Human-Computer Interaction -- INTERACT'03
M. Rauterberg et al. (Eds.)
Published by IOS Press, (c) IFIP, 2003, pp. 519-526

Embedding Interactions in a Retail Store Environment: The Design and Lessons Learned

Noi Sukaviriya, Mark Podlaseck, Rick Kjeldsen, Anthony Levas, Gopal Pingali, Claudio Pinhanez

IBM TJ Watson Research, 19 Skyline Drive, Hawthorne, NY 10532, USA
{noi, podlasec, fcmk, levas, gpingali, pinhanez}@us.ibm.com

Abstract: We developed a steerable interface system that can direct graphical displays to any desirable locations, can capture interactions at any desirable locations, and can track user locations in a 3-dimensional space. This paper discusses a retail store application where we applied this set of advanced technologies to brings more information to users in the shopping context. The paper presents the design challenges for this new interaction paradigm and reports the findings from our design walk-through sessions with users.

Keywords: Augmented reality, projected interfaces, vision-based recognition, user tracking.

1 Introduction

Current alternatives to desktop computing explore the issues of ubiquity, natural interfaces, and context sensitivity (Abowd and Mynatt, 2000). One approach to computing beyond the desktop is to project text and imagery into physical environments and then to detect user interactions with vision systems and/or sensors. Research prototypes that predated our work (Wellner, 1991; Koike et al, 2000; Patten, 2001; Jacob et al, 2002; Piper et al, 2002, Rekimoto, 2002) relied on fixed location projections while some of these prototypes (Raskar & Low, 2001) achieved multiple-location projections through deploying and coordinating multiple projectors.

The Everywhere Displays projector (Pinhanez et al, 2001) directs graphical displays to various surfaces in a given environment. Steerable vision (Kjeldsen et al, 2002) and user tracking (Pingali et al, 2002) enable the development of ambient computing environments built with a single projector and coordinated vision components.

Previously, we demonstrated a prototype of the Everywhere Display at the SIGGRAPH 2001 exhibit (Kjeldsen et al, 2002). This application guided users in extracting M&M candies from buckets and placing them at appropriate spots on a pre-calculated color matrix composition of a well-known painting. Users followed instructions projected on tables, buckets, and the composition matrix projected on a flat table. The application was used by over 600 users during a period of 6 days – each user picked up 10 candies; four "paintings" were completed.

The success of our previous prototype – in particular, how users managed multiple interactions in multiple locations – led us to believe that the technology may drive more sophisticated applications. One of the areas we believe ubiquitous information may be useful is in retail.

In this paper, we first give an overview of a retail store application and discuss the underlying technology. Then we focus on the design characteristics of the application, the challenges in designing this type of applications, and preliminary results from design walk-through studies.

2 Retail Store Application

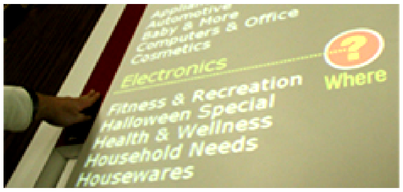

Figure 1. Product directory table.

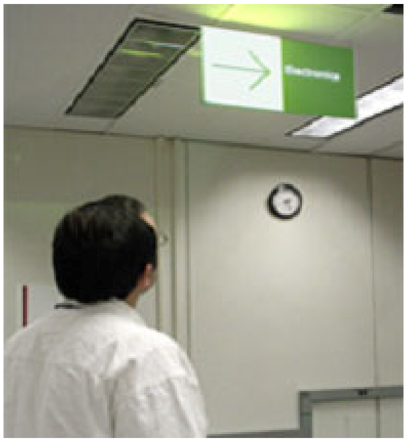

Figure 2. User follows signs for direction.

To explore different design concepts before tackling a real store, we replicated a 16x11 foot (5x3.3m) retail space in our laboratory. Three different sets of interactions were developed: product directories, interactive shelves (clothing bins) and a mixed-media product table.

Interactive product directories are designed to assist users with locating merchandise throughout the store. A product directory, projected on a table, is located at the entrance of the store. The user can move a physical slider on the left of the table to find a product as illustrated in Figure 1. Consequently, the product list located to the right of the slider scrolls up and down to mirror the motions of the slider. Once the user touches the "where" symbol, arrows pointing to the location of the highlighted product are projected on signage boards hanging from the ceiling as illustrated in Figure 2. Low cost

Figure 3. Product directory on the wall.

Figure 4. User examines information related to the pants in the bin in front of him.

paper versions of this interface are hung though out the store, as shown in Figure 3.

A second set of interactions uses an arrangement of bins containing women's pants as shown in Figure 4. When the user stands at a considerable distance from these bins, a circulating series of advertisements for women's clothing are projected on the panels, aiming to attract customers to the merchandise area. When the user approaches the bins, the display is changed to advertise the store's credit card and special promotions. Once the user starts examining pants in a bin, information about the pants is displayed in the proximity of the bin. Additional interactions can be called on to check the size chart or the availability of the stock.

Figure 5 shows the third set of interactions, a mixed-media table for Halloween products such as books, CD-ROMs, DVDs, and videotapes. While the user walks around the table, product information

Figure 5. Halloween products table.

appears highlighting products related to different user categories – witches, ghosts, vampires, and cats.

3 Technology behind the Scene

In order to support interactions anywhere, our system needs up to three major components to work together – the Everywhere Displays projector, the steerable vision system and the user tracking component. Depending on the demand and the requirements, an application may use these three system components singly, independently, or in synchrony.

Figure 6 shows our current prototype of the Everywhere Displays (ED) projector. The projector comprises of a 3000 lumen LCD projector, a computer-controlled pan/tilt mirror, and a pan/tilt/zoom camera. The 3000 lumen projector is sufficiently bright for typical indoor conditions. The pan/tilt mirror deflects the light from the projector to any surfaces in a space and is computer-controlled to serve the on-demand projection requested from an application. In order to eliminate distorted images due to oblique projection – projection with angles other than perpendicular to the projected surface – each graphical image is computationally pre-warped

Figure 6. Prototype of the ED-projector.

to compensate for the distortion before being sent to the projector (see detail in Pinhanez, 2001).

The pan-tilt camera attached to the ED projector feeds information to the steerable vision interface system. While the camera can be directed independently to view any area within a room, it is generally directed towards a projected graphical display. Our method to detect user activity involves the interpretation of the path of a user's moving hand. This method is preferred over the more common static background subtraction techniques since the dynamically projected images can significantly alter appearances of both the background and foreground objects. Motion is detected by looking at differences between adjacent images in the video stream. Hands are tracked by looking for fingertips in the filtered motion data. Candidate fingertips are identified based on shapes, and evaluated using spatial heuristics such as orientation, distance from the user's body and the previous fingertip locations.

To build an interface, application designers assemble "widgets" that specify areas where users can interact. Each type of widget responds differently to user's motions. A touch widget looks for an in-pause-back motion of the hand, while a continuous motion reports the relative location of a fingertip in one or two dimensions. The system can track objects other than users' hands. For example, in case of the slider, the physical slider on the product directory table is tracked.

The user tracking component determines the 3D position of a user in real time by analyzing the video feed from multiple static cameras. In principle, we determine a user in each frame in a camera view as the foreground regions by subtracting the previous frame, or by subtracting a pre-captured image of the background (Horprasert et al, 1999) – the latter approach yields better segmentation results but is confounded by projected displays appearing as foreground regions. We have developed a geometric masking technique to eliminate the projected displays from the scene based on the fact that they appear on planar surfaces relative to the background scene. Finally, a user's head is detected and tracked based on camera calibration parameters (Zhang, 2000) to determine the 3D position, hence the height of the user, with respect to a geometric model of the space. The accuracy of user position tracking is within a few inches. Currently, we do not deal effectively with multiple users occluding each other therefore have limited our application to single user tracking.

5 The Application Design

Although retail stores are starting to use lighting fixtures to project promotional patterns on walls and floors (Derksen, USA), the goal of our research is to explore how the full power of computing and interactive graphic displays can be used ubiquitously in a shopping context.

6 Design Context

We have looked at shelves and displays in various kinds of stores as a basis for our initial design. While higher-end stores appear to have more suitable display space with less merchandise in a comparable square footage, aesthetics and maintaining brand separation are key in such spaces. Lower-end stores tend to maximize shelf space; they often clutter the space, encourage competition among brands, and make searching more challenging. This poses different challenging design requirements altogether.

Our design guidelines that are unique to retail space and in-context interactions are:

- To make information appear as naturally and as serendipitously as possible. We want to keep the shopping experience the same when shoppers do not need the information and enhance the experience when additional information is needed. We do not require shoppers to interact with the system to start the process. Instead, we came up with ways that user's natural actions would trigger the interaction raising the possibility that false detection may come with such decisions.
- To display information as close to the user as possible, utilizing suitable space around her. Designers have to define suitable surfaces as part of the design solutions.
- To maximize the available display space, bringing out as much information as possible to minimize further information requests. We did not want to turn shopping into interactive computer sessions.

With these guidelines in mind, we attempted to meet the lower-end store requirements. The product displays we chose – the clothing bins and the mixed-media table – represent the typical cluttered display of multiple but related products. We also attempted to use only the potentially available space in our store replica such as signage, table space, floor space around a table, wall space, and shelf space. There are also a number of design innovations we experimented with in the design solution. One was to see whether interesting events can be triggered by utilizing user positions in the store. Another was to utilize the ability of the system to enable same interactions in various locations in different scales at different times. We have incorporated some of these innovations into the current design.

6.1 The Conceptual Design

As mentioned in Section 2, three sets of interaction designs were completed so far – product directories, interactive clothing bins and the mixed-media Halloween product table. Figure 7 shows the diagram of how these product areas are located in our store replica. The diagram also shows where our equipment is located.

Product Directory: We intended the product directory to be a multiple-scale interface that the user can call on throughout the store. A product directory table, a white table with a physical red slider on the left, is located at the entrance of the store. A smaller, non-moving replica of this interface is used in other areas of the store. To activate the product directory, the user touches the table and a list of products is displayed to the right of the slider.

We contradict our first guideline here of "natural interactions" by having users initiating the interactions for the following reason. While the user tracking system can detect where the user is at all times, the user will not always want a product directory at her finfertips. There is no other way to determine when the directory is needed unless the user makes the call. Our design solution designated areas for directories in multiple places in the store, visually identified as such by foam-core boards on walls. As the user moves around in the store, she can touch one of these boards to initiate the directory display.

For the product directory table, the user interacts with the red slider to scroll up and down the product list. Once the user touches the "Where" symbol, arrows pointing to the highlighted product are shown in sequence. The first arrow is displayed on the far end (from the user) of the table; the rest of the sequence is displayed on signage boards hung from the ceiling, starting from the one closest to the user to the one closest to the products. We assumed that

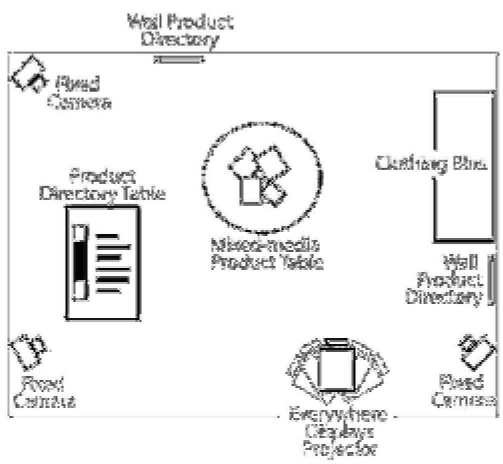

Figure 7. Floor Diagram of the Store Replica.

the user can see the sequence of the displays and can see the directions of how to get there from here.

Interactive Clothing Bins: Our clothing bins are assembled in a checkerboard pattern of alternating merchandise and displayable space. Each storage bin contains a type of women's pants with variations in colors and sizes. As we do not detect the user touching individual pants, a user's hand in the bin for a predetermined period of time is used as a trigger that there is an interest in pants in the bins. This causes information about the pants to be displayed. Since we want to minimize further interactions, we used multiple bin surfaces for information about each type of pants. Figure 8 shows the design of the pant information display. The storage bin with the focus is highlighted with a yellow box around its rim. The name and the description of the pants are displayed on the nearest bin surfaces surrounding the pants. A color chart fills the next closest bin. Touch icons for sizes, available stock, and customer service are displayed on the third closest bin. As the user browses pants in other bins, the display changes to reflect information relevant to each new product. (A critical assumption we make is that a user will understand that all the displayed information relates to the pants in the highlighted bin.)

Mixed-media Products Table: We placed a number of Halloween-related media products such as DVDs, CDs, and books on a round table, leaving a 6 inch "margin" of exposed surface around the edge of the table. User position is the primary interaction medium for the table. By shifting position, the user identifies herself with a demographic group. A display projected on the margins of the table points out products that might be of special interest to that particular group. Hoping to stimulate the imaginations of users and subjects, we chose cats, witches, ghosts, and vampires as our demographic groups. Four positions were marked around the table with footprints representing each group.

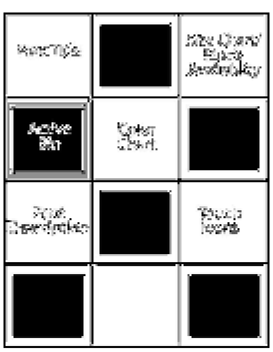

Figure 8. Example Design Patterns of the Bin Display.

It was intended that this application would function as a self-personalizing display; rather than being identified by the system through biometrics or badges, the user identifies herself through her position. Furthermore, by shifting positions, she could shift perspectives on the merchandise and see it, in effect, through the eyes of other kinds of shoppers. Thus, cats might locate themselves to view items of interest to themselves, but they might also want to change positions to see what ghosts were reading and watching.

7 Design Walk-through

7.1 The Evaluation Procedure

We ran four design walk-through sessions with internal users from our research labs. Two of the subjects were male and two were female. Subjects were told that they would be exploring designs in the context of a retail store. They were told that they would see information display appearing on different surfaces and that some might support further interactions.

Each subject went through the design of four different product areas – product directory, interactive clothing bin, mixed-media product table, and one other application (not described in this paper). Subjects were asked to verbalize what they saw, how they understood what they saw, their actions and sources of confusion. All subjects went through one product area at a time and were asked to summarize and critique the experience in each product area. A series of design checkpoints was compiled as a guide to walk the subjects through the design. Since the subjects were aware that they were not there to shop for any of the products, some task guidance was given periodically to move the subjects through all aspects of the design. Audio of the walk-through sessions was recorded. The designers stayed outside of the designated store area and talked to the subjects from a distance because of the possible interference to the vision and the user tracking systems, both of which can accommodate only a single user at the moment.

6.2 Analysis

After running about 4 sessions of design walk-through, users' comments started to repeat so we stopped the evaluation. Subjects were comfortable with information appearing right next to them on various surfaces. Information appearing as a consequence of their explicit actions, for example touching on the size chart symbol, or touching to initiate product directories, was clearly understood The product directories, in particular, is a case where no physical product is present, hence anticipation for something to appear is expected. On the contrary, the intended "natural interactions," walking around the product table and putting a hand in a pant bin, frequently caused confusion and disconnection between the interactions and the appearing results.

While a number of problems were spotted, a small number of them could be categorized as purely design problems. A majority of these design problems came from mismatches between design assumptions and the actual user behaviour. Yet another portion of these problems could be categorized as common design problems anticipated in this new sensing-based paradigm as pointed out in (Belloti, et al, 2002).

6.3 Mismatches between Design Assumptions and User Reactions

Subjects were not readily aware of their positions being part of the interactions: We initially thought that using user's position to trigger information changes would be more obvious and visible. Around the Halloween mixed-media product table, although subjects noticed that information changed over time, they could not figure out what caused the changes. Only one subject, after walking around the table for quite some time, started to assume such a possibility but was not conclusive about it. Similarly in the clothing bin studies, subjects approaching the bins noticed changing information but none had any ideas of why it happened. This problem has been reported before in the context of an interactive environment for children (Bobick et al., 1999). Nevertheless, in our case, the subjects did not find the apparent lack of a cause for the changes as a source of concern. Their perception was that the environment was a dynamic one where information may change at any time.

Subjects did not relate the information projected on multiple bins as related to single product in one bin: The clothing bin information design attempted to maximize display space and minimize further interactions. We intended the design to spread out information on one bin of pants over several bins. On the contrary, most subjects thought the information was general information for all pants in the bins and were not aware of the connection between their hands in a bin and the information change. One subject even thought that the detailed description of one item described all of the items in the various bins. This perception seemed to change after subjects were asked to discuss information on each bin face; until this point, they did not carefully read any of the information. None of the subjects saw the box highlighting the bin with which they most recently interacted. After they were asked to look closely at the information in all bin faces, two subjects perceived the yellow highlight box and understood its purpose. However, when asked why such information showed up, none of the subjects could associate their previous interactions with the highlighted bin and the information.

Subjects expected to touch and get responses after being familiar with the concept: We intentionally made all subjects go through product directories first as a way to familiarize them with the idea of projected interfaces and in situ interactions. While some subjects were a bit more reserved at the beginning, all of them eventually appeared rather comfortable with touching projected images. For better projection, we avoided 3D cues to indicate that images were touchable. As a result, some subjects had trouble discerning what could and could not be touched. For example, two subjects touched the stock availability and color charts, expecting that the system would highlight the pants in the bins of specified size and the color.

Subjects did not easily follow a sequence of displays in disconnected locations: As described in the design section, product directions are displayed as a sequence of arrows, starting from the far end of the directory table, then to signage boards hung from the ceiling. All but one subject could see the first directional sign on the far end of the table. None of the subjects saw the subsequent arrows. Two subjects, both males, physically followed the very first arrow by immediately walking to the far-end of the table where it was displayed. They both started looking for the product at that location.

It is evident that subjects did not visually find a sequence of displays as anticipated in the design, even if it related to directions. Utilization of this particular display technique merits careful consideration for visual search factors and visual connections within the user's field of view.

6.4 Common Problems related to Sensing Environments

Subjects did not easily relate their actions to projected information: It was clear that subjects could not relate their actions, be it walking around or putting their hands in the bins, to the change of display. It is possible that the subjects were engaged in activities that drew their attention elsewhere and that they therefore missed the display switching point.

Subjects were not aware why information changed: While the previous issue relates to subjects not recognizing the trigger of change, this issue emphasizes the disassociation between information at hand and the contextual object. It is evident from the walk-through that subjects could not recognize the associations between pants in a particular bin and information about them, nor the association between the product information on the table and the product classification of each table quarter. One explanation would be that subjects were not familiar with the new interaction paradigm and that they simply did not conceptualize these new types of triggers. We designed the system assuming the user can navigate to information they might want, yet the user has to be able to conceptualize these triggers as tools for navigation. We have yet to achieve t his design goal.

Inexplicable disappearance: One problem very specific to our technology is exactly what our steerable interface system was built for – moving displays from one location to another. In our designs, we implicitly assume that the user's visual attention matches the relocation of the displays. More often than not, this is not the reality. Subjects often perceived as mysterious the disappearance of what they were seeing in front of them. We need to address this issue better if we were to succeed in helping users form the correct conceptual model of the information presented to them.

Lack of immediate feedback and slow response on the interactions contributed to problematic interactions: The current environment is by no means a very fast and responsive one. The design was limited in terms of what can be done to provide immediate feedback because of delays in the computer vision detection system, in the projector hardware, and sometimes even because of network issues. Traditional feedback such as providing a visual cue where the user is pointing or a visual cue that a touch action is successful can be slow; our current design provided no feedback so as to avoid confusion from untimely feedback. We strongly suspect that such lack of immediate feedback contributed to hardships in conceptualizing the relationships between actions and emerging display. Interestingly, throughout our walkthrough, we left the audio trace from the vision system on; the system beeped when it successfully detected a touch by the user. At the conclusion of the walk-through, two subjects mentioned that they took the audio signal as a cue that they had successfully touched an image and would wait for something to happen.

Missing interactions and false positive detection: Because the vision system is not operating with 100% accuracy, there were times when the subjects' interactions were not detected or an interaction was detected where the subjects did not intend. Subjects were not aware of their success or failure (except when they heard the audio beep and discerned its meaning). In this walk-through, designers were present to ask subjects to repeat their interaction when the system failed to detect it. We recognize that error handling is crucial and it is obvious from this walk-through that the design was not ready to be on its own without a strategy to handle these type of errors.

7 Conclusion and Future Work

In this retail store design, many assumptions have been made so a prototype could be built for further exploration with real users. These assumptions have now been evaluated, and challenged by our results. The mismatches in our design assumptions and user reactions were in line with those questions raised in Bellotti et al, 2002. The lesson we learned was that we designers have taken the advancement of user interface technologies for granted. Our design thinking has changed to deal with more complex issues. When we deal with new innovative technology such as the one used in our application, we focused on advance issues but overlooked simpler issues such as intermediate feedback to allow users to learn and follow the interactions. This is utterly important in an interface which the user is not necessarily attending to. We need to understand better how to design for the benefits of users – allowing maximum usage of information while distracting the users the least.

A number of design changes have been planned as results of this design walk-through. For example, we plan to graphically amplify the relationship between information and objects further using pointing arrows and animation. A follow-up study is planned to probe further into issues related to the characteristics of steerable interfaces and "in-context" interactions – mainly how people would relate their actions or object with information in this context and how design can be achieved to enhance this relationship.

References

Abowd, G.D. & Mynatt, E.D.(2000), Charting Past, Present, and Future Research in Ubiquitous Computing, *ACM Transactions on Computer-Human Interaction* 7(1), 29-58.

Bellotti, V., Back, M., Edwards, W.K., Grinter, R.E., Henderson, A. & Lopes, C., (2002), Making Sense of Sensing Systems: Five Questions for Designers and Researchers, *in CHI'02: Proceedings of the ACM Conference on Human Factors in Computing Systems.*

Bobick, A.F., Intille, S.S., Davis, J.W., Baird, F., Pinhanez, C.S., Campbell, L.W., Ivanov, Y.A., Schutte, A. and Wilson, A. (1999) The KidsRoomL A Percaptually-based Interactive and Immersive Story Environment, in *Presence: Teleooperators and Virtual Environments*, 8(4), 367-391.

Derksen USA, Inc. http://www.derksen.com

Horpraset,T., Harwood, D. & Davis, L., (1999) A Statistical Approach for Real-time Robust Background Subtraction and Shadow Detection, in *Proceedings of IEEE Frame-rate Workshop*.

Ishii, H., Wisneski, C., Orbanes, J., Chun, B. & Paradiso, J., (1999), PingPongPlus: Design of an Athletic-Tangible Interface for Computer-Supported Cooperative Play, *in CHI'99: Proceedings of the ACM Conference on Human Factors in Computing Systems.*

Jacob, R.J.K., Ishii, H., Pangaro, G. & Patten, J., (2002), A Tangible Interface for Organizing Information Using a Grid, in *CHI'02: Proceedings of the ACM Conference on Human Factors in Computing Systems.*

Kjeldsen., F.C., Pinhanez, C., Pingali, G., Hartman, J., Levas, A. & Podlaseck, M., (2002) Interacting with Steerable Projected Displays, in *FG'02, Proceedings of the fifth International Conference on Automatic Face and Gesture Recognition*.

Koike, H., Sato, Y., Kobayashi, Y., Tobita, H. & Kobayashi, M., (2000), Interactive Textbook and Interactive Venn Diagram: Natural and Intuitive Interfaces on Augmented Desk System, *in CHI'00: Proceedings of the ACM Conference on Human Factors in Computing Systems.*

Patten, J., Ishii, H., Hines, J. & Pangaro, G., (2001), Sensetable: a Wireless Object Tracking Platform for Tangible Interfaces, *in CHI'01: Proceedings of the ACM Conference on Human Factors in Computing Systems*.

Pingali, G.S., Pinhanez, C.S., Levas, A., Kjeldsen, F.C. & Podlaseck, M., (2002), User-Following Displays, *in CME'02: Proceedings of the IEEE International Conference on Multimedia and Expo*.

Pinhanez, C.S., (2001), The Everywhere Displays Projector: a Device to Create Ubiquitous Graphical Interfaces, in *Ubicomp'01, Proceedings of an International Conference on Ubiquitous Computing*.

Piper, B., Ratii, C. & Ishii, H., (2002), Illuminating Clay: a 3-D Tangible Interface for Landscape Analysis, *in CHI'2002: Proceedings of the ACM Conference on Human Factors in Computing Systems.*

Raskar, R. & Low, K., (2001), Interacting with Spatially Augmented Reality, *in AFRIGRAPH'2001: Proceedings of the First International Conference on Virtual Reality, Computer Graphics, Visualization and Interaction in Africa.*

Rekimoto, J., (2002), SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces, *in CHI'02: Proceedings of the ACM Conference on Human Factors in Computing Systems.*

Wellner, P., (1991), The DigitalDesk Calculator: Tactile Manipulation on a Desktop Display, *in UIST'91: Proceedings of the ACM Conference on User Interface Software and Technology.*