the only gesture allowing for one handed multipoint, along with strong performance, we recommend the unimanual gesture for use in the design of remote multipoint systems for large displays.

The visual feedback provided in both experiments requires further investigation. Some users commented on the progress bars' purely utilitarian function, and how having feedback located in their periphery was at times confusing or unsatisfactory. This may have caused additional errors, although this increase should be proportional for all techniques as the feedback was uniform. In addition, in techniques using the breach selection, the clicking gesture provides no inherent physical feedback, unlike squeezing or pressing a button with laser pointers.

7. Conclusions

In this paper, we presented MultiPoint, a set of perspective-based interaction techniques for large displays. We discussed a number of perspective-based interaction techniques, including the squeeze gesture and the breach gesture. We empirically compared performance of these two in-air techniques with the trigger gesture, and laser pointing, in both single and multipoint interactions. The trigger gesture for single point conditions and the unimanual breach gesture for multipoint conditions were preferred, and were among the fastest for their respective experiment. Laser pointing obtained mixed results: in the single point experiment, it was a fast technique but obtained a large number of errors; in the multipoint experiment, it obtained the lowest ranking and performance.

Overall, MultiPoint techniques have been shown to be effective for interacting with graphical objects on a large display from a distance. Consequently, we believe that design of remote interaction techniques can be informed by the results of our evaluation. For exclusively single-point use cases, perspective-based pointing using the trigger gesture would be suitable. Perspective-based pointing invites casual walk-up-and-use; it is device-less, provides a cohesive mental model of pointing, and is more accurate. For multipoint scenarios, the unimanual breach is recommended due to lower fatigue levels resulting from the use of a single arm, and the higher accuracy it affords for affine transformations.

7.1. Future work and limitations

Unimanual gestures bring remote multipoint interactions to new scenarios, including meeting presentation systems and artistic performances (Banerjee et al., 2011). We believe the use of such gestures requires further investigation. For instance, it may be interesting to explore additional selection gestures for unimanual multipoint. It would be worthwhile to examine such gestures performed sitting down, simulating accessing a display from a desk during a meeting, as well as in conjunction with an interactive tabletop. We would also like to extend this work to collaborative situations, where multiple users could perform remote multipoint gestures on large displays at once.

Finally, it is important to note that currently available marker-less computer vision based tracking solutions, such as the Microsoft Kinect, do not have the fidelity to consistently support all the interaction techniques presented in this paper. Thus, the current work required the use of retro-reflective markers on gloves and glasses to perform an empirical evaluation. To fully realize the potential of these interaction techniques, it is essential that future embodiments include marker-less systems that allow users to apply these techniques unencumbered by gloves or glasses—thus becoming truly device-less.

Acknowledgments

Funding provided by SMART Technologies Inc., Ontario Centres of Excellence, Natural Sciences and Engineering Council Canada.

References

Ball, R., North, C., 2007. Realizing embodied interaction for visual analytics through large displays. Computers and Graphics 31 (3), 380–400.

Banerjee, A., Burstyn, J., Girouard, A., Vertegaal, R., 2011. WaveForm: remote video blending for VJs using in-air multi-touch gestures. In: Proceedings of the CHI Extended Abstracts. 1–6.

Baudisch, P., Sinclair, M. Wilson, Andrew D., 2007. Soap: a pointing and gaming device for the living room and anywhere else. In: Proceedings of the SIGGRAPH Emerging Technologies. 17.

Bi, X., Balakrishnan, R., 2009. Comparing usage of a large high-resolution display to single or dual desktop displays for daily work. Proceedings of the CHI. 1005–1014.

Bolt, R., 1980. "Put-that-there": Voice and gesture at the graphics interface. In: Proceedings of the SIGGRAPH. 262–270.

Boring, S., Baur, D., Butz, A., Gustafson, S., Baudisch, P., 2010. Touch projector: mobile interaction through video. In: Proceedings of the CHI, 2287–2296.

Bowman, D., Wingrave, C., Campbell, J., Ly, V., 2001. Using pinch gloves for both natural and abstract interaction techniques in virtual environments. In: Proceedings of the HCII. 629–633.

Brignull, H., Rogers, Y., 2003. Enticing people to interact with large public displays in public spaces. In: Proceedings of the INTERACT. 17–24.

Cao, X., Balakrishnan, R., 2003. VisionWand: interaction techniques for large displays using a passive wand tracked in 3D. In: Proceedings of the UIST. 173–182.

Dietz, P., Leigh, D., 2001. DiamondTouch: a multi-user touch technology. In: Proceedings of the UIST. 219–226.

Forlines, C., Balakrishnan, R., 2008. Evaluating tactile feedback and direct vs. indirect stylus input in pointing and crossing selection tasks. In: Proceedings of the CHI. 1563–1572.

Grossman, T., Wigdor, D., Balakrishnan, R., 2004. Multi-finger gestural interaction with 3D volumetric displays. In: Proceedings of the UIST. 61–70.

Johanson, B., Hutchins, G., Winograd, T., 2002. PointRight: experience with flexible input redirection in interactive workspaces. In: Proceedings of the UIST. 227–234.

Jota, R., Nacenta, M.A., Jorge, J., Carpendale, S., Greenberg, S., 2010. A comparison of ray pointing techniques for very large displays. In: Proceedings of the GI. 269–276.

Kendon, A., 2004. Gesture: visible action as utterance. Cambridge University Press, Cambridge, United Kingdom.

Khan, A., Matejka, J., Fitzmaurice, G., Kurtenbach, G., 2005. Spotlight: directing users? Attention on large displays. In: Proceedings of the CHI. 791–798.

Malik, S., Laszlo, J., 2004. Visual touchpad: a two-handed gestural input device. In: Proceedings of the ICMI 289–296.

Myers, B., Bhatnagar, R., Nichols, J., Peck, C.H., Kong, D., Miller, R., et al., 2002. Interacting at a distance: measuring the performance of laser pointers and other devices. Proceedings of the CHI. 33–40.

Nancel, M., Wagner, J., Pietriga, E., Chapuis, O., Mackay, W., 2011. Mid-air pan-and-zoom on wall-sized displays. In: Proceedings of the CHI. 177–186.

Oakley, I., Sunwoo, J., Cho, I.-Y., 2008. Pointing with fingers, hands and arms for wearable computing. In: Proceedings of the CHI Extended Abstracts. 3255–3260.

Oblong Industries. 〈http://www.oblong.com〉. (accessed 03.04.11).

Peltonen, P., Kurvinen, E., Salovaara, A., Jacucci, G., Ilmonen, T., Evans, J., et al., 2008. It's Mine, Don't Touch!: interactions at a large multi-touch display in a city centre. In: Proceedings of the CHI. 1285–1294.

Pierce, J., Forsberg, A., Conway, M., Hong, S., Zeleznik, R.C., Mine, M.R., 1997. Image plane interaction techniques in 3D immersive environments. In: Proceedings of the Interactive 3D Graphics. 39–43.

Qin, Y., Shi, Y., Jiang, H., 2010. Structured laser pointer: enabling wrist-rolling movements as a new interactive dimension. In: Proceedings of the AVI. 163–166.

Rekimoto, J., 2002. SmartSkin: an infrastructure for freehand manipulation on interactive surfaces. In: Proceedings of the CHI. 113–120.

Ringel, M., Berg, H., Jin, Y., Winograd, T., 2001. Barehands: implement-free interaction with a wall-mounted display. In: Proceedings of the CHI Extended Abstracts. 367–368.

Shoemaker, G., Tang, A., Booth, K.S., 2007. Shadow reaching: a new perspective on interaction for large wall displays. In: Proceedings of the UIST. 53–56.

Streitz, N.A., Geißler, J., Holmer, T., Konomi, S., Müller-Tomfelde, C., Reischl, W., et al., 1999. i-LAND: an interactive landscape for creativity and innovation. In: Proceedings of the CHI. 120–127.

Subramanian S., Aliakseyeu D., Lucero A., 2006. Multi-layer interaction for digital tables. In: Proceedings of the UIST. 269–272.

Tate, A., Chen-Burger, Y.H., Dalton, J., Potter, S., Richardson, D., Stader, J., et al., 2010. I-Room: a virtual space for intelligent interaction. IEEE Intelligent Systems 25 (4), 62–71.

Vogel, D., Balakrishnan, R., 2005. Distant freehand pointing and clicking on very large, high resolution displays. In: Proceedings of the UIST. 33–42.

Wienss, C., Nikitin, I., Goebbels, G., Troche, K., 2006. Sceptre: an infrared laser tracking system for virtual environments. In: Proceedings of the VRST. 45–50.

Wilson, A.D., 2004. TouchLight: an imaging touch screen and display for gesture-based interaction. In: Proceedings of the ICML. 69–76.

Wilson, A.D., 2006. Robust computer vision-based detection of pinching for one and two-handed gesture input. In: Proceedings of the UIST. 255–258.

Zigelbaum, J., Browning, A., Leithinger, D., Bau, O., Ishii, H., 2010. G-stalt: a chirocentric, spatiotemporal, and telekinetic gestural interface. In: Proceedings of the TEI. 261–264.