Touching the Void: Direct-Touch Interaction for Intangible Displays

Li-Wei Chan[2] **Hui-Shan Kao**[1] **Mike Y. Chen**[2] **Ming-Sui Lee**[1,2]
Jane Hsu[1,2] **Yi-Ping Hung**[1,2]
[1]Department of Computer Science and Information Engineering
[2]Graduate Institute of Networking and Multimedia
National Taiwan University
{mikechen, mslee, yjhsu, hung}@csie.ntu.edu.tw

ABSTRACT
In this paper, we explore the challenges in applying and investigate methodologies to improve direct-touch interaction on intangible displays. Direct-touch interaction simplifies object manipulation, because it combines the input and display into a single integrated interface. While traditional tangible display-based direct-touch technology is commonplace, similar direct-touch interaction within an intangible display paradigm presents many challenges. Given the lack of tactile feedback, direct-touch interaction on an intangible display may show poor performance even on the simplest of target acquisition tasks. In order to study this problem, we have created a prototype of an intangible display. In the initial study, we collected user discrepancy data corresponding to the interpretation of 3D location of targets shown on our intangible display. The result showed that participants performed poorly in determining the z-coordinate of the targets and were imprecise in their execution of screen touches within the system. Thirty percent of positioning operations showed errors larger than 30mm from the actual surface. This finding triggered our interest to design a second study, in which we quantified task time in the presence of visual and audio feedback. The pseudo-shadow visual feedback was shown to be helpful both in improving user performance and satisfaction.

Author Keywords
Direct-touch interaction, virtual panel, intangible display.

ACM Classification Keywords
H.5.2 Information interfaces and presentation: User Interfaces; Graphical user interfaces.

General Terms
Design, Human Factors.

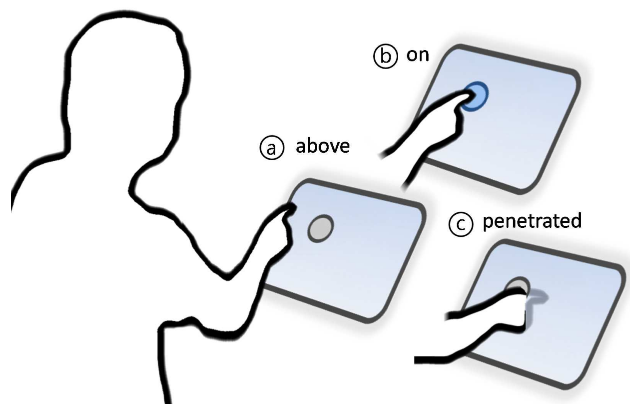


Figure 1. Three states of a direct-touch operation for intangible displays: (a) above the surface, (b) on the surface (c) penetrating the surface. The penetrable property is the main distinctive feature of the intangible display with respect to the tangible display.

INTRODUCTION

This paper explores direct-touch interaction for intangible displays. In recent years, direct-touch display devices have become more widespread, with devices available in varying form factors ranging from palm-size devices such as Apple's iPhone to tabletop displays such as Microsoft's Surface Computer. Several research and commercial intangible displays have also been introduced, offering displays which appear in mid-air. These displays appear to be touchable, but when touched provide no tactile feedback.

The notion of intangible displays has been around for decades, take the instantiations of such displays we have witnessed on Star Trek. Imagine that, in the near future, we could achieve this dream, that anyone could initiate an intangible display anywhere in front of their eyes by simply performing an action via gesture or voice. The elderly could effortlessly manipulate the displays without using physical input devices. Moreover, direct-touch intangible displays do not suffer from the hygiene issues of their tangible counterparts. For example, shared touch-displays, such as kiosks and ATM machines, increase the likelihood of the communication of infectious disease due to direct-contact cross contamination. Intangible displays are also an ideal solution in environments such as kitchens and factories, where it is easy for the tangible displays to become greasy and dirty.

In our prior work, we developed an intangible display, Virtual Panel [5], based on image formation by use of a Fresnel lens. The Fresnel lens functions as a convex lens which transmits images from one side of the lens, to appear in mid-air on the other side. Facing the lens, an intangible planar display is seen in front of the user's eyes. Touch detection for the display was achieved via two infrared cameras. We attached a water ripple effect as the visual feedback on each touch point detected. Two simple interactive games were implemented. Users completed the games by simply performing several touches.

In an official exhibition of novel display technology, we were invited to present this prior work, and had the opportunity to observe a large number of user reactions. We were surprised by the wide range of user performance levels. The main causes for low performance levels could be due to users' unfamiliarity with intangible displays in general, or the awkward feeling of "touching" a mid-air display. Based on our observations, more than half of the users could correctly finish the games after about one-minute of training, but others took longer and still others did not improve even after training. These observations strengthened our resolve to find the root cause of the performance discrepancy among users, and we began to think of ways we could help users by adding supportive feedback to the display.

In this work, we are interested in exploring the use of direct-touch manipulation as a preliminary focus point, and in particular how visual and audio feedback affect user performance on direct-touch manipulation for intangible displays. If with the support of some appropriate feedback to the user, the user can more easily manipulate intangible displays, then the previous results of touch-based interaction studies can be readily applied.

In our initial study, we collected data associated with user's ability to interpret the 3D location of targets shown on our intangible display prototype. We found that among the participants, performance was very inconsistent in determining the z-coordinate of the targets, and that participants reported uncertainty about the success of their touches in the questionnaire. Strikingly, more than 30 percent of positioning operations made errors in the z-coordinate larger than 30mm. In the second study, we quantified task time with respect to different conditions of visual and audio feedback. According to the second study, the pseudo-shadow visual feedback was helpful both in improving task time and user satisfaction. Audio feedback was effective as well, but was more distracting according to subjective responses. We also received a variety of preferences from participants with respect to each feedback condition. Based on these findings, we also define a series of guidelines for the design of intangible displays.

RELATED WORK

Intangible Display Techniques

There are already several intangible displays available in commercial products. We classify these displays into two categories. The first category includes intangible displays which display true 3D space. Fogscreen [13] produces a thin curtain of dry fog that serves as a translucent projection screen, enabling projected images in free space. Based on the same principle, HelioDisplay creates a mid-air rear projection display by emitting condensed air. However, to our knowledge, these teams have not provided user study results on their displays and the applied interactions. The second category includes intangible displays that only appear directly in front of the user. iBall [10] creates a real image in air by transmitting an LCD display via a Fresnel lens. The real image is then encompassed by a transparent ball to give the illusion of a spherical display. Chan et. al. proposes Magic Crystal Ball [4] based on the same display mechanism, and investigates touch-based interactions that users can directly use with their bare hands. Intangible displays made by the Fresnel lens are only able to produce 2D planar images, or in the case of iBall and Magic Crystal Ball, warped transformations thereof. While stereoscopic or autostereoscopic displays, which produce 3D images that can be viewed by the users with or without the use of special glasses, can also be classified in this category. We focus on the former Fresnel lens-based intangible display technology in this work.

It is worth mentioning again that one of the unique features of our intangible display is that only 2D planar images can be displayed, in comparison with other intangible displays which provide true depth perception. This is an important distinction, as our study results in this paper may not apply directly to situations where users can rely on their true depth cues [8][7].

Manipulation of Intangible Objects

As we have previously stated, the main problem in interacting with intangible objects is the inherent lack of tactile feedback. Related work can be found in the field of Virtual Reality (VR). In addition to the aforementioned techniques which create displays which literally float in mid-air, most VR systems create the illusion of virtual world by generating a pair of images, one for each eye. This stereoscopic imagery provides a true 3D image so virtual objects appear to float in front of and behind the physical display surface. In order to improve object selection and manipulation in the virtual world, researchers have proposed and evaluated a variety of 3D interaction techniques [15][3]. One category of such techniques applies indirect manipulation such as *ray-casting* [15][16] to facilitate the access of distant objects. Other techniques [12][19][14] enhance the presence of human body parts in the virtual world, simulating a paradigm closer to direct manipulation.

Instead of being immersed in a virtual environment, however, within our intangible display prototype environment, a flat intangible display is seen in front of the user's eyes. Moreover, users can touch intangible virtual objects via use of their physical hands rather than via virtual substitutes.

Use of Shadows as a Depth Cue

There are several bodies of research which investigate the use of shadows as depth cues to enhance illusions of direct-manipulation in interactive applications. Kenneth [9] proposes using shadows for visualizing the spatial relationships

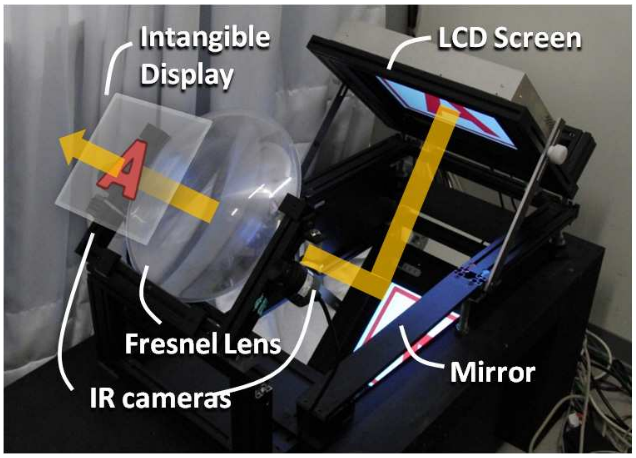


Figure 2. The prototype system of our intangible display, annotated for clarity. The Fresnel Lens transmits the image on the LCD monitor into the air. The yellow arrow represents the optic path of image formation.

between objects in three-dimensional graphics applications. The shadows are also used as interfaces for manipulating objects. Wanger [21] further investigates the effect of shadow sharpness and shadow shape on the perception of spatial relationships. Shadow Reaching [18] employs a perspective-projected shadow of the user on the display. The technique was designed in interactions over large distances. LucidTouch [22] introduces generating shadows of the user's hands onto the screen to improve the user's depth perception performance.

Eye-hand Coordination

Eye-hand coordination refers to synergy of hand movement control with visual feedback. This ability is important for operating with intangible displays, because users have to perform touches which do not rely on tactile feedback. Roland and Colin [1] report that stereoscopic perception and force feedback improve user performance in tabbing tasks. Andrea et. al. [11] ask users to perform a grasp gesture toward target objects in conditions where visual and haptic feedback was either present or absent. Prablanc et. al. [17] and Evan et. al. [6] report pointing errors in situations with or without vision of the physical hands. Daniel et. al. [23] report touch errors for under the table interactions where vision of the operating hands was occluded. This error for back-side interaction was recapped in Baudisch's work [2] on small displays. For intangible displays, however, it is unclear whether the error is consistent with previous reports. In our first study, we have conducted an experiment which records the accuracy across participants in performing target positioning tasks on our intangible display.

THE INTANGIBLE DISPLAY DEVICE

Our prototype of an intangible display is created by image formation using a Fresnel lens. In order to study direct-touch interaction with intangible displays, we implemented touch detection by use of stereo vision techniques with two infrared cameras.

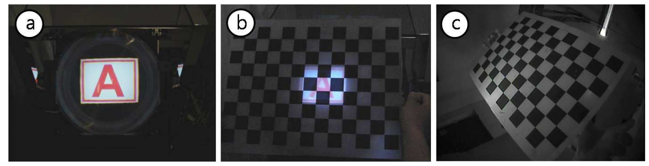


Figure 3. The translucent checker-board pattern helps the users and the cameras identify the position of the intangible display during the calibration process. (a) the LCD monitor displays an image 'A'. (b) users adjust the pattern sheet in order to see the 'A' image clearly. (c) the checker-board pattern seen by the cameras.

Hardware Configuration

The architecture of our prototype system can be divided into the display module and the detection module as shown in Fig. 2. For the display module, we use a Fresnel lens which is able to concentrate beams to form a real image in the air. The optical path of the display is as follows: the displayed content in the LCD screen is first reflected by the mirror, penetrates the Fresnel lens, and then forms a real image in the air. In this work, we use a 17 inch LCD screen which provides a display resolution of 1280 by 1024 pixels, and an 11-inch Fresnel lens with 8.2 inch focal length. The distances from the screen to the lens is 20.5 inches which produces a real image in the air at 13.66 inches away from the optical center of the lens. The detection module consists of two infrared cameras coupled with infrared illuminators attached on either side of the Fresnel lens. The cameras are used to detect touch positions on the intangible display surface.

Note that due to the nature of the intangible display design used in this paper, the display suffers from a restricted range of viewing angles. In other words, users can only perceive a full and undistorted view of the displayed content if they remain within this viewing angle range. Although we expect this limitation in viewing angle to be solved for future intangible displays, this limitation does not affect our present study. To achieve this, we make sure each participant in our study perceives the correct presentation of the displayed content before they begin the user tests.

System Calibration

We apply homographic transformation techniques to compute planar coordinate mappings between the two cameras, the intangible display, and the LCD screen coordinate systems. Homographic transformation provide mappings of two planar coordinate systems by at least four corresponding points collected for computing a homographic matrix. This technique, also applied in TouchLight [24], warps the two camera views onto the projection surface for detecting touch points.

Direct-touch Detection

Figure 4 shows the process of our touch detection approach. After removing the lens distortion of the left and right camera views, we warped the two views onto the display coordinates of the intangible display. Since the two cameras are attached with narrow-angle infrared illuminators, only the foreground objects presented in the overlap of the two illuminations will show brightly in the camera views. There-

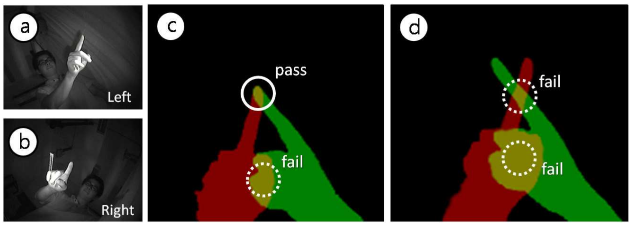


Figure 4. Direct-touch detection for our intangible display. (a)(b) The camera images after removing lens distortion. (c) After image rectification, foreground regions of the two camera images are extracted for analyzing finger touches. (d) The user's finger penetrates the display surface and is not detected as a valid touch.

fore, the intersection of the foreground regions of the two warped views can represent where the touch points take place.

Currently the intersections could be the fingertip, the middle part of a finger, or palm. For each intersection component, we further recognize touch points. For each center of an intersection component, we collect two intensity profiles along with a circle on each of the two warped images. Both profiles are analyzed according to the following procedure. If there is a single segment on the profile, and the length of the segment is smaller than a preset value, then the segment is labeled valid. The center of an intersection component is reported as a direct-touch finger point, if both profiles of the intersection are labeled valid. This idea is based on the smaller structure of fingertips comparing to a palm. If an intersection observes fingertip structures in both camera views, the intersection is reported as a finger touch. Otherwise, the intersection could be the middle part of a finger or lower part of a palm. This approach effectively takes the intersections as regions of interest for direct-touch detection, and can achieve real-time performance.

Resulting Size
Due to the Fresnel lens, the produced floating image suffers from slight distortion in its border region. In order to optimize the display quality, we masked out the border region of the lens by displaying black pixels in that region, narrowing down the display area to two-thirds of its original size. The resulting size of the display seen by the viewers is about 4 x 3 inches.

USER STUDY 1: DEPTH DISCREPANCY
In this study, we examined users' ability to correctly determine the 3D locations of targets by stereoscopic cues without tactile feedback support. Our main hypothesis was that users would be able to accurately determine target positions with varying degrees of success, however no clear correlation between touch confidence and the resulting errors would be found.

Task
Participants performed a target positioning task. For each trial, the system performed an on-screen five second countdown before target display began. Each target appeared for 3 seconds and then disappeared automatically, followed by a 2-second black screen. The size of target was 1.5 inch by 1.5 inch, slightly larger than the width of normal fingers. When the target appeared, participants were instructed to smoothly and steadily bring a single finger into virtual contact with the surface of the target. Participants were also instructed to withdraw their hands and rest when the target disappeared. At the moment the target disappeared, we captured images of the two infrared cameras where we assumed participants' fingers were touching the intangible target. During the entire positioning procedure, the target did not provide any feedback to the participants.

Before the task began, we attached a circular marker to the tip of the participant's finger responsible for performing positioning. The use of the marker allowed us to calculate the accurate 3D position of the pointing fingertip using stereo triangulation.

Experimental Design
The study design was 9 target positions with 3 repetitions for each cell. Target positions were the 9 centroids of a regular 3 x 3 grid. Target positions were randomized. Prior to the study, users filled out a background questionnaire, and received up-front training with at least 5 trials until they were familiar with the entire process. After the study, they were asked to provide subjective feedback. Using a 7-point Likert scale, participants rated their level of confidence that their touches were correctly positioned on the targets.

After the entire study, we calculated the 3D positions of the pointing fingertips. The positions of circular markers in the two camera images were manually labeled in order to obtain reliable and accurate 3D positions. For each trial we recorded errors of the pointing positions in the x, y, and z axes with respect to the ground-truth target positions.

Here, we report the positional accuracy of our fingertip localization procedure. Since the errors in stereo triangulation are mainly from the corresponding points, we simulate the effect of random error on the labeled positions using Gaussian with 1, 3 and 5 pixel distances, and report the average errors of 1.8, 3.7 and 6.3 mm, respectively. As the marker radius in camera images is not more than 3 pixels, the manual labeling assures the positional errors below 5 mm.

Apparatus
The experiment was run on our prototype intangible display. The operating system used was Windows XP and the study was implemented in C++ and OpenCV library. The valid viewing angles of the intangible display were 50 and 38 degrees in the horizontal and vertical viewing directions, respectively. Participants sat on a chair without wheels. Prior to the study, participants were guided to obtain a correct view of the intangible display, and were told the limitations of the display. After participants adjusted their posture and the chair position, the chair helped to stabilize their viewing position during the study.

Participants
We recruited 39 participants (25 male) between ages of 21 and 36. Thirty-one were graduate students; eight were un-

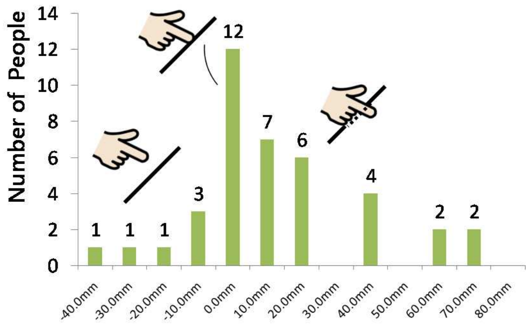


Figure 5. Histogram showing the number of participants vs. average error.

	x	y	z
Average (mm)	2.41	2.46	13.25
Std (mm)	1.12	1.08	31.01

Table 1. Average distance and standard deviation from participants' touches to the target positions in the x, y, and z axis.

dergraduates. All participants reported at least 5 hours of computer usage daily. Nine were using touch screen devices daily; twenty-seven reported occasional use; while three had no experience with touch screens. All but three were right-handed. All participants had normal or corrected-to-normal vision.

Results

The findings of this study are organized in the following table and charts.

Table 1 shows errors aggregated across all participants in the x, y, and z axes with respect to the target positions. As expected, errors in z axis were obviously larger than errors in x and y axes. This supports our observations in a pilot study that, the lack of tactile feedback made participants uncertain in determining at what depth to stop their fingertips in the air.

Figure 5 shows a wide difference across participants in locating the z-coordinate of the targets. The result shows a right-skewed distribution, which reveals the fact that participants were inclined to pass through the display surface. One reason for this could be that before reaching the display surface, participants could easily locate the targets based on perceived stereo cues. In the XY plane, performing a touch on the intangible display is similar to the tangible touch screen. This helps to explain why errors in the x and y axes were relatively small. Once the surface was penetrated, however, participants no longer had a point of reference to indicate the depth at which to stop their fingers.

Figure 6 aggregates the targeting data from all participants. The curve shows the percentages of touches located within a given displacement in z-coordinate away from the actual surface. The data shows that about 70 percent of touches performed in mid-air achieved less than a 3-centimeter error along the z axis. The result suggests that the applied

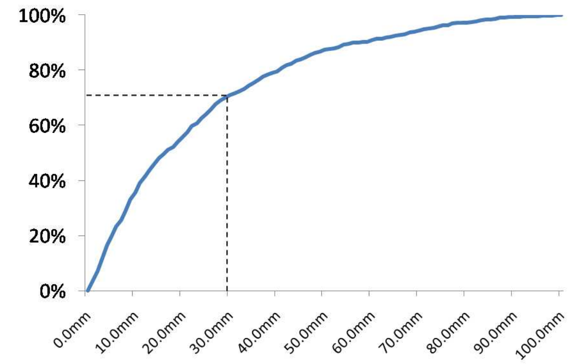


Figure 6. Percentage of touches that are located within a given displacement along the z axis.

touch detection approach should provide sufficient tolerance to minimize positioning errors across users. On the other hand, this also suggests an estimate of the error rate that a narrow detection zone approach would offer. The estimation, however, should be viewed as providing a rough upper bound for the error rate, since the applied interaction techniques can greatly affect performance.

Figure 7 shows the confidence level ratings from participants related to their touches performed in this study. The average confidence level was 5.3. There is no clear correlation between the rated confidence and the resulting errors. Confident participants did not correlate with smaller errors. And more interestingly, confident participants often made larger errors than more conservative participants.

Subjective Opinions

In the following, we report some difficulties and tricks of achieving mid-air touches, described by participants after the study.

"When I was trying to touch the target, I saw that the target stayed behind my fingertip. So I reached out my finger further, but the target was still shown as behind it. I reached out further still until my arm was fully extended."

When passing through the display surface, fingers occlude the light (Figure 8). As a result, targets were always showed

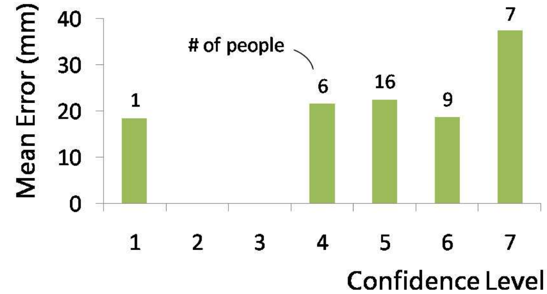


Figure 7. Errors (mm) made by participants in different levels of confidence.

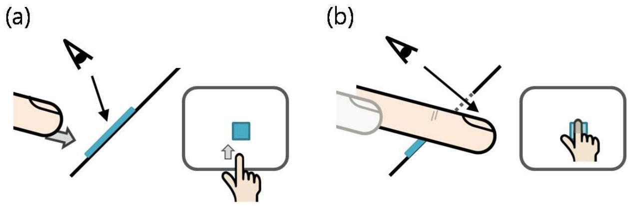


Figure 8. Passing through the intangible display causes a double vision problem. (a) Before reaching the surface, the user's focus remains on the target. (b) The users' focus is transferred to the fingertip after the finger penetrates the display surface.

behind the fingers. One could consider it a limitation for Fresnel lens-based intangible displays. However, this is not a limitation for intangible displays based on projecting light toward the human eye, such as the increasingly popular stereoscopic 3D displays with goggles or auto-stereoscopic 3D displays. For an ideal realization of the intangible display, when passing through the display surface, fingers should sink into the display surface and users should still see the content above the finger. In practice, however, thus functionality is not yet achievable at the present time. Therefore we shall consider the limitation more of a problem to conquer than a defect to inventory.

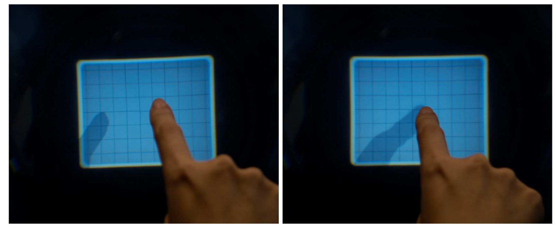
"I saw the content split up into two counterparts, which made me feel uncomfortable and a little dizzy in the head."

Participants started receiving a visual conflict, named double vision, also known as diplopia, when their fingers penetrated the display surface by too far a margin. We use Figure 8 to describe this problem. Before reaching the target, participants were focusing on the target for which they stretched out their hands. Once their fingers passed through the surface, the focus transferred to the fingertip because the target which was just in focus was now occluded by the fingers. If the fingers reached still further, the increasing displacement between the surface and the focus of vision system (on the fingertip) made the vision system receive simultaneous perception of two identical images of the display content. At this time, participants may become uncomfortable. Also, this problem also holds across intangible displays made by projecting lights from opposite directions towards the human eye.

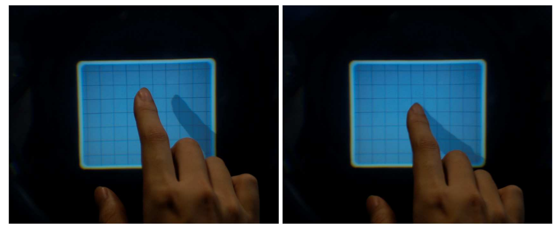
"I reached out my finger and stopped at a position where I found that the finger just barely occluded the target. I did this based on my visual experience of touching a button in the real world."

Some participants performed the positioning tasks using visual tricks from similar experiences in the real world. Most reported that they saw the target and reached to the target by instinct.

It is worth mentioning that, in the experiments, we provided a grid background as shown in Figure 9. A rich-textured background may help to alleviate some of the effects of occlusion within the display environment.



(a) Pseudo-shadow from Non-dominant side



(b) Pseudo-shadow from Dominant side

Figure 9. Pseudo-shadow effect: the user is operating with her right hand. The shadow helps visualize the proximity of her fingers to the intangible display.

We consider that the errors made by the participants in the experiment could initially be due to the difference in participants' ability to judge exact distance, and further affected by the cognition conflict of fingers often occluding buttons, and at last the lack of tactile feedback.

In summary, the study supports our observations in our pilot study that participants showed inconsistent accuracy when it came to localizing targets shown in mid-air and also showed a lack of confidence in the success of their touches. These findings drove our interest to investigate whether supportive feedback could assist participants in using intangible displays, to further help them improve direct-manipulation performance while reducing their uncertainty.

USER STUDY 2: COMPARING SUPPORTIVE FEEDBACK

Participants performed a target acquisition task. The purpose of the study was to investigate whether visual and audio feedback helped improve user performance in target acquisition while reducing their uncertainty of touching in mid-air.

Pseudo-shadow

One of our objectives was to help users with no previous experience operate an intangible display. We integrated a pseudo-shadow effect on the intangible display, suggesting a physical surface captured the hand shadow. The pseudo-shadow was produced from rectifications of the two infrared cameras. We extracted the hand region in the rectified image, and displayed the region in a semi-transparent gray color, imitating the shadow of the user's hand projected onto the intangible display. In Figure 9, two types of shadows, one on either side of the hand, were created respectively. Because the camera images were rectified with respect to the display space of the intangible display, the pseudo-shadow would behave like a real shadow. When the users' fingers reached the surface, the pseudo-shadow of the fingers on the screen would also reach the real-world fingers.

The pseudo-shadow provides excellent discoverability which greatly simplified the learning phase. Users were only instructed to raise their hand and smoothly reach to the display surface. They could effortlessly understand what was meant by a mid-air display. Our initial plan was to study the effectiveness of the three types of pseudo-shadows, produced by mimicking a light source placed at above-left, above, and above-right of the intangible display, suggesting real shadows from right, below, and left of the hands, respectively. Pilot studies, however, showed that the pseudo-shadow from below the hand was useless and distracting as the shadow, when included, was mostly occluded by the real hands. As a result, the dominant-side shadow and non-dominant-side shadow were the two types of visual feedback used in our second study.

Interfaces

There were four interface conditions. Two of which were adding pseudo-shadow as supportive visual feedback, one was using audio feedback, and the other was without supportive feedback.

The *Dominant-side shadow* interface simulates a pseudo-shadow image attached on the dominant-hand side of the user.

The *Non-dominant side shadow* interface simulates a pseudo-shadow image attached on the non-dominant-hand side of the user (Figure 9). The pseudo-shadow images were generated at 30 frames per second. Participants could easily recognize that the pseudo-shadow corresponded to their hand.

The *Audio* interface played a short non-speech audio sound whenever participants touched the surface of the intangible display. If participants had fingers hovering above or passing through the display surface, there was no sound feedback rendered. In other words, if participants hovered their fingers on the surface, the sound was repeatedly played.

The *Baseline* interface was not supported by feedback, and was served as a baseline performance in this study.

Task

Participants performed a target acquisition task. For each trial, the system performed an onscreen five second countdown, and then started to display the target (blue circular button).

The target was implemented with a dwell time interface. Participants were instructed to place single finger on the target shown in mid-air. Unlike the first study, targets in this study responded to touches by expanding a yellow outer circle enclosing the target, as shown in Figure 10. Once the expanding yellow circle enclosed the target, the target turned red, prompting participants to release their finger to complete the click operation. The expanding duration was set to 2-second dwell time. During the expanding duration, if participants released their fingers, the expanding restarted.

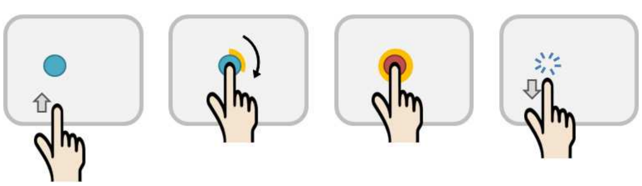


Figure 10. Dwell-click acquisition of a target.

In User Study 1, we observed that users had difficulty in determining the exact distance from the intangible surface. Straightforward penetration without dwelling might induce unintended target acquisitions. To avoid this situation, we used the dwelling-and-select interface as a tool for measurement rather than as an intended interaction requirement. The target size was 1.5 inch by 1.5 inch, purposely chosen larger than average finger size to avoid the fat finger problem [20].

Several existing intangible display technologies, including ours and auto-stereoscopic displays, suffer from the double-vision effect that happens when users' fingers over-penetrate the display surface. In User Study 1, several participants reported this issue. In order to control the effect of double vision on the study results, we defined a target zone in which users could comfortably perform target acquisition without feeling the effects of double vision.

Experimental Design

The design of the study was 4 x 9 [Interface x Target Position] with 3 repetitions. Target positions were the 9 centroids of a regular 3 x 3 grid. For each trial, we recorded task completion time. Interface was a within-subject variable, the interface order was counterbalanced, and the target positions were randomized.

Prior to the study, participants trained on all interface conditions, each time performing at least 3 target acquisitions. After each condition, they were asked to provide subjective feedback on the condition just used. After the study, participants reported ranks of overall satisfaction with regards to all conditions.

We had considered that the interface order could influence user performance. Because the intangible display was shown at an identical position in mid-air, within a single condition session, participants could build familiarity from previous successful target acquisitions via muscle memory and posture, or via referencing nearby physical objects. To alleviate these effects, participants were required to leave the experiment room and take at least a 5-minute break before continuing on to the next condition. At the start of each condition, participants were only informed of the type of up-coming condition, and no additional training was given.

Apparatus

The hardware setting was same as in User Study 1. In this study, we performed real-time touch detection. The radius parameter in the touch detection determined the available range of the touch-zone along the z-axis. The parameter was set such that a 2-centimeter thick touch-zone was established between 0.0mm to 20.0mm. In other words, the detection re-

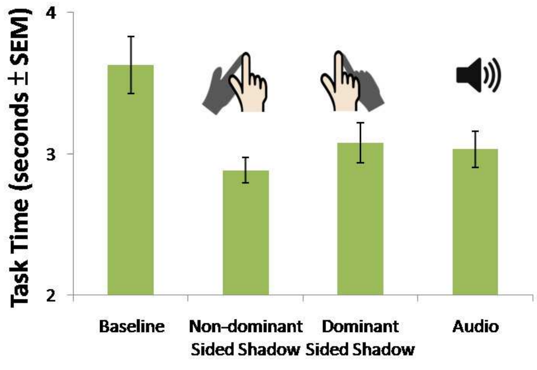


Figure 11. Task times for the *Baseline*, *Dominant-side Shadow*, *Non-dominant-side Shadow*, and *Audio* conditions.

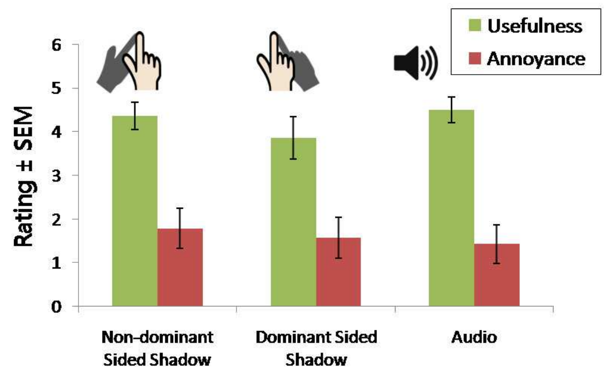


Figure 12. Subjective usefulness and disturbance ratings from participants for three feedback conditions.

ported touch positions in x and y coordinates only when the fingers stayed within the touch-zone. Hovering above the zone or passing beyond the zone would not produce touch events.

Participants

16 participants were recruited, 8 male, between ages of 21 and 32. Seven were graduate students; nine were undergraduates. All participants were familiar with computers. Fourteen use touch screen devices daily; two reported occasional use. All but two were right-handed.

Hypotheses

There were three hypotheses: (1) we expect that the three conditions with supportive feedback will be more effective than the *Baseline* condition. (2) Participants will find the *Non-dominant-side shadow* condition more distracting than the *Dominant-side shadow* condition. (3) Participants will find the *Audio* condition more distracting than the two pseudo-shadow conditions.

Results

Figure 11 shows the completion time with respect to interface condition. We performed a repeated measures ANOVA evaluating the within-subjects effects of interface type on completing time. A significant main effect of interface $(F = 14.3, p < 0.0005)$ was found.

H1: Paired-sample t-test between interface conditions was performed. The difference between the three interface conditions with supportive feedback with respective to the *Baseline* condition was significant. The *Dominant-side shadow* and the *Non-dominant-side shadow* conditions were both more effective than the *Baseline* condition $(p < 0.0001, p < 0.00001)$. The *Audio* condition was also more effective than the *Baseline* condition $(p < 0.0008)$. This supports our hypothesis, that supportive feedback improved user performance.

The difference between the *Non-dominant-side shadow* and *Dominant-side shadow* conditions was borderline significance $(p < 0.03)$. The benefit of the *Audio* condition in the presence of the *Non-dominant* and the *Dominant-side shadow* conditions was not significant.

Figure 12 shows subjective ratings in usefulness and distraction for the three feedback conditions. The result suggests that all feedback was useful. No clear preference, however, between feedback conditions was found across participants. From subjective opinions, we had found that participants' views toward specific feedback conditions varied widely. We report some interesting findings later in *Subjective Opinions*.

H2, H3: We applied the Wilcoxon test on ranks of distraction. *Non-dominant-side shadow* was not significantly distracting compared to the *Dominant-side shadow* ($\chi^2 = 7.6364$, $\alpha = .17, p > \alpha$). *Audio* was also not more distracting than two pseudo shadow conditions ($\chi^2 = 6.8869, \alpha = .22, p > \alpha$) ($\chi^2 = 6.4843, \alpha = .26, p > \alpha$).

Somewhat surprisingly, participants rated little distraction for all feedback conditions. Initially we hypothesized that the audio would be more distracting, and non-dominant sided shadow would be more distracting than its counterpart. The result, however, does not support these hypotheses. The reason for this could be the short period use of the conditions and the fact that, in general, participants thought the feedback to be helpful. To gain more insight about any distraction caused by the feedback itself, we asked participants an additional question of whether they agree that feedback should be added to products of intangible displays as a permanent feature. A majority of participants reported that they would like to add it as a feature, but that it should also be able to be toggled off.

In the study, we had observed that participants quickly became familiar with the display, or more specifically, with the presence of a display in 3D. The difference in display use performance across participants might relate to their hand-eye coordination ability. For most participants, after completing a few successful acquisitions, participants were able to perform faster for the remaining acquisitions. With the help of feedback, not only did participants perform faster than in the previous acquisitions, but were also more stable in the remaining ones. The observation can be supported in the measure of overall task completion time, where the standard error for the *Baseline* condition was larger than other conditions with built-in feedback.

Subjective Opinions

In the following, we report some findings from the subjective opinions gathered from users.

"I like the shadow added from my right hand side [Dominant-handed side]. Because it looks like when I was writing on desk, I usually like to put the lamp on the opposite side [to my hand] so the shadow won't occlude anything."

The participant described that dominant-side shadow was intuitive since it echoed his past experience of writing on a desk.

"I would prefer that the shadow came from my left hand side [Non-dominant-handed side]. Since the approaching of the shadow finger to my fingertip provided strong reference of proximity to the surface"

The participant reported *Non-dominant side shadow* was better because it provided obvious reference compared to its counterpart from the dominant side. The pseudo-shadow effect provided great discoverability in that it continuously reveals the proximity of the real finger to the intangible display surface. For *Dominant sided shadow*, the sense of approaching was weaker because a large part of the shadow was occluded by the hand. In comparison, *Non-dominant side shadow* suffered from less occlusion problems since its shadow was cast from the opposite side of the hand. *Non-dominant side shadow*, however, has the drawback that it occludes more content.

It is interesting that for some participants who after minimal practice performed very well, all feedback soon become useless. As a result, feedback which offered less intrusion was preferred as this would be less distracting. On the other hand, some participants did not as readily pick up usage of the display, and they reported that some form of applied feedback was always helpful and never distracting.

As far as the *Audio* feedback is concerned, some participants reported that audio feedback was a complementary source of feedback with visual feedback, which made them feel confident that the touch just performed was successful. This was partly due to the fact that the target itself provided visual feedback for the dwell click already; the pseudo-shadows could increase the burden of vision.

DISCUSSION

In the following, we propose design guidelines for designers of intangible displays. The guidelines shall hold for general implementations of intangible displays such as immersive virtual environments and stereoscopic 3D displays.

Continuous Feedback vs. Discrete Feedback

In study 2, pseudo-shadow feedback smoothly and continuously visualizes states of users' hands against the surface, while audio feedback discretely reacts only when users touch the surface. Discrete feedback clearly confirms a touch by playing a sound effect or an instant visual change, but provides little help before users touch the surface. On the other hand, continuous feedback offers references during the entire process, but is ambiguous for signaling when a touch has occurred. User interfaces for intangible displays shall combine the two types of feedback to improve usability. However, designers should also consider that continuous feedback can easily become distracting to users.

Distraction Associated with Feedback

In some cases, continuous feedback is more distracting than discrete feedback. In the study, we found that participants could operate with the intangible display quite well after several successful touches. However, most participants revealed awkwardness at the beginning of each new condition. Continuous feedback helped them to recapture the familiarity, but after that, the feedback again become unnecessary. Designers should use continuous feedback very carefully in order not to distract their users. One simple approach would be to provide continuous feedback only in the first few minutes of use of intangible displays for each session. A more sophisticated approach would be to analyze the behavior of touches over time, and provide feedback whenever abnormal operations occurred.

Penetration Avoidance

One challenge for intangible display technology is the penetrable property of the intangible display surface. In the total absence of feedback, display manipulation performance is low for most users. Penetrating the surface causes the problem of double vision, which is described in the discussion of the Study 1. Pseudo-shadow feedback provides continuous interpretation of the proximity of the hand to the surface. However, if users, under any circumstance, penetrated the surface, pseudo-shadow is not able to provide appropriate indications. For this situation, audio feedback provides a clearer indication to penetration by simply not playing a sound effect. Audio feedback, however, leads to more severe environmental restrictions. To indicate display penetration to users, alternative feedback may be applied. For example, globally darkening the intensity of the entire screen image could indicate that the display is being penetrated; or locally deforming or hollowing out the content where the penetration occurred on the display could also be used as an indication as well.

CONCLUSION

In this paper we explored how to deploy direct-touch interaction for intangible displays. Unlike tangible displays, intangible displays suffer from lack of tactile feedback. With the absence of tactile feedback, direct-touch interaction has been shown to perform poorly even for simple target acquisition tasks. In order to investigate this problem, we have conducted two user studies. In the first study, we have reported that the large difference in positioning accuracy across users could cause failure of traditional direct-touch interaction. The second study suggests pseudo-shadow and audio feedback are effective in improving user performance while reducing their uncertainty.

The main difference between intangible displays and their counterparts is the penetrable property of the display surface. This paper unveiled this basic problem, and describes user reactions about mid-air display interaction with the supports of continuous and discrete feedback. As future work, we are planning to investigate different touch-based interactions such as land-on and take-off techniques for intangible displays. We also plan on exploring visual and audio designs to reveal to users the level of penetration so as to help them stay away or recover from any problems.

ACKNOWLEDGEMENTS

This work was supported in part by the National Science Council, Taiwan, under grants NSC 97-3114-E-002-002, and by the Excellent Research Projects of National Taiwan University, under grants 98R0062-04.

REFERENCES

1. R. Arsenault and C. Ware. Eye-hand coordination with force feedback. In *Proc. of CHI'00*, pages 408–414, 2000.
2. P. Baudisch and G. Chu. Back-of-device interaction allows creating very small touch devices. In *Proc. of CHI'09*, pages 1923–1932, 2009.
3. D. A. Bowman, D. B. Johnson, and L. F. Hodges. Testbed evaluation of virtual environment interaction techniques. *Presence: Teleoper. Virtual Environ.*, 10(1):75–95, 2001.
4. L.-W. Chan, Y.-F. Chuang, M.-C. Yu, Y.-L. Chao, M.-S. Lee, Y.-P. Hung, and J. Hsu. Gesture-based interaction for a magic crystal ball. In *Proc. of VRST07*, pages 157–164, 2007.
5. L.-W. Chan, T.-T. Hu, J.-Y. Lin, Y.-P. Hung, and J. Hsu. On top of tabletop: A virtual touch panel display. In *Proc. of TABLETOP'08*, pages 169–176, Oct. 2008.
6. E. D. Graham and C. L. MacKenzie. Physical versus virtual pointing. In *Proc. of CHI'96*, pages 292–299, 1996.
7. T. Grossman and R. Balakrishnan. The design and evaluation of selection techniques for 3d volumetric displays. In *Proc. of UIST06*, pages 3–12, 2006.
8. T. Grossman and R. Balakrishnan. An evaluation of depth perception on volumetric displays. In *Proc. of AVI'06*, pages 193–200, 2006.
9. K. P. Herndon, R. C. Zeleznik, D. C. Robbins, D. B. Conner, S. S. Snibbe, and A. van Dam. Interactive shadows. In *Proc. of UIST92*, pages 1–6, 1992.
10. H. Ikeda, H. Naemura, T.and Harashima, and J. Ishikawa. i-ball: Interactive information display like a crystal ball. In *SIGGRAPH '01 Emerging Technologies*, page 122, 2001.
11. A. H. Mason, M. A. Walji, E. J. Lee, and C. L. MacKenzie. Reaching movements to augmented and graphic objects in virtual environments. In *Proc. of CHI'01*, pages 426–433, 2001.
12. M. R. Mine, F. P. Brooks, Jr., and C. H. Sequin. Moving objects in space: exploiting proprioception in virtual environment interaction. In *Proc. of SIGGRAPH'97*, pages 19–26, 1997.
13. A. Olwal, S. DiVerdi, N. Candussi, I. Rakkolainen, and T. Hollerer. An immaterial, dual-sided display system with 3d interaction. In *Proc. of VR'06*, pages 279–280, 2006.
14. J. S. Pierce, A. S. Forsberg, M. J. Conway, S. Hong, R. C. Zeleznik, and M. R. Mine. Image plane interaction techniques in 3d immersive environments. In *Proc. of SI3D'97*, 1997.
15. I. Pouprey, S. Weghorst, M. Billunghurst, and T. Ichikawa. Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques. *Computer Graphics Forum*, 17(3):41, 1998.
16. I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The go-go interaction technique: non-linear mapping for direct manipulation in vr. In *Proc. of UIST96*, pages 79–80, 1996.
17. C. Prablanc, J. E. Echallier, M. Jeannerod, and E. Komilis. Optimal response of eye and hand motor systems in pointing at a visual target. *Biological Cybernetics*, 35(3):183–187, 1979.
18. G. Shoemaker, A. Tang, and K. S. Booth. Shadow reaching: a new perspective on interaction for large displays. In *Proc. of UIST07*, pages 53–56, 2007.
19. R. Stoakley, M. J. Conway, R. Pausch, and Y. Pausch. Virtual reality on a wim: Interactive worlds in miniature. In *Prof. of CHI'95*, pages 265–272, 1995.
20. D. Vogel and P. Baudisch. Shift: a technique for operating pen-based interfaces using touch. In *Proc. of CHI'07*, pages 657–666, 2007.
21. L. Wanger. The effect of shadow quality on the perception of spatial relationships in computer generated imagery. In *Proc. of SI3D'92*, pages 39–42, 1992.
22. D. Wigdor, C. Forlines, P. Baudisch, J. Barnwell, and C. Shen. Lucid touch: a see-through mobile device. In *Proc. of UIST07*, pages 269–278, 2007.
23. D. Wigdor, D. Leigh, C. Forlines, S. Shipman, J. Barnwell, R. Balakrishnan, and C. Shen. Under the table interaction. In *Proc. of UIST06*, pages 259–268, 2006.
24. A. D. Wilson. Touchlight: an imaging touch screen and display for gesture-based interaction. In *Proc. of ICMI'04*, pages 69–76, 2004.